Парадигма развития науки

Методологическое обеспечение

А.Е. Кононюк

ДИСКРЕТНО-НЕПРЕРЫВНАЯ МАТЕМАТИКА

Книга 5

Матрицы

Часть 1

Киев
Освіта України
2011


Структурная схема развивающейся панмедийной системы наук
I. Группа междисциплинарных наук
ПН Теория понятий
ЛГ Теория лингвистики
ПЗ Теория познания
СЗ Теория созидания
ПТ Общая теория понятоведения
II. Группа наук познания (фундаментальных наук)
ИН Теория информации
СТ Теория систем
РС Теория распознавания
КС Теория консалтинга (консультирования)
ПР Теория принятия решения
РР Теория реализации принятых решений
III. Группа наук созидания (прикладных наук)
СС Теория социальных систем
ЖО Теория жизнеобеспечения
КМ Теория коммуникаций и транспортирования
ОБ Теория образования (обучения)
СР Теория сохранения (защиты и охраны)
РЗ Теория развития
IV. Группа общедисциплинарных наук (обеспечивающих наук)
ДМ Теория дискретно-непрерывной математики (четкая и нечеткая)
НИ Теория научных исследований
ММ Теория моделей и моделирования
ОМ Теория оптимизации

УДК 51 (075.8)
ББК В161.я7
К 213

Рецензенты: *М.К.Печурин* - д-р техн. наук, проф. (Национальный авиационный университет).

Кононюк А.Е.
К65 Дискретно-непрерывная математика. Матрицы. К.5.Ч.1.
К.4:"Освіта України", 2011. - 612 с.
ISBN 978-966-7599-50-8

Многотомная работа содержит систематическое изложение математических дисциплин, исспользуемых при моделировании и исследованиях математических моделей систем.

В работе излагаются основы теории множеств, отношений, поверхностей, пространств, алгебраических систем, матриц, графов, математической логики, теории формальных грамматик и автоматов, теории алгоритмов, которые в совокупности образуют единную методолгически взамосвязанную математическую систему «Дискретно-непрерывная математика».

Для бакалавров, специалистов, магистров, аспирантов, докторантов и просто ученых и специалистов всех специальностей.

ББК В161.я7

ISBN 978-966-7599-50-8

Оглавление

Введение

Сразу отметим, что хотя нет такой задачи, которую нельзя было бы решить без помощи матриц, представление совокупностей чисел и других объектов в виде таблиц оказалось чрезвычайно удобным и эффективным способом упорядочения информации. Это обусловило быстрое развитие матричного аппарата и его широкое применение в науке и технике. Многие выкладки и результаты без использования матриц выглядели бы слишком громоздкими и трудно обозримыми. Работа с матрицами не только экономит время, но и означает более высокий уровень математической культуры и мышление.

Теория матриц основана на простых положениях. Однако овладение матричным аппаратом требует значительных усилий и тренировки. Необходимо не только понимать смысл основных матричных соотношений, но и научиться уверенно оперировать с матрицами как объектами больше общего характера по сравнению с числами и функциями. С этих позиций и излагается материал в настоящей работе.

Вначале подробно рассматриваются основные действия над матрицами, вычисление определителей и обращение матриц. Излагаемые числовые методы служат не только средством решения этих задач. Они позволяют глубже проникнуть в сущность основных понятий и соотношений теории матриц и определителей.

Развитие матричного аппарата связано, прежде всего, с решением систем линейных алгебраических и дифференциальных уравнений. При изложении этих вопросов предполагается, что читатель знаком с основными положениями теории таких систем и методами их решения без применения матриц. Матричный аппарат позволяет представить процессы решения и исследования систем уравнений в удобной и лаконичной форме, а также построить алгоритмы для реализации этих процессов на электронных вычислительных машинах (компьютерах).

При рассмотрении дифференциальных уравнений выясняется смысл понятия экспоненциальной функции от матрицы в простейшем случае, когда все корни характеристического уравнения простые. В дальнейшем это понятие распространяется на случай кратных корней и приводится обзор методов определения аналитических функций от матриц в общем виде.

Использование матриц для решения многих прикладных задач часто сводится по существу к их преобразованиям. Систематическое рассмотрение этих вопросов позволяет выяснить взаимосвязь между различными типами таких преобразований, установить их особенности и области применения.

Последний материал посвящен рассмотрению основных соотношений в пространстве переменных состояний, значение которого сильно возросло в связи с использованием современных средств вычислительной техники для математического моделирования физических систем. Вместе с общими методами представления таких систем, затрагиваются некоторые вопросы их исследования (наблюдательность, управляемость, устойчивость).

Модуль 1
Типы матриц и действия над ними

Микромодуль 1
Введение в матрицы

1.1. Основные определения

Пусть дано некоторое числовое поле K.

Под *числовым полем* понимают любое множество чисел, в пределах которого всегда выполними и однозначны четыре операции: сложение, вычитание, умножение и деление на число, которое отлично от нуля.

Примерами числовых полей может служить множество всех рациональных чисел, множество всех действительных чисел или множество всех комплексных чисел.

Предполагается, что все встречающиеся в дальнейшем числа принадлежат данному исходному числовому полю.

Определение 1. Прямоугольную таблицу чисел из поля K

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}$$

будем называть *матрицей*

Или, другими словами, *матрица* — это множество чисел или объектов другой природы, расположенных в виде прямоугольной таблицы:

$$A = \begin{array}{|c|c|c|c|} \hline a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline a_{m1} & a_{m2} & \cdots & a_{mn} \\ \hline \end{array}\,.$$

Такая таблица, которая состоит из m строк и n столбцов, содержит mn клеток (позиций). При этом говорят, что матрица имеет *размер* $m \times n$ и ее называют *$(m \times n)$-матрицей.* Позицию на пересечении i-й строки и j-го столбца будем называть *ij-клеткой.*

Числа или любые другие объекты, которые расположены в клетках таблицы, называют *элементами* матрицы. Положение элементов строго фиксировано: в каждой клетке должен располагаться только один элемент и ни одна клетка не должна оставаться свободной.

В общем обозначении элемента a_{ij} первый индекс i всегда указывает номер строки, а второй — номер столбца. Элемент, который расположен в ij-клетке, называют *ij-элементом.*

Матрица обозначается одной буквой (часто буквы, которые обозначают матрицы, набирают жирным шрифтом или обозначают какими-нибудь дополнительными символами). Однако независимо от принятого способа обозначения **матрица всегда является совокупностью таблично упорядоченных элементов**. Две матрицы равны, если и только если равны их соответствующие элементы, т.е. $A=B$ при условии $a_{ij}=b_{ij}$ $(i = 1,2, \ldots , m;\ j = 1,2, \ldots , n)$. Ясно, что сравнивать можно только матрицы одного и того же размера, между элементами которых определено **отношение равенства**.

Матрицы, элементами которых являются вещественные или комплексные числа, называют соответственно *вещественными* или *комплексными.* Пусть A — комплексная $(m \times n)$-матрица с элементами $a_{ij}=\alpha_{ij}+i\beta_{ij}$. Матрица A того же размера с элементами $a^*_{ij}=\alpha_{ij}-i\beta_{ij}$ называется *комплексно-сопряженной* с A.

Часто для упрощения нулевые элементы в таблицу не записывают, но при этом имеют в виду, что пустые клетки также содержат числа (нули).

Кроме приведенной выше клеточной записи, используют и другие способы представления матриц, например:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}; \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}; \quad \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}.$$

Матрицы впервые появились в середине позапрошлого века в роботах английских математиков А. Кэли и У. Гамильтона. Представление совокупностей элементов в виде матриц и разработанные правила операций над ними оказались довольно плодотворными в математике и нашли широкое применение в физике, технике, экономике. Существенный вклад в разработку общей теории матриц и ее приложений внесли математики И. А. Лаппо-Данилевський, А. Н. Крылов, Ф. Р. Гантмахер, М. Г. Крейн.

Типы матриц. Матрица может иметь любое количество строк и столбцов (конечное или бесконечное). В дальнейшем при отсутствии оговорок будут рассматриваться конечные матрицы с числовыми элементами.

Если матрица состоит из одного столбца или одной строки, то она соответственно называется *столбцевой* или *строчной* (употребляются также названия *матрица-столбец* и *матрица-строка)*. В таких случаях достаточно отмечать элементы одним индексом:

$$x = \begin{Vmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{Vmatrix}; \qquad y = \begin{Vmatrix} y_1 & y_2 & \dots & y_n \end{Vmatrix}.$$

Столбцевую и строчную матрицы называют также *векторами* и сокращенно обозначают как $x = (x_1, x_2, ..., x_n)$ и $y = (y_1, y_2, \dots, y_n)$. Обычно из контекста ясно, идет ли речь о векторе-столбце или о векторе-строке. В противном случае используют приведенные высшее обозначения.

Матрица, количество строк и столбцов которой одинаково и равно n, называется *квадратной матрицей* порядка n. Совокупность ii-клеток ($i = 1, 2, ..., n$) образует *главную диагональ* квадратной матрицы. Матрица, все элементы которой вне главной диагонали равны нулю, т.е.

$$D = \left|\begin{array}{|c|c|c|c|}\hline d_1 & & & \\ \hline & d_2 & & \\ \hline & & \cdots & \\ \hline & & & d_n \\ \hline\end{array}\right|,$$

называется диагональной и кратко записывается $D=\mathrm{diag}\{d_1, d_2, \ldots, d_n\}$ или $\|d_i\delta_{ik}\|^n_1$. Здесь δ_{ik} символ Кронекера:

$$\delta_{ik} = \begin{cases} 1, (i = k), \\ 0, (i \neq k). \end{cases}$$

Если в диагональной матрице $d_1=d_2=\ldots=d_n=1$, то *имеем единичную* матрицу n-го порядка

$$E_n = \left|\begin{array}{|c|c|c|c|}\hline 1 & & & \\ \hline & 1 & & \\ \hline & & \cdots & \\ \hline & & & 1 \\ \hline\end{array}\right|,$$

которая часто обозначается также через 1_n или просто цифрой 1 (не следует вопринимать это обозначение как число, которое равно единице).

Матрица, все элементы которой равны нулю, называется *нулевой* и обозначается цифрой 0. Заметим, что нулевая матрица может иметь любой размер $m\times n$, в то время как **единичная** матрица — **всегда квадратная**. Матрица, которая состоит только из одного элемента, обычно отождествляется с этим элементом.

Квадратная матрица называется *верхней (нижней) треугольной,* если равны нулю все элементы, которые расположены под (над) главной диагональю:

$$A = \begin{array}{|c|c|c|c|} \hline a_{11} & a_{12} & \cdots & a_{1n} \\ \hline & a_{22} & \cdots & a_{2n} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline & & & a_{nn} \\ \hline \end{array}; \qquad B = \begin{array}{|c|c|c|c|} \hline b_{11} & & \cdots & \\ \hline b_{21} & b_{22} & \cdots & \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline b_{n1} & b_{n2} & \cdots & b_{nn} \\ \hline \end{array}.$$

Диагональная матрица является частным случаем как верхней *(А),* так и нижней *(В)* треугольных матриц.

1.2. Действия над матрицами

Сложение матриц.

Пусть величины $y_1, y_2, ..., y_m$ выражаются через величины $x_1, x_2, \ldots, x_n$ с помощью линейного преобразования

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, \ldots, m), \tag{1.1}$$

а величины $z_1, z_2, ..., z_m$ — через те же величины $x_1, x_2, \ldots, x_n$ с помощью преобразования

$$z_i = \sum_{k=1}^{n} b_{ik} x_k \quad (i = 1, 2, \ldots, m). \tag{1.2}$$

Тогда

$$y_i + z_i = \sum_{k=1}^{n} (a_{ik} + b_{ik})\, x_k \quad (i = 1, 2, \ldots, m). \tag{1.3}$$

В соответствии с этим мы устанавливаем:

- сумма двух прямоугольных матриц A и B одинаковых размеров определяется как матрица C тех же размеров, каждый элемент которой равен сумме соответствующих элементов матриц, т.е. $C=A+B$, если $c_{ij} = a_{ij}+ b_{ij}$.

Пример 1:

$$\begin{Vmatrix} 3 & 0 & -7 \\ 1 & 2 & 0{,}5 \end{Vmatrix} + \begin{Vmatrix} -2 & -3 & 10 \\ 0 & 0{,}5 & -0{,}5 \end{Vmatrix} = \begin{Vmatrix} 1 & -3 & 3 \\ 1 & 2{,}5 & 0 \end{Vmatrix}.$$

Пример 2:

$$\begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{Vmatrix} + \begin{Vmatrix} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{Vmatrix} = \begin{Vmatrix} a_1+c_1 & a_2+c_2 & a_3+c_3 \\ b_1+d_1 & b_2+d_2 & b_3+d_3 \end{Vmatrix}.$$

Из приведенного определения следует, что операция сложения матриц коммутативна, т.е. $A+B=B+A$, и ассоциативна, т.е. $(A+B)+C=A+(B+C)$. Она естественным образом распространяется на любое число слагаемых. Очевидно также, что матрица A не изменяется при суммировании ее с нулевой матрицей тех же размеров, т.е. $A+0 = A$.

В силу этого же определения матрица коэффициентов в преобразовании (1.3) есть сумма матриц коэффициентов в преобразованиях (1.1) и (1.2).

Умножение матрицы на число.

Умножим в преобразовании (1.1) величины $y_1, y_2, \ldots, y_m$ на некоторое число α из K. Тогда

$$\alpha y_i = \sum_{k=1}^{n} (\alpha a_{ik})\, x_k \quad (i = 1, 2, \ldots, m).$$

Согласно этому имеет место определения.

Определение. Произведением матрицы A на число α (в отличие от матриц и векторов, числа часто называют *скалярами)* является матрица $C=\alpha A$, элементы которой получаются умножением соответствующих элементов матрицы A на это число α, т.е. $c_{ij}= \alpha a_{ij}$.

Пример 1:

$$2 \cdot \begin{Vmatrix} 5 & 4 & -1 \\ 1 & 2 & 0 \end{Vmatrix} = \begin{Vmatrix} 10 & 8 & -2 \\ 2 & 4 & 0 \end{Vmatrix}.$$

Пример 2:

$$\alpha \begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{Vmatrix} = \begin{Vmatrix} \alpha a_1 & \alpha a_2 & \alpha a_3 \\ \alpha b_1 & \alpha b_2 & \alpha b_3 \end{Vmatrix}.$$

Очевидно, справедливы следующие соотношения:

$$\alpha(A + B) = \alpha A + \alpha B;$$

$$(\alpha + \beta)A = \alpha A + \beta A;$$

$$(\alpha\beta)A = \alpha(\beta A),$$

где A и B — матрицы одинакового размера; α и β - числа (скаляры). Общий множитель элементов можно выносить за знак матрицы, считая его скалярным множителем.

Разность двух матриц одинаковых размеров сводится к уже рассмотренным операциям соотношением $A - B = A + (-1)B$, т.е. $C = A - B$, если $c_{ij} = a_{ij} + b_{ij}$.

Умножение матриц.

Пусть величины $z_1, z_2, ..., z_m$ выражаются через величины $y_1, y_2, ..., y_n$ с помощью преобразования

$$z_i = \sum_{k=1}^{n} a_{ik} y_k \quad (i = 1, 2, \ldots, m), \tag{1.4}$$

а величины $y_1, ..., y_n$ выражаются через величины $x_1, x_2, ..., x_q$ с помощью формул

$$y_k = \sum_{j=1}^{q} b_{kj} x_j \quad (k = 1, 2, \ldots, n). \tag{1.5}$$

Тогда, подставляя эти выражения для y_k (k = 1, 2, ..., n) в формулы (1.4), мы выразим $z_1, z_2, ..., z_m$ через $x_1, x_2, ..., x_q$ при помощи «составного» преобразования:

$$z_i = \sum_{k=1}^{n} a_{ik} \sum_{j=1}^{q} b_{kj} x_j = \sum_{j=1}^{q} \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) x_j \quad (i = 1, 2, \ldots, m).$$

По многим соображениимям целесообразно определить операцию умножения матриц следующим образом: *произведением* матрицы A размера $(m \times n)$ на матрицу B размера $(n \times r)$ является матрица $C = AB$ размера $(m \times r)$, элемент c_{ij} которой, расположенный в ij-клетке, равен сумме произведений элементов i-й строки матрицы A на соответствующие элементы j-го столбца матрицы B, т.е.

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{1.6}$$

Умножение A на B допустимо (произведение AB существует), если число столбцов A равно числу строк B (в таких случаях говорят, что эти две матрицы *согласуются по форме).*

Пример 1:

$$\left\|\begin{array}{c|c|c|c} 2 & 0 & 3 & 1 \\ \hline 5 & 1 & 2 & 0 \\ \hline 0 & 0 & 4 & 1 \end{array}\right\| \left\|\begin{array}{c|c} 1 & 3 \\ \hline 2 & 1 \\ \hline 4 & 0 \\ \hline 3 & 5 \end{array}\right\| = \left\|\begin{array}{c|c} 2\cdot 1+0\cdot 2+3\cdot 4+1\cdot 3 & 2\cdot 3+0\cdot 1+3\cdot 0+1\cdot 5 \\ \hline 5\cdot 1+1\cdot 2+2\cdot 4+0\cdot 3 & 5\cdot 3+1\cdot 1+2\cdot 0+0\cdot 5 \\ \hline 0\cdot 1+0\cdot 2+4\cdot 4+1\cdot 3 & 0\cdot 3+0\cdot 1+4\cdot 0+1\cdot 5 \end{array}\right\| =$$

$$= \left\|\begin{array}{c|c} 17 & 11 \\ \hline 15 & 16 \\ \hline 19 & 5 \end{array}\right\|.$$

Пример 2:

$$\left\|\begin{matrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{matrix}\right\| \left\|\begin{matrix} c_1 & d_1 & e_1 & f_1 \\ c_2 & d_2 & e_2 & f_2 \\ c_3 & d_3 & e_3 & f_3 \end{matrix}\right\| =$$

$$= \left\|\begin{matrix} a_1c_1+a_2c_2+a_3c_3, & a_1d_1+a_2d_2+a_3d_3, & a_1e_1+a_2e_2+a_3e_3, & a_1f_1+a_2f_2+a_3f_3 \\ b_1c_1+b_2c_2+b_3c_3, & b_1d_1+b_2d_2+b_3d_3, & b_1e_1+b_2e_2+b_3e_3, & b_1f_1+b_2f_2+b_3f_3 \end{matrix}\right\|.$$

Для матриц $A(m\times n)$ и $B(n\times m)$ существует как произведение AB размера $m\times m$, так и произведение BA размера $n\times n$. Ясно, что при $m\neq n$ эти произведения не могут быть равными уже вследствие различных размеров результирующих матриц. Но даже при $m=n$, т.е. в случае квадратных матриц одинакового порядка, произведения AB и BA не обязательно равны между собой. Например, для матриц

$$A = \left\|\begin{array}{c|c} -1 & 2 \\ \hline 0 & 4 \end{array}\right\|; \quad B = \left\|\begin{array}{c|c} 3 & 1 \\ \hline 2 & 3 \end{array}\right\|;$$

имеем:

$$AB = \begin{vmatrix} 1 & 5 \\ 8 & 12 \end{vmatrix}; \qquad BA = \begin{vmatrix} -3 & 10 \\ -2 & 16 \end{vmatrix}.$$

Отсюда следует, что вообще операция умножения матриц не подчиняется коммутативному закону *(AB≠BA).* Если же выполняется соотношение *AB=BA,* то матрицы *A* и *B* называют *коммутирующими* или *перестановочными.* Ассоциативный и дистрибутивный законы для матричного умножения выполняются во всех случаях, когда размеры матриц допускают соответствующие операции: *(AB)C=A(BC)=ABC* (ассоциативность), *A(B+C)=AB+ACC* и *(A+B)C=ACC+BC* (дистрибутивность умножения слева и справа относительно сложения).

Умножение $(m \times n)$-матрицы A на единичную матрицу m-го порядка слева и на единичную матрицу n-го порядка справа не изменяет этой матрицы, т.е. $E_mA=AE_n=A$. Если хотя бы одна из матриц произведения *AB* является нулевой, то в результате получим нулевую матрицу.

Отметим, что из $AB=0$ не обязательно следует, что $A=0$ или $B=0$. В этом можно убедиться на следующем примере:

$$\begin{vmatrix} 4 & 1 \\ 2 & 0,5 \end{vmatrix} \begin{vmatrix} 1 & -2 \\ -4 & 8 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}.$$

Умножение всегда выролнимо, если оба сомножителя - квадратные матрицы одного и того же порядка. Обращаем внимание читателя и на то, что даже в этом частном случае умножения матриц не обладает переместительным свойством. Так, например,

$$\begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix} \begin{Vmatrix} 2 & 0 \\ 3 & -1 \end{Vmatrix} = \begin{Vmatrix} 8 & -2 \\ 18 & -4 \end{Vmatrix}, \quad \begin{Vmatrix} 2 & 0 \\ 3 & -1 \end{Vmatrix} \begin{Vmatrix} 1 & 2 \\ 3 & 4 \end{Vmatrix} = \begin{Vmatrix} 2 & 4 \\ 0 & 2 \end{Vmatrix}.$$

Пример. Матрицы

$$A = \begin{Vmatrix} 1 & 2 \\ -2 & 0 \end{Vmatrix} \text{ и } B = \begin{Vmatrix} -3 & 2 \\ -2 & -4 \end{Vmatrix}$$

перестановочны между собой, так как

$$AB = \begin{Vmatrix} -7 & -6 \\ 6 & -4 \end{Vmatrix}, \quad BA = \begin{Vmatrix} -7 & -6 \\ 6 & -4 \end{Vmatrix}.$$

Операция умножения матриц естественно распространяется на случай нескольких сомножителей.

Свойства действий над матрицами

1. $A+B=B+A$. 2. $A+(B+C)=(A+B)+C$.
3. $(\alpha+\beta)A=\alpha A+\beta A$. 4. $\alpha(A+B)=\alpha A+\alpha B$.
5. $\alpha(\beta A)=(\alpha\beta)A$. 6. $(\alpha A)B=A(\alpha B)$.
7. $(A+B)C=CA+BC$. 8. $(AB)C=A(BC)$.
9. $(A^T)^T=A$. 10. $(A+B)^T=A^T+B^T$.

Транспонирование матрицы. Преобразование матрицы *A,* которое состоит в замене строк столбцами (или столбцов строками) при сохранении их нумерации, называется *транспонированием.* Полученная в результате такого преобразования матрица называется *транспонированной* к матрице *A* и обозначается A^t или *A':*

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix}; \quad A^t = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{vmatrix}.$$

Произвольная $(m\times n)$-матрица при транспонировании становится $(n\times m)$-матрицей, а элемент a_{ij} занимает ji-клетку, т.е. $a_{ij}=a^t_{ji}$

Если матрица (квадратная) совпадает со своей транспонированной, т.е. $A=A^t$, то она называется *симметричной* и ее элементы связанны соотношениям $a_{ij}=a_{ji}$ (симметрия относительно главной диагонали). Матрица, для которой $A=-A^t$, называется *кососимметричной,* и ее элементы связаны соотношениям $a_{ij}=-a^t_{ji}$. Она, как и симметричная матрица, всегда квадратная, но диагональные элементы равны нулю, т.е. $a_{ii}=0$ ($i = 1, 2, \ldots, n$). Ниже приведены примеры симметричной и кососимметричной матриц:

2	0,5	3	—5
0,5	0	0	7
3	0	0,1	0
—5	7	0	—4

;

0	2	0,1	0
—2	0	—3	0
—0,1	3	0	—7
0	0	7	0

.

Ясно, что не все элементы таких матриц могут быть выбраны произвольно. Можно убедиться, что с n^2 элементов для симметричной матрицы независимыми могут быть только $\frac{1}{2}n(n + 1)$, а для кососимметричной — $\frac{1}{2}n(n - 1)$ элементов.

Комплексно-сопряженная и транспонированная матрица $(A)^t$ называется *сопраженной* с *A* и обозначается через *A**. Матрица, которая равна своей сопраженной, т.е. $A=(\bar{A})^t=A^*$, называется *эрмитовой*. Если $A=-(\bar{A})^t$, то *A* - *косоэрмитовая* матрица.

Легко показать, что транспонирование произведения *AB* равно произведению транспонированных матриц, взятых в обратном порядке: $(AB)^t=B^tA^t$. Дважды транспонированная матрица равна исходной, т.е. $(A^t)^t= A$.

Матричная запись системы линейных уравнений. Первоначально матрицы были введены для упрощения записи систем линейных уравнений, что обусловило и определение основных матричных операций. Система линейных уравнений:

$$\left.\begin{array}{l} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{array}\right\} \qquad (1.7)$$

записывается одним матричным равенством

$$\begin{Vmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{Vmatrix} = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{Vmatrix}.$$

Действительно, перемножив в правой части равенства $(m\times n)$-матрицу на столбцевую матрицу, получим

$$\begin{Vmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{Vmatrix} = \begin{Vmatrix} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n \end{Vmatrix}.$$

Из равенства матриц-столбцов следуют равенства для соответствующих элементов, которые совпадают с исходной системой уравнений. Если обозначить

$$y = \begin{Vmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{Vmatrix}; \quad A = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix}; \quad x = \begin{Vmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{Vmatrix},$$

это матричное равенство запишется еще короче

$$y=Ax.$$

Такое представление системы линейных уравнений оказалось возможным благодаря правилу умножения матриц, которое наилучшим образом подходит для этой цели. Однако исторически дело обстояло как раз наоборот: правила действий над матрицами определялись, прежде всего, исходя из удобства представлений систем линейных уравнений.

Линейные преобразования. Систему уравнений (1.7) можно рассматривать как линейное преобразование совокупности величин $x_1, x_2,...,x_n$ в совокупность $y_1, y_2,......,y_m$. Это преобразование полностью определяется коэффициентами a_{ij} $(i = 1, 2, ..., m;\ j = 1, 2, ..., n)$. На языке матриц линейное преобразование $y=Ax$ означает преобразование столбца x в столбец y, что определяется *матрицей преобразования А.*

Пусть величины $x_1, x_2,..., x_n$ получаются из некоторой совокупности величин $z_1, z_2, ..., z_r$ с помощью линейного преобразования $x=Bz$, где x и z — столбцы соответствующих величин; B — матрица их преобразования. Тогда формальной подстановкой x в первое матричное уравнение получаем

$$y = Ax = A(Bz) = (AB)z = Cz,$$

где $C = AB$ — матрица преобразования величин z в y. К этому же результату можно прийти путем подстановки значений $x_1, x_2 ...,x_n$ из второй системы уравнений в первую с учетом введенного ранее правила умножения прямоугольных матриц.

Равенства (1.7) выражают собой тот факт, что столбец y является линейной комбинацией столбцов матрицы A с коэффициентами $x_1, x_2, \ldots, x_n$:

$$y = x_1 a_{\cdot 1} + x_2 a_{\cdot 2} + \ldots + x_n a_{\cdot n} = \sum_{k=1}^{n} x_k a_{\cdot k}. \tag{1.8}$$

Вернемся теперь к равенствам (1.6), которые эквивалентны одному матричному равенству

$$C = AB. \tag{1.9}$$

Эти равенства могут быть записаны в виде

$$c_{\cdot j} = \sum_{k=1}^{n} b_{kj} a_{\cdot k} \quad (j = 1, 2, \ldots, q) \tag{1.10}$$

или в виде

$$c_{i\cdot} = \sum_{k=1}^{n} a_{ik} b_{k\cdot} \quad (i = 1, 2, \ldots, m). \tag{1.11}$$

Таким образом, любой j-й столбец матрицы-произведения $C=AB$ является линейной комбинацией столбцов первого сомножителя, т.е.

матрицы *A,* причем коэффициенты этой линейной зависимости образуют j-1 столбец во втором сомножителе *B.* Аналогично, любая i-я строка в матрице *C* является линейной комбинацией строк матрицы *B,* а коэффициентами этой линейной зависимости являются элементы i-й строки матрицы *A.*

Следовательно, матричное уравнение $AX= C$, где *A* и *C* — заданные соответственно $m\times n$- и $m\times q$-*матрицы*, а *X*— искомая $m\times q$-*матрица*, имеет решение в том и только в том случае, когда столбцы матрицы *C* являются линейными комбинациями столбцов матрицы *A.* Для уравнения $XB=C$ необходимое и достаточное условие существования решения *X* заключается в том, что строки матрицы *C* должны быть линейными комбинациями строк матрицы *B.*

Остановимся еще на том частном случая, когда в произведении $C=AB$ второй сомножитель является квадратной и притом диагональной матрицей $B=\{d_1, d_2, ..., d_n\}$. Тогда из формул (1.6) следует:

$$c_{ij} = d_{ij}dj \quad (i = 1, 2, ..., m; j = 1, 2, ..., n),$$

т.е.

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix} \begin{Vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{Vmatrix} = \begin{Vmatrix} a_{11}d_1 & a_{12}d_2 & \dots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \dots & a_{2n}d_n \\ \dots & \dots & \dots & \dots \\ a_{m1}d_1 & a_{m2}d_2 & \dots & a_{mn}d_n \end{Vmatrix}.$$

Аналогично

$$\begin{Vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_m \end{Vmatrix} \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix} = \begin{Vmatrix} d_1a_{11} & d_1a_{12} & \dots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \dots & d_2a_{2n} \\ \dots & \dots & \dots & \dots \\ d_ma_{m1} & d_ma_{m2} & \dots & d_ma_{mn} \end{Vmatrix}.$$

Таким образом, *при умножении прямоугольной матрицы A справа (слева) на диагональную матрицу* $\{d_1, d_2, ...\}$, *все столбцы (соответственно строки) матрицы A умножаются на числа* $d_1, d_2, \dots$

Обратная матрица. В обычной алгебре два числа, произведение которых равно единице, называют взаимно обратными. Число, которое обратное числу *a*, обозначается через a^{-1} и по определению $aa^{-1}=1$.

Аналогично в матричной алгебре две квадратные матрицы, произведение которых равно единичной матрице, т.е. $AA^{-1}=A^{-1}A=1$, называют *взаимно обратными* (A^{-1} обратна *A*). Однако дальше этого аналогия не проходит.

Выражение $a^{-1}b$, где a и b — числа, можно представить как частное от деления b на a, но для матриц такое представление не имеет смысла и в общем случае $A^{-1}B \neq BA^{-1}$. Поэтому вместо операции деления B на A различают левое частное $A^{-1}B$ и правое частное BA^{-1}, которые сводятся к умножению слева или справа на обратную матрицу A^{-1}.

Способ обращения матрицы проще установить, рассматривая решение системы n линейных уравнений с n неизвестными:

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = q_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = q_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = q_n \end{array}\right\}$$

В матричной форме эта система уравнений запишется как $Ax=q$, где A — квадратная матрица n-го порядка, называемая *матрицей системы;* x и q — столбцевые матрицы неизвестных переменных и свободных членов:

$$A = \begin{array}{|c|c|c|c|} \hline a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline a_{n1} & a_{n2} & \cdots & a_{nn} \\ \hline \end{array}; \quad x = \begin{array}{|c|} \hline x_1 \\ \hline x_2 \\ \hline \cdots \\ \hline x_n \\ \hline \end{array}; \quad q = \begin{array}{|c|} \hline q_1 \\ \hline q_2 \\ \hline \cdots \\ \hline q_n \\ \hline \end{array}.$$

Матричное уравнение $Ax=q$ решается умножением обеих его частей слева на обратную матрицу A^{-1}, т.е. $A^{-1}Ax = A^{-1}q$, в результате чего получаем $x = A^{-1}q$.

Согласно правилу Крамера неизвестные x_k $(k=1, 2, \ldots, n)$ определяются соотношением:

$$x_k = \frac{1}{\Delta}(\Delta_{1k}q_1 + \Delta_{2k}q_2 + \cdots + \Delta_{nk}q_n) = \frac{1}{\Delta}\sum_{s=1}^{n}\Delta_{sk}q_s,$$

где Δ - *определитель* системы уравнений и Δ_{sk} — *алгебраические дополнения.*

Определитель Δ представляет собой числовую функцию, которая вычисляется по определенным правилам на основании квадратной таблицы, которая состоит из коэффициентов системы уравнений

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Табличное представление определителя Δ по форме совпадает с матрицей системы уравнений, т.е. состоит из тех же элементов и в том же порядке, что и матрица *А*. В таких случаях его называют *определителем матрицы А* и записывают Δ=det *A*.

Алгебраическое дополнение Δ_{sk}, вычисляется как определитель матрицы, полученной удалением из матрицы *А* *s*-й строки и *k*-го столбца, причем этот определитель умножатеся еще на $(-1)^{s+k}$. Величину Δ_{sk} называют также *алгебраическим дополнением элемента* a_{sk} матрицы *А*. Часто определитель матрицы *А* обозначается через $|A|$, а алгебраическое дополнение - через A_{sk}.

Записав для всех элементов столбцевой матрицы *x* выражение по правилам Крамера, получим решение системы уравнений в виде:

$$\begin{Vmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{Vmatrix} = \begin{Vmatrix} \frac{1}{\Delta}(\Delta_{11}q_1+\Delta_{21}q_2+\cdots+\Delta_{n1}q_n) \\ \frac{1}{\Delta}(\Delta_{12}q_1+\Delta_{22}q_2+\cdots+\Delta_{n2}q_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{1}{\Delta}(\Delta_{1n}q_1+\Delta_{2n}q_2+\cdots+\Delta_{nn}q_n) \end{Vmatrix} =$$

$$= \frac{1}{\Delta}\begin{Vmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{Vmatrix} \begin{Vmatrix} q_1 \\ q_2 \\ \cdots \\ q_n \end{Vmatrix},$$

откуда, сравнивая с $x = A^{-1}q$, имеем

$$A^{-1} = \frac{1}{\Delta}\begin{vmatrix} \Delta_{11} & \Delta_{21} & \dots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \dots & \Delta_{n2} \\ \dots & \dots & \dots & \dots \\ \Delta_{1n} & \Delta_{2n} & \dots & \Delta_{nn} \end{vmatrix}.$$

Из полученного выражения следует правило определения обратной матрицы: 1) элементы a_{ij} данной матрицы A n-го порядка заменяют их алгебраическими дополнениями Δ_{ij}; 2)матрица алгебраических дополнений транспонируется, в результате чего получаем *присоединенную* или *взаимную матрицу* к A (она обозначается через Adj$A$); 3) вычисляется определитель Δ матрицы A и присоединенная матрица Adj$A$ умножается на величину, обратную этому определителю.

Обратная матрица существует для матрицы A при условии, что det $A\neq 0$. Такие матрицы называются *неособенными,* в отличие от *особенных (вырожденных),* определитель которых равен нулю. Ниже вычисление обратной матрицы иллюстрируется примером:

$$A = \underset{\det A = -94}{\begin{vmatrix} 2 & 1 & 0 \\ -3 & 0 & 7 \\ -5 & 4 & -1 \end{vmatrix}} \to \underset{(1)}{\begin{vmatrix} -28 & -38 & -12 \\ 1 & -2 & -13 \\ 7 & -14 & 3 \end{vmatrix}} \to$$

$$\to \underset{(2)}{\begin{vmatrix} -28 & 1 & 7 \\ -38 & -2 & -14 \\ -12 & -13 & 3 \end{vmatrix}} \to \underset{(3)}{\begin{vmatrix} \frac{14}{47} & -\frac{1}{94} & -\frac{7}{94} \\ \frac{19}{47} & \frac{1}{47} & \frac{7}{47} \\ \frac{6}{47} & \frac{13}{94} & -\frac{3}{94} \end{vmatrix}} = A^{-1}.$$

Матрица, обратная произведению двух матриц, равна переставленному произведению матриц, обратных исходным, т.е. $(AB)^{-1}=B^{-1}A^{-1}$. Действительно, умножив обе части этого равенства на AB, приходим к тождеству $E=B^{-1}A^{-1}(AB)$, так как $B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E$, где E — единичная матрица n-го порядка.

Блочные матрицы. Часто матрицу удобно разбить вертикальными и горизонтальными линиями на *блоки,* которые являются матрицами меньших размеров и при выполнении операций рассматриваются как элементы исходных матриц. Операции над *блочными матрицами* выполняются по сформулированными выше правилам при условии, что эти операции допускаются размерами соответствующих матрцц.

Пусть, например, матрицы A и B разбиты на блоки (жирными линиями) так, чтобы для соответствующих блоков имела смысл операция умножения, т.е.

$$A = \left[\begin{array}{cc|cc} 1 & -2 & 0 & 0 \\ 4 & 3 & 0 & 0 \\ \hline 3 & 0 & 2 & -1 \end{array}\right] = \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right]; \quad B = \left[\begin{array}{cc} 1 & 0 \\ 3 & -1 \\ \hline 0 & 4 \\ 2 & 1 \end{array}\right] = \left[\begin{array}{c} B_{11} \\ B_{21} \end{array}\right].$$

По правилу умножения прямоугольных матриц можно записать:

$$C = AB = \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right] \left[\begin{array}{c} B_{11} \\ B_{21} \end{array}\right] = \left[\begin{array}{c} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \end{array}\right] = \left[\begin{array}{c} C_{11} \\ C_{21} \end{array}\right].$$

Вычислим блоки C_{11} и C_{21} матрицы C:

$$C_{11} = \left[\begin{array}{cc} 1 & -2 \\ 4 & 3 \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ 3 & -1 \end{array}\right] + \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right] \left[\begin{array}{cc} 0 & 4 \\ 2 & 1 \end{array}\right] = \left[\begin{array}{cc} -5 & 2 \\ 13 & -3 \end{array}\right];$$

$$C_{21} = \left[\begin{array}{cc} 3 & 0 \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ 3 & -1 \end{array}\right] + \left[\begin{array}{cc} 2 & -1 \end{array}\right] \left[\begin{array}{cc} 0 & 4 \\ 2 & 1 \end{array}\right] = \left[\begin{array}{cc} 1 & 7 \end{array}\right].$$

В результате имеем

$$C = \left\| \begin{array}{c} C_{11} \\ \hline C_{21} \end{array} \right\| = \left\| \begin{array}{c|c} -5 & 2 \\ \hline 13 & -3 \\ \hline 1 & 7 \end{array} \right\|$$

Конечно, тот же результат получается и при непосредственном перемножении матриц.

Формула Бине — Коши. Пусть квадратная матрица $C=\|c_{ij}\|_1^m$ является произведением двух прямоугольных матриц $A=\|a_{ik}\|$ и $B=\|b_{kj}\|$ соответственно размеров $m \times n$ и $n \times m$:

$$\left\| \begin{matrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{matrix} \right\| = \left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix} \right\| \left\| \begin{matrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nm} \end{matrix} \right\|, \tag{1.12}$$

т.е.

$$c_{ij} = \sum_{\alpha=1}^{n} a_{i\alpha} b_{\alpha j} \quad (i, j = 1, 2, \dots, m). \tag{1.13}$$

Установим важную формулу Бине — Коши, которая выражает определитель $|C|$ через миноры матриц A и B:

$$\left| \begin{matrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{matrix} \right| = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} \left| \begin{matrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{matrix} \right| \left| \begin{matrix} b_{k_1 1} & \dots & b_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & b_{k_m m} \end{matrix} \right| \tag{1.14}$$

или в специальных обозначениях

$$C\begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_n \leq n} A\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \tag{1.15}$$

Согласно этой формуле *определитель матрицы С равен сумме произведений всех миноров максимального (т-го) порядка) матрицы А на соответствующие миноры того же порядка матрицы В.*

Если $m > n$, то матрицы A и B не имеют миноров m-го порядка. В этом случае правые части формул (1.14) и (1.15) следует заменить нулями.

Вывод формулы Бине — Коши. На основании формулы (1.13) определитель матрицы C можно представить в виде

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \begin{vmatrix} \sum\limits_{\alpha_1=1}^{n} a_{1\alpha_1} b_{\alpha_1 1} & \dots & \sum\limits_{\alpha_m=1}^{n} a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ \sum\limits_{\alpha_1=1}^{n} a_{m\alpha_1} b_{\alpha_1 1} & \dots & \sum\limits_{\alpha_m=1}^{n} a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix} =$$

$$= \sum_{\alpha_1,\dots,\alpha_m=1}^{n} \begin{vmatrix} a_{1\alpha_1} b_{\alpha_1 1} & \dots & a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ a_{m\alpha_1} b_{\alpha_1 1} & \dots & a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix} =$$

$$= \sum_{\alpha_1,\dots,\alpha_m=1}^{n} A\begin{pmatrix} 1 & 2 & \dots & m \\ \alpha_1 & \alpha_2 & \dots & \alpha_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m}. \tag{1.16}$$

Если $m>n$, то среди чисел $\alpha_1, \alpha_2, \dots, \alpha_m$ всегда найдутся равные между собой числа и, следовательно, каждое слагаемое в правой части равенства (1.16) будет равно нулю. Значит, в этом случае $|C| = 0$.

Пусть теперь $m \le n$. Тогда в сумме, которая стоит в правой части равенства (1.16), будут равны нулю те слагаемые, у которых хотя бы два из индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ равны между собой. Все же остальные слагаемые этой суммы можно разбить на группы по m! слагаемых в каждой, объединяя в одну группу те слагаемые, которые отличаются друг от друга только порядком индексов $\alpha_1, \alpha_2, \dots, \alpha_m$ (индексы $\alpha_1, \alpha_2, \dots, \alpha_m$ в пределах каждой группы слагаемых имеют одну и ту же совокупность значений). Тогда в пределах одной такой группы сумма соответствующих слагаемых будет равна

$$\sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) A\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} =$$

$$= A\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} \sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} =$$

$$= A\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}.$$

Поэтому из (1.16) получаем (1.15).

Здесь $k_1 < k_2 < \dots < k_m$ — нормальное расположение индексов $\alpha_1, \alpha_2, \dots, \alpha_m$, а $\varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) = (-1)^N$, где N — число транспозиций индексов, необходимых для преобразования перестановки $\alpha_1, \alpha_2, \dots, \alpha_m$ к нормальному расположению $k_1 < k_2 < \dots < k_m$.

Пример 1.

$$\begin{Vmatrix} a_1c_1 + a_2c_2 + \dots + a_nc_n & a_1d_1 + a_2d_2 + \dots + a_nd_n \\ b_1c_1 + b_2c_2 + \dots + b_nc_n & b_1d_1 + b_2d_2 + \dots + b_nd_n \end{Vmatrix} = \begin{Vmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{Vmatrix} \begin{Vmatrix} c_1 & d_1 \\ c_2 & d_2 \\ \vdots & \vdots \\ c_n & d_n \end{Vmatrix}.$$

Поэтому формула (1.15) дает так называемое тождество Коши

$$\begin{vmatrix} a_1c_1+a_2c_2+\ldots+a_nc_n & a_1d_1+a_2d_2+\ldots+a_nd_n \\ b_1c_1+b_2c_2+\ldots+b_nc_n & b_1d_1+b_2d_2+\ldots+b_nd_n \end{vmatrix} = \sum_{1\le i<k\le n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix} \begin{vmatrix} c_i & d_i \\ c_k & d_k \end{vmatrix}.$$

Полагая в этом тождестве $a_i=c_i$, $b_i=d_i$ $(i=1, 2, ..., n)$, получим:

$$\begin{vmatrix} a_1^2+a_2^2+\ldots+a_n^2 & a_1b_1+a_2b_2+\ldots+a_nb_n \\ a_1b_1+a_2b_2+\ldots+a_nb_n & b_1^2+b_2^2+\ldots+b_n^2 \end{vmatrix} = \sum_{1\le i<k\le n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix}^2.$$

В случае, когда a_i и b_i $(i = 1,2, ..., n)$ - вещественные числа, отсюда следует известное неравенство

$$(a_1b_1+a_2b_2+\ldots+a_nb_n)^2 \le (a_1^2+a_2^2+\ldots+a_n^2)(b_1^2+b_2^2+\ldots+b_n^2).$$

При этом знак равенства имеет место тогда и только тогда, когда все числа a_i пропорциональны соответствующим числам b_i $(i = 1,2, ..., n)$.

Пример 2.

$$\left\| \begin{matrix} a_1c_1+b_1d_1 \ldots a_1c_n+b_1d_n \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ a_nc_1+b_nd_1 \ldots a_nc_n+b_nd_n \end{matrix} \right\| = \left\| \begin{matrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{matrix} \right\| \left\| \begin{matrix} c_1 \ldots c_n \\ d_1 \ldots d_n \end{matrix} \right\|.$$

Поэтому при $n > 2$

$$\begin{vmatrix} a_1c_1+b_1d_1 \ldots a_1c_n+b_1d_n \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ a_nc_1+b_nd_1 \ldots a_nc_n+b_nd_n \end{vmatrix} = 0.$$

Рассмотрим частный случай, когда A и B — квадратные матрицы одного и тот же порядка n, и положим в (1.15) $m=n$. Тогда приходим к известной теореме об умножении определителей:

$$C\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} = A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} B\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix},$$

или в других обозначениях:

$$|\,C\,| = |\,AB\,| = |\,A\,| \cdot |\,B\,|. \qquad (1.17)$$

Таким образом, *определитель произведения двух квадратных матриц равен произведению определителей перемножаемых матриц.*

Формула Бине — Коши дает возможность в самом общем случае выразить миноры произведения двух прямоугольных матриц через миноры сомножителей. Пусть

$$A = \|\,a_{ik}\,\|,\ B = \|\,b_{kj}\,\|,\ C = \|\,c_{ij}\,\|$$

$$(i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n;\ j = 1, 2, \ldots, q)$$

и

$$C = AB.$$

Рассмотрим произвольный минор матрицы C:

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \qquad \left(\begin{matrix} 1 \leqslant i_1 < i_2 < \dots < i_p \leqslant m \\ 1 \leqslant j_1 < j_2 < \dots < j_p \leqslant q \end{matrix};\ p \leqslant m, q\right).$$

Матрица, которая составлена из элементов этого минора, представляет собой произведение двух прямоугольных матриц

$$\begin{Vmatrix} a_{i_1 1} & a_{i_1 2} & \dots & a_{i_1 n} \\ \dots & \dots & \dots & \dots \\ a_{i_p 1} & a_{i_p 2} & \dots & a_{i_p n} \end{Vmatrix}, \qquad \begin{Vmatrix} b_{1 j_1} & \dots & b_{1 j_p} \\ b_{2 j_1} & \dots & b_{2 j_p} \\ \dots & \dots & \dots \\ b_{n j_1} & & b_{n j_p} \end{Vmatrix}.$$

Поэтому, применяя формулу Бине — Коши, получаем:

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = \sum_{1 \leqslant k_1 < k_2 < \dots < k_p \leqslant n} A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \tag{1.18}$$

Из той же формулы Бине — Коши следует, что миноры p-го порядка матрицы C при $p>n$ (если миноры таких порядков существуют) все равны нулю. В этом случае правую часть формулы (1.18) следует заменить нулем.

При p=1 формула (1.18) переходит в формулу (1.6). При p>1 формула (1.18) является естественным обобщением формулы (1.6).

Отметим еще одно следствие из формулы (1.18).

Ранг произведения двух прямоугольных матриц не превосходит ранга любого из сомножителей.

Если $C = AB$ и r_A, r_B, r_c — ранги матриц A, B, C, то

$$r_c \leq r_A, r_B$$

Если X — решение матричного уравнения $AX=C$ (размеры матриц A, X и C соответственно $m \times n$, $n \times q$ и $m \times q$), то $r_X \geq r_C$. Покажем, что *среди решений матричного уравнения $AX=C$ существует решение X_0 минимального ранга, для которого $r_{X_0} = r_C$.*

Действительно, пусть $r=r_C$. Тогда среди столбцов матрицы C имеется r линейно независмых. (Мы ссылаемся на положение: ***<u>ранг матрицы</u> равен числу линейно независимых столбцов (строк) матрицы***. Доказательство этого положения будет приведено позже).

Пусть для конкретности первые r столбцов $C_{.1}, \dots, C_{.r}$ линейно независимы, а другие столбцы $C_{.r+1}, \dots, C_{.q}$ являются линейными комбинациями первых r:

$$C_{\cdot j}=\sum_{k=1}^{r}\alpha_{jk}C_{\cdot k}\quad (j=r+1,\ldots,q). \tag{1.19}$$

Пусть X — произвольное решение уравнения $AX=C$.

$$AX_{\cdot k}=C_{\cdot k}\qquad (k=1,\ldots,r). \tag{1.20}$$

Определим столбцы $\widetilde{X}_{\cdot r+1},\ldots,\widetilde{X}_{\cdot q}$ равенствами

$$\widetilde{X}_{\cdot j}=\sum_{k=1}^{r}\alpha_{jk}X_{\cdot k}\quad (j=r+1,\ldots,q).$$

Умножая эти равенства слева почленно на A, в силу равенств (1.19) и (1.20) находим:

$$A\widetilde{X}_{\cdot j}=C_{\cdot j}\qquad (j=r+1,\ldots,q). \tag{1.21}$$

Система из q равенств (1.20) и (1.21) эквивалентна одному матричному равенству

$$AX_0=C,$$

где $X_a=(X_{\cdot 1\ldots},X_{\cdot r},\ \widetilde{X}_{\cdot r+1},\ldots,\ \widetilde{X}_{\cdot q})$ — матрица ранга r.

В матрице X_0 последние n-r столбцов являются линейными комбинациями первых r; первые же r столбцов $X_{\cdot 1}\ \ldots,\ X_{\cdot r}$ линейно независимы, так как линейная зависимость между этими столбцами в силу равенств (1.20) послужила бы причиной линейной зависимости между столбцами $C_{\cdot 1}\ \ldots,\ C_{\cdot r}$.

Решение X_0 минимального ранга r_C матричного уравнения $AX=C_0$ всегда представимо в виде

$$X_0=VC,$$

где V — некоторая $m\times n$ -матрица.

Действительно, из равенства $AX_0=C$ следует, что строки матрицы C являются линейными комбинациями строк матрицы X_0. Поскольку как среди строк матрицы C, так и среди строк матрицы X_0 имеется одно и то же число r_C линейно независимых, то и, обратно, строки матрицы X_0 являются линейными комбинациями строк матрицы C, а отсюда уже следует равенство $X_0=VC$.

Докажем теперь следующее предложение.

Матричное уравнение

$$AXB=C, \tag{1.22}$$

где A, B — заданы, а X — искомая прямоугольная матрица (предполагается, что размеры матриц A, X, B, C таковы, что произведение AXB имеет смысл и имеет размеры матрицы C), *имеет решение в том и только в том случае, когда одновременно имеют решение матричные уравнения*

$$AY=C,\qquad ZB=C, \tag{1.23}$$

т.е. когда столбцы матрицы С являются линейными комбинациями столбцов матрицы А, а строки матрицы С являются линейными комбинациями строк матрицы В.

В самом деле, если матрица X — решение уравнения (1.22), то матрицы $Y = XB$ и $Z = AX$ являются решениями уравнений (1.23).

Обратно, пусть существуют решения Y, Z уравнений (1.23). Тогда первое из этих уравнений имеет решение Y_0 минимального ранга r_C, которое по доказанному представимо в виде

$$Y_0 = VC.$$

Поэтому

$$C= AY_0 = AVC= AVZB.$$

Тогда матрица $X= VZ$ будет решением уравнения (1.22).

Ранг матрицы. Пусть задана матрица A размером $m\times n$. Выделим в матрице A любые k строк и столько же столбиков, где число k не больше чисел m и n.

Определитель порядка k, составленный из элементов, которые стоят на пересечении выделенных строк и столбцов, называют *минором k-го порядка матрицы А*.

Рангом матрицы А называют наибольший порядок отличного от нуля минора этой матрицы и обозначают $r(A)$.

Ранг матрицы находится в пределах $0 \le r(A) \le \min(m, n)$.

Например, матрица

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 4 & 2 & 1 \\ 2 & -4 & 0 \end{pmatrix}$$

имеет такие миноры второго порядка:

$$\begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix}, \begin{vmatrix} 1 & -1 \\ 4 & 1 \end{vmatrix}, \begin{vmatrix} 3 & -1 \\ -4 & 0 \end{vmatrix}.$$

Минором третьего порядка данной матрицы является ее определитель.

Ранг матрищ можно находить так. Если в матрице указан отличный от нуля минор k — го порядка, то ранг матрицы не меньше k. При этом, если все миноры $(k+1)$ - го порядка равны нулю, то ранг матрицы равен k. Если встретился ненулевой минор $(k+1)$-го порядка, то переходим к исследованию миноров порядка $k+2$, т.е. процедура продолжается.

На практике для нахождения ранга высоких порядков удобнее использовать другой метод, который базируется на следующем утверждении.

Ранг матрицы не изменится, если над матрицей выполнить *элементарные преобразования,* а именно:
1) переставить местами две строки (столбца);
2) умножить каждый элемент строки (столбца) на ненулевой множитель;
3) прибавить к элементам строки (столбца) соответственно элементы другой строки (столбца), умноженные на одно и то же число.

Воспользовавшись элементарными преобразованиями, матрицу можно привести к виду, когда все ее элементы, кроме элементов $a_{11}, a_{22},...,a_{rr}$, где $r \le \min(m,n)$, равны нулю. Тогда ранг матрицы равен r.

Микромодуль 1
Примеры решения типовых задач

1. Найдите произведение матриц AB, если

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 4 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 0 & 1 \\ 5 & 1 & -1 \\ 4 & 3 & 2 \end{pmatrix}$$

Решение. Поскольку количество столбцов матрицы A равно количеству строк матрицы B, т.е. матрицы согласованы, то операция умножения AB имеет смысл и произведение матриц вычисляем так:

$$AB = \begin{pmatrix} 1\cdot(-2)+3\cdot5+2\cdot4 & 1\cdot0+3\cdot1+2\cdot3 & 1\cdot1+3\cdot(-1)+2\cdot2 \\ 4\cdot(-2)+1\cdot5+3\cdot4 & 4\cdot0+1\cdot1+3\cdot3 & 4\cdot1+1\cdot(-1)+3\cdot2 \end{pmatrix} = \begin{pmatrix} 21 & 9 & 2 \\ 9 & 10 & 9 \end{pmatrix}$$

2. Найдите $f(A)$, если

$$A = \begin{pmatrix} 2 & -3 \\ 4 & 5 \end{pmatrix}, \quad f(x)=(x^2-3x)(3x+2)$$

Решение. Необходимо найти значение выражения

$$f(A)=(A^2-3A)\cdot(3A-2E)$$

Имеем

$$A^2-3A = \begin{pmatrix} -8 & -21 \\ 28 & 13 \end{pmatrix} - 3\begin{pmatrix} 2 & -3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} -8 & -21 \\ 28 & 13 \end{pmatrix} - \begin{pmatrix} 6 & -9 \\ 12 & 15 \end{pmatrix} = \begin{pmatrix} -14 & -12 \\ 16 & -2 \end{pmatrix};$$

$$3A+2E = \begin{pmatrix} 2 & -3 \\ 4 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 4 & 5 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 4 & 7 \end{pmatrix};$$

$$f(A) = \begin{pmatrix} -14 & -12 \\ 16 & -2 \end{pmatrix}\begin{pmatrix} 4 & -3 \\ 4 & 7 \end{pmatrix} = \begin{pmatrix} -14\cdot4-12\cdot4 & -14\cdot(-3)-12\cdot7 \\ 6\cdot4-2\cdot4 & 16\cdot(-3)-2\cdot7 \end{pmatrix} = \begin{pmatrix} -104 & -42 \\ 16 & -62 \end{pmatrix}.$$

Замечание. Если $f(x) = a_0 x^n + a_1 x^{n-1} + a_{n-1} x + a_n$ и A – некоторая квадратная матрица, то

$$f(A) = a_0 A^n + a_1 A^{n-1} + a_{n-1} A \ + a_n E.$$

3. Найдите обратную к матрице A матрицу, если

$$A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 3 & 1 \\ 2 & -1 & 0 \end{pmatrix}$$

Решение. В первую очередь вычисляем определитель матрицы A:

$$\det(A) = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 3 & 1 \\ 2 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 5 & 0 \\ 2 & -1 & 0 \end{pmatrix} = 10 \neq 0.$$

Поскольку A - невырожденная матрица, то обратная матрица существует. Находим алгебраические дополнения:

$$A_{11} = (-1)^{1+1}\begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix} = 1; \quad A_{21} = (-1)^{2+1}\begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix} = 1; \quad A_{31} = (-1)^{3+1}\begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} = 5;$$

$$A_{12} = \begin{pmatrix} -1 & 1 \\ 2 & 0 \end{pmatrix} = 2; \quad A_{22} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} = 2; \quad A_{32} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = 0; \quad A_{13} = \begin{pmatrix} -1 & 3 \\ 2 & -1 \end{pmatrix} = -5;$$

$$A_{23} = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} = 5; \quad A_{33} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = 5;$$

Следовательно, обратная матрица имеет вид

$$A^{-1} = \frac{1}{10}\begin{pmatrix} 1 & 1 & 5 \\ 2 & 2 & 0 \\ -5 & 5 & 5 \end{pmatrix} = \begin{pmatrix} 0{,}1 & 0{,}1 & 0{,}5 \\ 0{,}2 & 0{,}2 & 0 \\ -0{,}5 & 0{,}5 & 0{,}5 \end{pmatrix}$$

4. Решите матричное уравнение $X \cdot A \cdot B = C$, если

$$A = \begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix}, \ B = \begin{pmatrix} -1 & 4 \\ 2 & -7 \end{pmatrix}, \ C=(1,-2).$$

Решение. Последовательно получаем

$$X\cdot A\cdot B=C,\ X\cdot A\cdot B\,B^{-1}=CB^{-1},\ X\cdot A\cdot E=CB^{-1},\ X\cdot A=CB^{-1},$$
$$X\cdot AA^{-1}=CB^{-1}A^{-1},\ X\cdot E=CB^{-1}A^{-1},\ X=CB^{-1}A^{-1}.$$

Находим обратные матрицы A^{-1} и B^{-1}:

$$\Delta(A)=\begin{vmatrix}2 & 2\\ 3 & 1\end{vmatrix}=-4\neq 0,\quad A_{11}=1,\quad A_{21}=-2,\quad A_{12}=-3,\quad A_{22}=2,$$

$$A^{-1}=-\frac{1}{4}\begin{pmatrix}1 & -2\\ -3 & 2\end{pmatrix};$$

$$\Delta(B)=\begin{vmatrix}-1 & 4\\ 2 & -7\end{vmatrix}=-1,\quad B_{11}=-7,\quad B_{21}=-4,\quad B_{12}=-2,\quad B_{22}=-1,$$

$$B^{-1}=-\begin{pmatrix}-7 & -4\\ -2 & -1\end{pmatrix}=\begin{pmatrix}7 & 4\\ 2 & 1\end{pmatrix}.$$

Тогда

$$\mathrm{X}=\mathrm{CB}^{-1}\mathrm{A}^{-1}=(1-2)\begin{pmatrix}7 & 4\\ 2 & 1\end{pmatrix}\left(-\frac{1}{4}\right)\begin{pmatrix}1 & -2\\ -3 & 2\end{pmatrix}=$$

$$=-\frac{1}{4}(1\cdot 7-2\cdot 2\quad 1\cdot 4-2\cdot 1)\begin{pmatrix}1 & -2\\ -3 & 2\end{pmatrix}=-\frac{1}{4}(3\quad 2)\begin{pmatrix}1 & -2\\ -3 & 2\end{pmatrix}=$$

$$=-\frac{1}{4}(3\cdot 1+2\cdot(-3)\quad 3\cdot(-2)+2\cdot 2)=-\frac{1}{4}(-3\quad -2)=\begin{pmatrix}\frac{3}{4} & \frac{1}{2}\end{pmatrix}.$$

5. Найдите ранг матрицы
$$A=\begin{pmatrix}1 & -2 & 0 & 4\\ -1 & 3 & -1 & 2\\ 0 & 1 & -1 & 6\end{pmatrix}$$

Решение. Выделим в матрице минор второго порядка

$$M^{(2)}=\begin{vmatrix}1 & -2\\ -1 & 3\end{vmatrix}=1\neq 0$$

Обводными для него минорами третьего порядка являются:

$$\begin{vmatrix}1 & -2 & 0\\ -1 & 3 & -1\\ 0 & 1 & -1\end{vmatrix}\quad i\quad \begin{vmatrix}1 & -2 & 4\\ -1 & 3 & 2\\ 0 & 1 & 6\end{vmatrix}$$

Оба минора третьего порядка равны нулю, а минор второго порядка не равен нулю, следовательно, *r(A)*=2.

6. Найдите ранг матрицы
$$A=\begin{pmatrix}2 & -2 & 1 & -1 & 3\\ -4 & -4 & 2 & -4 & -2\\ 3 & 0 & 0 & 2 & 3\\ -2 & -6 & 3 & -5 & 1\end{pmatrix}$$

Решение. Выполнив элементарные преобразования, получим

$$\begin{pmatrix} 2 & -2 & 1 & -1 & 3 \\ -4 & -4 & 2 & -4 & -2 \\ 3 & 0 & 0 & 2 & 3 \\ -2 & -6 & 3 & -5 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 1 & -1 & 3 \\ -4 & 0 & 2 & -4 & -2 \\ 3 & 0 & 0 & 2 & 3 \\ -2 & 0 & 3 & -5 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 3 & -1 & 3 \\ -4 & 0 & -2 & -4 & -2 \\ 3 & 0 & 3 & 2 & 3 \\ -2 & 0 & 1 & -5 & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 2 & -4 & -2 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 3 & -5 & 1 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & -2 & -8 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 0 & -2 & -8 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & -2 & -8 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & -2 & -8 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Определитель третьего порядка, составленный из элементов, которые стоят на пересечении первых трех строк и столбцов последней матрицы, не равны нулю, а все миноры четвертого порядка равны нулю. Следовательно, $r(A)$=3.

Микромодуль 1
Примеры тестовых задач

1. Любая матрица является прямоугольной таблицей. Справедливо ли обратное утверждение, т.е. можно ли считать любую прямоугольную таблицу матрицей? Если нет, то какие дополнительные требования предъявляются с позиций матричной алгебры?

2. Какие из приведенных ниже совокупностей объектов представляют собой матрицы:

$$\text{а)} \begin{bmatrix} 4 & 2 & 7 \\ 0 & 3 & -1 \\ 1 & -5 & 0 \end{bmatrix}; \quad \text{б)} \begin{bmatrix} \sin x \\ x \\ 0 \end{bmatrix}; \quad \text{в)} \begin{bmatrix} & 12 & \\ 0 & & 3 \\ 2 & -1 & \end{bmatrix}; \quad \text{г)} \begin{matrix} & 1 & & \\ 1 & 1 & & \\ 1 & 2 & 1 & \\ 1 & 3 & 3 & 1 \end{matrix}.$$

3. Укажите, какие из приведенных ниже матриц являются равными между собой (при x = 2):

$$A = \begin{bmatrix} x^2+1 & 2 \\ 2^x & (x-1)^3 \end{bmatrix}; \quad B = \begin{bmatrix} 5 & 2x-1 \\ 3x-2 & 1 \end{bmatrix}; \quad C = \begin{bmatrix} 2^x+1 & x \\ (4-x)^2 & 1 \end{bmatrix}.$$

4. При каком значении x матрицы A и B равны:

$$A = \begin{bmatrix} (x-3)^2 & 3 \\ x^2-1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 4 & 2x+1 \\ (x-1)(x+1) & 0 \end{bmatrix}.$$

5. Найти сумму $A+B$ и разность $A - B$ матриц:

$$A = \begin{bmatrix} 3 & 4 & -3 & 5 \\ 0 & 1 & 2 & -1 \\ -2 & 5 & 0 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 4 & 0 & 2 & 3 \\ 2 & 3 & 1 & -2 \\ 2 & -5 & 7 & 0 \end{bmatrix}.$$

6. Найти произведение *AB* и *BA* и сравнить полученные результаты для матриц:

а) $A = \begin{bmatrix} -2 & 3 \\ 1 & 4 \end{bmatrix}$; $B = \begin{bmatrix} 5 & 2 \\ -3 & 0 \end{bmatrix}$.

б) $A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$; $B = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$.

в) $A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -2 \end{bmatrix}$; $B = \begin{bmatrix} 5 & -3 \\ 1 & 2 \\ 0 & -1 \end{bmatrix}$.

7. Проверить дистрибутивность умножения слева *A (B + C) = AB+ACC* и справа *(A + B) C = ACC + BC* относительно сложения для следующих матриц:

$$A = \begin{bmatrix} 4 & 3 \\ -2 & 1 \end{bmatrix};\quad B = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix};\quad C = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}.$$

8. Найти все матрицы, которые являются перестановочными с матрицей

$$\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}.$$

9. Каким условиям в общем случае должны удовлетворять элементы квадратных матриц *A* и *B* второго порядка, чтобы они были перестановочными *(AB=BA)?* Как выглядят эти условия для случая, когда *A* — симметричная матрица?

10. При каких условиях справедливы матричные соотношения:

$$(A + B)^2 = A^2 + 2AB + B^2;\ (A - B)(A + B) = A^2 - B^2?$$

11. Каким условиям должны удовлетворять элементы ненулевых квадратных матриц *A* и *B,* чтобы $AB = 0$?

12. К каким типам относятся матрицы:

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 0 & 3 & 4 \end{bmatrix};\ B = \begin{bmatrix} 0 & 2 & -1 \\ -2 & 0 & 4 \\ 1 & -4 & 0 \end{bmatrix};\ C = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}?$$

13. Построить транспонированную A^t, комплексно-сопряженную $\bar{A}$ и сопряженную A^* для матрицы

$$A = \begin{bmatrix} 1+2i & 4 & 2-3i \\ -i & 2 & 4+2i \\ 5-3i & 1 & -6i \end{bmatrix}.$$

14. Показать, что матрица

$$A = \begin{bmatrix} 4 & 1+i & 2+3i \\ 1-i & 2 & 3-4i \\ 2-3i & 3+4i & 3 \end{bmatrix}$$

является эрмитовой. Что можно сказать о диагональных элементах любой эрмитовой матрицы?

15. Какого типа должна быть квадратная матрица *A,* чтобы она была перестановочной с диагональной матрицей *D* того же порядка, т.е. чтобы $AD = DA$?

16. К какому типу относятся треугольные матрицы, если они кроме того: а) симметричны, б) кососимметричны?

17. Показать, что $(AB) = AB$ и $(AB)^* = B^*A^*$.

18. Проверить соотношение $(AB)^* = B^*A^*$ для матриц задачи 17.

19. Показать, что произведение AA^t существует для любой матрицы A и является симметричной матрицей.

20. Для заданных матриц найти обратную и проверить соотношение $AA^{-1} = 1$:

$$\text{а)}\ \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{bmatrix}; \quad \text{б)}\ \begin{bmatrix} 1 & 2 & 1 \\ 4 & 3 & -2 \\ -5 & -4 & -1 \end{bmatrix}.$$

21. Найти матрицы, обратные заданным, и проверить соотношение $(AB)^{-1} = B^{-1}A^{-1}$;

$$A = \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}; \quad B = \begin{bmatrix} 3 & 5 \\ 7 & -9 \end{bmatrix}.$$

22. Дана система уравнений:

$$\left.\begin{aligned} 2x_1 - 3x_2 + x_3 &= -1 \\ x_1 + x_2 + x_3 &= 6 \\ 3x_1 + x_2 - 2x_3 &= -1 \end{aligned}\right\}.$$

Записать эту систему в матричной форме $Ax=q$, вычислить обратную матрицу A^{-1} и записать решение системы.

23. Зависимости между токами и напряжениями четырехполюсника (рис. 1.1, *а*) можно представить одной из систем уравнений:

$$\left.\begin{aligned} U_1 &= a_{11}U_2 + a_{12}I_2 \\ I_1 &= a_{21}U_2 + a_{22}I_2 \end{aligned}\right\}; \quad \left.\begin{aligned} I_1 &= y_{11}U_1 + y_{12}U_2 \\ I_2 &= y_{21}U_1 + y_{22}U_2 \end{aligned}\right\}.$$

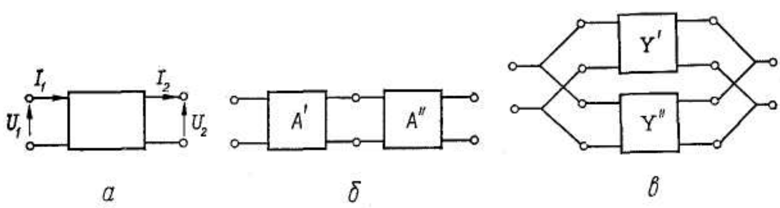


Рис. 1.1. Соединение четырехполюсника:
а - четырехполюсник; *б* - последовательное соединение; *в* - параллельное соединение.

а) Записать эти уравнения в матричной форме и установить зависимости между элементами матриц:

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix};\quad Y=\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}.$$

б) Показать, что матрица A последовательного соединения четырехполюсников (рис. 1.1, *б)* равна произведению их матриц A' и A'', т.е. $A = A'A''$ (в порядке следования).

в) Показать, что матрица Y параллельного соединения четырехполюсников (рис. 1.1, *в)* равна сумме их матриц Y' и Y'', т.е. $Y = Y' + Y''$.

24. Выполнить умножение матриц, воспользовавшись разбивкой их на блоки:

$$\left[\begin{array}{cc|c} 3 & 1 & 4 \\ -2 & -1 & 3 \\ \hline 0 & 0 & 5 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ -2 & 3 & 0 & 0 \\ \hline 1 & -1 & 4 & 1 \end{array}\right].$$

Проверить результат непосредственным умножением матриц.

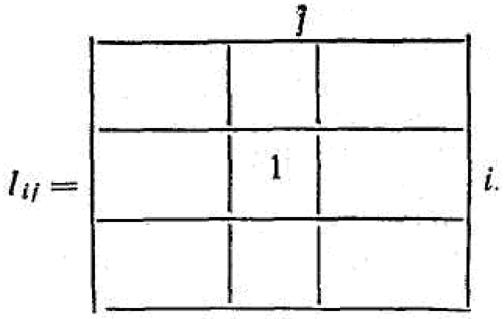
<u>**Микромодуль 2**</u>
Алгебра матриц

1. Матрицы как объекты алгебраических систем. Если элементами матриц являются числа, то говорят о **множестве матриц над числовым полем**. **Такое множество в совокупности с определенными на нем операциями и составляет <u>матричную алгебру</u>**. Основные операции над матрицами определены в микромодуле 1. В дальнейшем будут рассмотрены некоторые вопросы, которые связаны с техникой их выполнения. Здесь же уместно обратить внимание на матрицы как объекты алгебраических систем.

Множество всех прямоугольных матриц одинакового размера $m\times n$ образуют линейное пространство над числовым полем с внутренней операцией - сложением матриц и внешней операцией — умножением матрицы на число. Каждая $(m\times n)$-матрица A из этого множества может быть единственным образом представлена в виде:

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} I_{ij},$$

где I_{ij} *-матрица* размера $m\times n$, все элементы которой равны нулю, кроме ij-элемента, равного единице, т.е.,



Очевидно, матрицы I_{ij} $(i = 1, 2, ..., m; j=1, 2, ..., n)$ образуют базис линейного пространства, причем его размерность равна mn.

Множество всех квадратных матриц порядка n с операциями сложения и умножения можно рассматривать как унитарное кольцо. Роль единицы играет единичная матрица E_n (нейтральный элемент относительно умножения). Кольцо имеет делители нуля, так как при $A\neq 0$ и $B\neq 0$ может иметь место $AB=0$, но оно не коммутативно в силу того, что операция умножения матриц не подчиняется коммутативному закону. Согласно определениям, данным раньше, множество квадратных матриц n-го порядка является также некоммутативной линейной алгеброй над числовым полем ранга n^2.

Множество всех неособенных квадратных матриц данного порядка над числовым полем образуют относительно операции умножения мультипликативную группу, в которой нейтральным элементом является единичная матрица, а симметричным элементом любой матрицы из этого множества - обратная матрица.

2. Схема умножения матриц. При выполнении операции умножения по правилу, изложенному в микромодуле 1, удобно пользоваться схемой (рис. 1.2), которая иллюстрируется следующим примером:

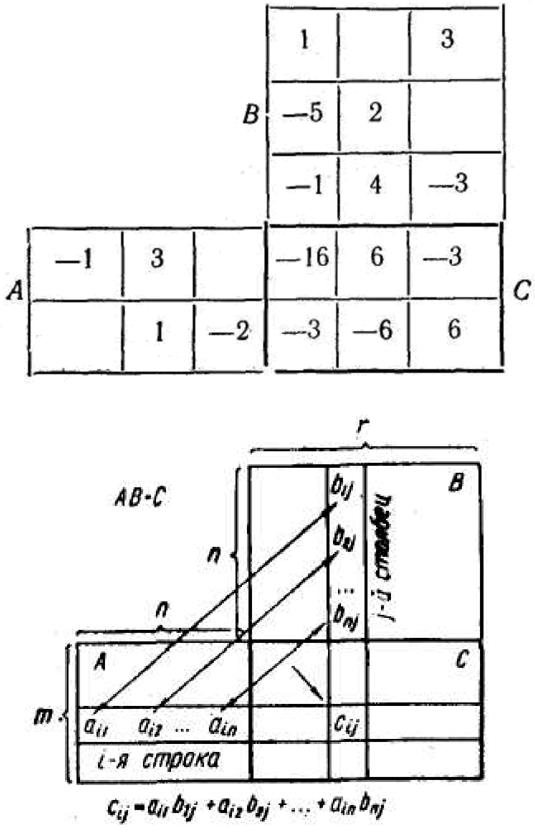


Рис. 1.2. Схема умножения матриц

На рис. 1.3 показана схемы для типичных случаев умножения матриц. В результате умножения двух матриц получаем также матрицу (рис. 1.3, *а).* Произведение матрицы на столбец дает столбец (рис. 1.3, *б),* а строки на матрицу — строку (рис. 1.3, *в).* При умножении строки на столбец получаем матрицу с единичным элементом (рис. 1.3, *г),* которая отождествляется с этим элементом. Наконец, в результате умножения столбца на строку получаем матрицу (рис. 1.3,*д)*

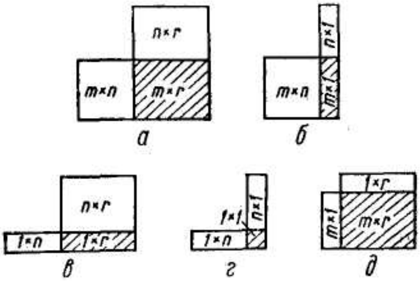


Рис. 1.3. Схема умножения:
а - матриц; *б* - матрицы на столбец;
в - строки на матрицу; *г* — строки на столбец;
д - столбца на строку.

В последнем случае: строки и столбцы результирующей матрицы пропорциональны, так как по правилу матричного произведения

$$\begin{vmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1r} \end{vmatrix} = \begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1r} \\ a_{21}b_{11} & a_{21}b_{12} & \cdots & a_{21}b_{1r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \cdots & a_{m1}b_{1r} \end{vmatrix}.$$

3. Ассоциативность матричного произведения. При перемножении нескольких матриц можно воспользоваться ассоциативностью этой операции для выбора наиболее целесообразной схемы выполнения. Пусть, например, необходимо найти произведение четырех матриц *ABCD.* Возможны следующие пути решения этой задачи.

Умножение слева: сначала *А* умножается на *В,* а затем произведение *АВ* — на *С,* наконец, *ABC* — на *D,* что записывается как *((AB)C)D* и иллюстрируется схемой на рис. 1.4.

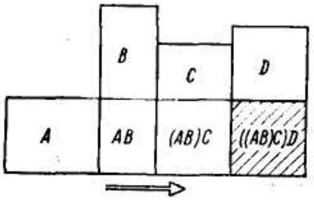


Рис. 1.4. Схема умножения матриц слева.

Этот способ удобен, когда первая матрица *A* имеет сравнительно небольшое количество строк (особенно, если она является строчной матрицей).

Умножение справа: предпоследняя матрица *C* умножается на *D*, затем *B* — на произведение *CD* и, наконец *A* — на *BCD*, что записывается как *A(B(CD))* и иллюстрируется схемой на рис. 1.5.

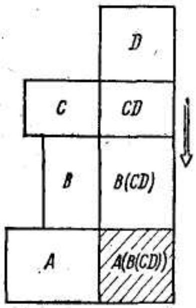


Рис. 1.5. Схема умножения матриц справа.

Этот способ удобен, когда последняя матрица *D* имеет сравнительно небольшое число столбцов (особенно, если она является столбцевой матрицей).

Комбинирование сомножителей: иногда удобно разбить произведение на группы сомножителей, не изменяя их порядка следования, и выполнить умножение сначала по группам, а потом перемножить полученные результаты (слева или справа). Так, для рассмотренной цепочки матриц этот путь приводит к следующим схемам:

$$ABCD = A((BC)D) = (A(BC))D = (AB)(CD).$$

Приведенные схемы естественно обобщаются на произведение любого количества матриц, размеры которых допускают эту операцию. Пусть, например, требуется вычислить произведение *ABCDE*, где

$$A=\begin{bmatrix}4 & 0 & 1 & -2\\ 1 & 3 & -2 & 0\\ 0 & -5 & 4 & 3\end{bmatrix};\quad B=\begin{bmatrix}1\\0\\-1\\3\end{bmatrix};\quad C=[-3,\ 1];$$

$$D=\begin{bmatrix}-1 & 3 & -2\\ 1 & 0 & 4\end{bmatrix};\quad E=\begin{bmatrix}5\\0\\-3\end{bmatrix}.$$

Воспользуемся следующим комбинированием сомножителей: *ABCDE=(AB)(C(DE)),* в соответствии с чем имеем:

$$AB=\begin{bmatrix}-3\\3\\5\end{bmatrix};\ DE=\begin{bmatrix}1\\-7\end{bmatrix};\ C(DE)=-10;\ ABCDE=\begin{bmatrix}30\\-30\\-50\end{bmatrix}.$$

4. Операции над столбцами и строками. Обозначим i-ю строку матрицы A через $a_{(i)}$ и j-й столбец — через $a^{(j)}$. Тогда каждую $(m\times n)$-матрицу можно представить как столбец из m ее строк или как строку с n ее столбцов:

$$A=\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{m1} & a_{m2} & \dots & a_{mn}\end{bmatrix}=\begin{bmatrix}a_{(1)}\\ a_{(2)}\\ \\ a_{(m)}\end{bmatrix}=[a^{(1)},\ a^{(2)},\ \dots,\ a^{(n)}].$$

При таком представлении произведение матриц сводится к операциям над их строками и столбцами. В зависимости от способа представления матриц-сомножителей $A(m\times n)$ и $B(n\times r)$ имеем два способа образования произведения AB:

$$AB=\begin{bmatrix}a_{(1)}\\ a_{(2)}\\ \dots\\ a_{(m)}\end{bmatrix}[b^{(1)},\ b^{(2)},\ \dots,\ b^{(r)}]=\begin{bmatrix}a_{(1)}b^{(1)} & a_{(1)}b^{(2)} & \dots & a_{(1)}b^{(r)}\\ a_{(2)}b^{(1)} & a_{(2)}b^{(2)} & \dots & a_{(2)}b^{(r)}\\ \dots & \dots & \dots & \dots\\ a_{(m)}b^{(1)} & a_{(m)}b^{(2)} & \dots & a_{(m)}b^{(r)}\end{bmatrix};$$

$$AB=[a^{(1)},\ a^{(2)},\ \dots,a^{(n)}]\begin{bmatrix}b_{(1)}\\ b_{(2)}\\ \dots\\ b_{(n)}\end{bmatrix}=\sum_{s=1}^{n}a^{(s)}b_{(s)}.$$

В первом случае ij-элемент результирующей $(m\times r)$-матрицы равен произведению i-й строки первой матрицы на j-й столбец второй матрицы, т.е.

$$a_{(i)}b^{(j)} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{s=1}^{n} a_{is}b_{sj},$$

что совпадает с правилом, приведенным в микромодуле 1. Во втором случае результат равен сумме n произведений столбцов первой матрицы на соответствующие строки второй $a^{(s)}b_{(s)}$ $(s = 1, 2, ..., n)$. Каждое такое произведение представляет собой $(m\times r)$-матрицу, ij-элемент которой равен $a_{is}b_{sj}$, а их сумма дает матрицу, элементы которой выражаются так же, как и в первом случае.

Представив столбец $a^{(s)}$ или строка $b_{(s))}$ в развернутом виде, получим:

$$AB = \begin{bmatrix} \sum_{s=1}^{n} a_{1s}b_{(s)} \\ \sum_{s=1}^{n} a_{2s}b_{(s)} \\ \cdots \\ \sum_{s=1}^{n} a_{ms}b_{(s)} \end{bmatrix} = \left[\sum_{s=1}^{n} a^{(s)}b_{s1}, \quad \sum_{s=1}^{n} a^{(s)}b_{s2}, \ \ldots, \ \sum_{s=1}^{n} a^{(s)}b_{sr} \right].$$

Отсюда следует соотношение для строки и столбца матрицы $C= AB$:

$$c_{(i)} = \sum_{s=1}^{n} a_{is}b_{(s)}; \qquad c^{(j)} = \sum_{s=1}^{n} a^{(s)}b_{sj}.$$

Как видно, i-я строка произведения двух матриц получается суммированием строк второй матрицы, умноженных на соответствующие элементы i-й строки первой матрицы. Аналогично j-й столбец произведения получается суммированием столбцов первой матрицы, умноженных на соответствующие элементы j-го столбца второй матрицы. Эти соотношения особенно удобны, если одна из матриц разрежена, т.е. значительное количество ее элементов равно нулю, например:

$$A=\begin{vmatrix} 2 & -1 & 4 \\ -1 & 1 & 3 \\ -3 & 2 & -2 \end{vmatrix}; \quad B=\begin{vmatrix} 2 & & & 3 \\ & 1 & & \\ -3 & & 4 & \end{vmatrix};$$

$$C=\begin{vmatrix} 2a^{(1)}-3a^{(3)} & a^{(2)} & 4a^{(3)} & 3a^{(1)} \end{vmatrix}=\begin{vmatrix} -8 & -1 & 16 & 6 \\ -11 & 1 & 12 & -3 \\ 0 & 2 & -8 & -9 \end{vmatrix}.$$

5. Произведения с диагональной матрицей. Если в произведении $C=AB$ матрица A диагональна, т.е. $a_{is}=0$ для всех значений s, кроме $s=i$, то $c_{(i)}=a_{ii}b_{(i)}$. Это означает, что при умножении матрицы на диагональную слева строки этой матрицы умножаются на соответствующие элементы диагональной матрицы:

$$\begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \dots & \\ & & & a_n \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1r} \\ b_{21} & b_{22} & \dots & b_{2r} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nr} \end{bmatrix}=\begin{bmatrix} a_1b_{11} & a_1b_{12} & \dots & a_1b_{1r} \\ a_2b_{21} & a_2b_{22} & \dots & a_2b_{2r} \\ \dots & \dots & \dots & \dots \\ a_nb_{n1} & a_nb_{n2} & \dots & a_nb_{nr} \end{bmatrix}.$$

Если диагональной является вторая матрица, т.е. $b_{sj}=0$ для всех значений s, кроме $s=j$, то $c^{(j)}=a^{(j)}b_{jj}$. Это значит, что для умножения матрицы на диагональную справа достаточно столбцы этой матрицы умножить на соответствующие элементы диагональной матрицы:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}\begin{bmatrix} b_1 & & & \\ & b_2 & & \\ & & \dots & \\ & & & b_n \end{bmatrix}=\begin{bmatrix} a_{11}b_1 & a_{12}b_2 & \dots & a_{1n}b_n \\ a_{21}b_1 & a_{22}b_2 & \dots & a_{2n}b_n \\ \dots & \dots & \dots & \dots \\ a_{m1}b_1 & a_{m2}b_2 & \dots & a_{mn}b_n \end{bmatrix}.$$

Произведение двух диагональных матриц является также диагональной матрицей. Очевидно, в этом случае $AB=BA$, *т.е.* диагональные матрицы коммутируют.

Диагональная матрица, все ненулевые элементы которой равны между собой, называется *скалярной.* Такую матрицу можно представить как единичную, умноженную на скаляр, т.е.

$$\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \cdots & \\ & & & \lambda \end{bmatrix} = \lambda \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \cdots & \\ & & & 1 \end{bmatrix} = \lambda E_n.$$

Умножение матрицы справа или слева на скалярную, элементы которой равны *X*, сводится к умножению этой матрицы на скаляр λ.

6. Степени матрицы. Произведения одинаковых квадратных матриц A можно записать как ее *степень:* $AA = A^2$, $AAA = A^2A = A^3$ и т.д. Вообще $A^r = AA^{r-1}$ представляет собой r-ю степень матрицы A (r — целое положительное число). Если A - неособенная матрица, то $A^{-r} = (A^{-1})^r$. Как и для чисел имеют место обычные свойства:

$$A^pA^q = A^{p+q}; \ (Ap)^q = A^{pq},$$

где p и q — положительные числа для произвольной квадратной матрицы и любые целые числа (положительные, отрицательные и нуль) для неособенной матрицы. В частности, $A^0 = AA^{-1} = 1$.

Никакая степень числа, отличная от нуля, не может равняться нулю. В то же время степень квадратной матрицы A^r может равняться нулю, даже если A — ненулевая матрица. Если $A^r=0$ для некоторого положительного числа r, то A называется *нильпотентной матрицей.* Например

$$H = \left\|\begin{array}{c|c|c|c|c} & 1 & & & \\ \hline & & 1 & & \\ \hline & & & 1 & \\ \hline & & & & 1 \\ \hline & & & & \end{array}\right\|; \quad H^2 = \left\|\begin{array}{c|c|c|c|c} & & 1 & & \\ \hline & & & 1 & \\ \hline & & & & 1 \\ \hline & & & & \\ \hline & & & & \end{array}\right\|.$$

Продолжая процесс умножения матрицы на себя, получаем $H^5 = 0$, т.е. матрица H является нильпотентной.

Условились называть p – й *наддиагональю* совокупность (i, i + 1)-клеток матрицы при $i = p, p + 1, ..., n — p$. Как видно, рассмотренная выше матрица H характеризуется тем, что элементы ее первой наддиагонали равны единице, а остальные - нулю. При каждом умножении этой матрицы самой на себя ее ненулевые элементы смещаются на следующую наддиагональ, так что H^p (при $p<n$) имеет

единичные элементы только на p–й наддиагонали, а при $p=n$ все элементы матрицы становятся нулевыми.

Аналогично совокупность $(i$ -1, $i)$-клеток матрицы при
$i = p, p + 1, ..., n - p$ называют p–й *поддиагональю*. Матрица F
n-го порядка, у которой все элементы первой поддиагонали равны единицы, а остальные равны нулю, также нильпотентна и $F^r=0$ при $r> n$.

7. Многочлены от матрицы. Подобно многочлену от числовой переменной $f(x)=a_0+a_1x+\ldots+a_mx^m$ водится понятие *многочлена от матрицы:*

$$f(A)=a_0 E+ a_1A + ... + a_mA^m,$$

где коэффициенты a_0, a_1, ..., a_m, являются вещественными или комплексными числами.

Формально многочлен от матрицы можно рассматривать как результат подстановки в алгебраический многочлен $f(x)$ вместо переменной x квадратной матрицы A. При этом матричный многочлен $f(A)$ также является квадратной матрицей того же порядка, что и матрица A.

Правила действий над многочленами от матрицы подобны соответствующим правилам для обычных (скалярных) многочленов. Так, если $f(x) = g(x) \pm h(x)$, то $f(A) = g(A) \pm h(A)$; если
$f(x) = g(x)h(x)$, то $f(A) = g(A)h(A) = h(A)g(A)$ (два многочлена от одной и той же матрицы всегда перестановочны между собой в силу перестановочносты степеней матрицы).

В качестве примеров приведем выражения многочленов от нильпотентных матриц, рассмотренных в (6). Для матриц пятого порядка имеем:

$$f(H)=\begin{vmatrix} a_0 & a_1 & a_2 & a_3 & a_4 \\ & a_0 & a_1 & a_2 & a_3 \\ & & a_0 & a_1 & a_2 \\ & & & a_0 & a_1 \\ & & & & a_0 \end{vmatrix}; \quad f(F)=\begin{vmatrix} a_0 & & & & \\ a_1 & a_0 & & & \\ a_2 & a_1 & a_0 & & \\ a_3 & a_2 & a_1 & a_0 & \\ a_4 & a_3 & a_2 & a_1 & a_0 \end{vmatrix}.$$

Аналогичные выражения можно записать для нильпотентных матриц любого порядка, причем они справедливы для любой степени матричного многочлена.

8. Прямая сумма квадратных матриц. Эта операция ставит в соответствие двум матрицам A и B порядков m и n квадратную матрицу

C порядка $m + n$ и обозначается $A \oplus B$. Она сводится к присоединению правого нижнего угла матрицы A к левому верхнему углу матрицы B и заполнению других клеток таблицы размера $(m + n) \times (m + n)$ нулями, т.е.:

$$C = A \oplus B = \begin{array}{|c|c|} \hline A & \\ \hline & B \\ \hline \end{array}.$$

Элементы *прямой суммы* определяются соотношениями:
$c_{ij} = a_{ij}$ $(i, j = 1, 2.....m)$, $c_{m+i,\ m+j} = b_{ij}$ $(i,\ \ j = 1, 2.....n)$ и $c_{ij} = 0$ для других элементов. Ясно, что эта операция не коммутативна, но ассоциативна. Она распространяется на любое количество матриц $A_1 \oplus A_2 \oplus A_3 = (A_1 \oplus A_2) \oplus A_3$ и т.д., причем порядок результирующей матрицы равен сумме порядков исходных матриц.

Если A_i — матрицы первого порядка, которые отождествляются со скалярами a_i $(i = 1, 2, ..., m)$, то их прямая сумма равна диагональной матрице m-го порядка $\mathrm{diag}(a_1, a_2, ..., a_m)$. В общем случае прямая сумма матриц A_i произвольных порядков образует квазидиагональную матрицу

$$\Lambda = A_1 \oplus A_2 \oplus \ldots \oplus A_m = \mathrm{diag}(A_1 \oplus A_2 \oplus \ldots \oplus A_m),$$

т.е.

$$\Lambda = \begin{array}{|c|c|c|c|} \hline A_1 & & & \\ \hline & A_2 & & \\ \hline & & \cdots & \\ \hline & & & A_m \\ \hline \end{array}$$

9. Кронекерово произведение прямоугольных матриц. Эта операция может выполняться над прямоугольными матрицами любых размеров. *Кронекерово (прямое, тензорное) произведение* $(m \times n)$*-матрицы* A на $(p \times q)$-матрицу B обозначается $A \otimes B$ и выражается матрицей размера $(mp \times nq)$:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}.$$

Можно показать, что при существовании обычных матричных произведений *АСС* и *BD* справедливо соотношение:

$$(A \otimes B)(C \otimes D) = AC \otimes BD.$$

Имеют место также следующие свойства кронекерового произведения:

$$(A \otimes B)^t = A^t \otimes B^t; \quad (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

10. Произведение векторов. *Скалярное произведение* комплексных n-мерных векторов $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ можно выразить как произведение строки на столбец

$$\langle x, y \rangle = [x_1, x_2, \dots, x_n] \begin{bmatrix} y_1^* \\ y_2^* \\ \cdots \\ y_n^* \end{bmatrix} =$$

$$= x_1 y_1^* + x_2 y_2^* + \cdots + x_n y_n^* = \sum_{i=1}^{n} x_i y_i^*,$$

где y_i^* — комплексно-сопряженное y_i. Отсюда следует также соотношение:

$$[x_1, x_2, \dots, x_n] \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i = \langle x, y^* \rangle.$$

Рассматривая строки и столбцы как векторы, произведения матриц согласно (4) можно представить в виде:

$$AB \begin{bmatrix} \langle a_{(1)}, b^{(1)*} \rangle & \langle a_{(1)}, b^{(2)*} \rangle & \cdots & \langle a_{(1)}, b^{(r)*} \rangle \\ \langle a_{(2)}, b^{(1)*} \rangle & \langle a_{(2)}, b^{(2)*} \rangle & \cdots & \langle a_{(2)}, b^{(r)*} \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle a_{(m)}, b^{(1)*} \rangle & \langle a_{(m)}, b^{(2)*} \rangle & \cdots & \langle a_{(m)}, b^{(r)*} \rangle \end{bmatrix},$$

причем общий элемент матрицы $C = AB$ выражается как $c_{ij} = \langle a_{(i)}, b^{(j)*} \rangle$.

Кроме скалярного (внутреннего) произведения, вводится также понятие *внешнего произведения векторов* x и y, которое обозначается $\{x, y\}$ и соответствует умножению столбца на строку:

$$\{x, y\} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix} [y_1^*, y_2^*, \dots, y_r^*] = \begin{bmatrix} x_1y_1^* & x_1y_2^* & \cdots & x_1y_r^* \\ x_2y_1^* & x_2y_2^* & \cdots & x_2y_r^* \\ \dots & \dots & \dots & \dots \\ x_my_1^* & x_my_2^* & \cdots & x_my_r^* \end{bmatrix}.$$

Следует обратить внимание на то, что скалярное произведение определяется как операция над векторами одинаковой размерности, результатом которой является число (скаляр). Внешнее произведение допускает различные размерности векторов (m и r), а его результат — матрица размера ($m \times r$).

Произведение ($m \times n$)-матрицы A на ($n \times r$)-матрицу B выражается через внешние произведения столбцов первой матрицы и строк второй следующим образом:

$$AB = \sum_{s=1}^{n} \{a^{(s)}, b_{(s)}^*\}.$$

11. Дифференцирование и интегрирование матриц. Часто, например при рассмотрении дифференциальных уравнений в матричной форме, приходится иметь дело с матрицами, элементами которых есть не числа, а функции от скалярного аргумента (времени t или любой другой переменной). Дифференцирование и интегрирование таких матриц сводится к правилам, аналогичным обычным правилам дифференцирования и интегрирования с одним существенным отличием. Так как произведение матриц в общем случае некоммутативно, то необходимо следить за сохранением первоначального порядка следования сомножителей.

Пусть матрица $X(t)$ размера $m \times n$ имеет своими элементами дифференцируемые функции $x_{ij}(t)$ скалярного аргумента t. *Производная матрице* $X(t)$ по переменной t определяется как

$$\frac{dX(t)}{dt} = \begin{bmatrix} \frac{dx_{11}(t)}{dt} & \frac{dx_{12}(t)}{dt} & \dots & \frac{dx_{1n}(t)}{dt} \\ \frac{dx_{21}(t)}{dt} & \frac{dx_{22}(t)}{dt} & \dots & \frac{dx_{2n}(t)}{dt} \\ \dots & \dots & \dots & \dots \\ \frac{dx_{n1}(t)}{dt} & \frac{dx_{n2}(t)}{dt} & \dots & \frac{dx_{nn}(t)}{dt} \end{bmatrix},$$

т.е. дифференцирование матрицы сводится к дифференцированию всех ее элементов по той же переменной. Имеют место также соотношения:

$$\frac{d}{dt}(X + Y) = \frac{dX}{dt} + \frac{dY}{dt};$$
$$\frac{d}{dt}(XY) = \frac{dX}{dt}Y + X\frac{dY}{dt}.$$

Если в первом из приведенных соотношений порядок следования матриц и их производных безразличен, то во втором он должен быть строго выдержан.

Если матрица $X(t)$ — дифференцируема и имеет обратную $X^{-1}(t)$, то *производная от обратной матрицы* определяется соотношением:

$$\frac{d}{dt}[X^{-1}(t)] = -X^{-1}(t)\frac{dX(t)}{dt}X^{-1}(t).$$

Действительно, так как $X(t)X^{-1}(t)=E$, то после дифференцирования с учетом того, что производная единичной матрицы равна нулевой матрице, имеем

$$\frac{dX(t)}{dt}X^{-1}(t) + X(t)\frac{dX^{-1}(t)}{dt} = 0,$$

откуда

$$X(t)\frac{dX^{-1}(t)}{dt} = -\frac{dX(t)}{dt}X^{-1}(t).$$

Умножая обе части равенства слева на $X^{-1}(t)$, получаем приведенное высшее соотношение для производной обратной матрицы.
Пример:

$$X = \begin{bmatrix} t+1 & t^2 \\ 1 & t-1 \end{bmatrix}; \quad X^{-1} = \begin{bmatrix} -t+1 & t^2 \\ 1 & -t-1 \end{bmatrix}; \quad \frac{dX^{-1}}{dt} = \begin{bmatrix} -1 & 2t \\ 0 & -1 \end{bmatrix}.$$

По формулам дифференцирования обратной матрицы получаем тот же результат:

$$\frac{dX^{-1}}{dt} = -\begin{bmatrix} -t+1 & t^2 \\ 1 & -t-1 \end{bmatrix}\begin{bmatrix} 1 & 2t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -t+1 & t^2 \\ 1 & -t-1 \end{bmatrix} = \begin{bmatrix} -1 & 2t \\ 0 & -1 \end{bmatrix}.$$

Интеграл от матрицы X(t) определяется как матрица, элементы которой равны интегралам от соответствующих элементов исходной матрицы, т.е.

$$\frac{dX^{-1}}{dt} = -\begin{bmatrix} -t+1 & t^2 \\ 1 & -t-1 \end{bmatrix}\begin{bmatrix} 1 & 2t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -t+1 & t^2 \\ 1 & -t-1 \end{bmatrix} = \begin{bmatrix} -1 & 2t \\ 0 & -1 \end{bmatrix}.$$

Микромодуль 2
Примеры тестовых задач

1. Выразить $(m\times n)$-матрицы A и B через базис линейного пространства I_{ij} (i = 1, 2, ..., m; j = 1, 2, ..., n):

$$A = \begin{bmatrix} 5 & -4 \\ 2 & 1 \end{bmatrix}; \quad B = \begin{bmatrix} 2 & 0 & 7 \\ -6 & 2 & 5 \end{bmatrix}.$$

2. Показать, что в линейной алгебре квадратных матриц n-го порядка нейтральными элементами относительно сложения и умножения являются соответственно нулевая и единичная матрицы того же порядка.
3. Даны матрицы:

$$A = \begin{bmatrix} 1 & 3 \\ -2 & 1 \\ 0 & 2 \end{bmatrix}; \quad B = \begin{bmatrix} 6 \\ 1 \\ 3 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 2 & -1 \end{bmatrix}; \quad D = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}.$$

а) Выполнить следующие действия (если они допустимы): $A - 2B$; $B - (2D + B) + D$; $2C + A$; $C - B + 2D$.

б) Определить матрицу X из уравнений: $6X = B$; $A + X = C$; $D - 2X = 3B$.

4. Найти произведение $ABCD$ матриц:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix}; \quad B = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}; \quad C = [-1 \ \ 2 \ \ 1]; \quad D = \begin{bmatrix} 3 & 1 \\ -1 & 0 \\ 2 & 4 \end{bmatrix}$$

а) умножением слева;

б) умножением справа;

в) наиболее рациональным комбинированием сомножителей.

5. Найти произведение AB матриц:

$$A=\begin{bmatrix} 1 & 4 & 1 & 0 \\ 0 & 3 & 5 & 3 \\ 2 & 0 & 1 & -1 \\ -1 & 1 & 2 & 1 \end{bmatrix}; \quad B=\begin{bmatrix} 3 & 1 & 3 \\ 2 & 0 & 2 \\ -1 & 4 & 1 \\ 0 & 5 & -3 \end{bmatrix}$$

а) умножением строк на столбцы;

б) умножением столбцов на строки.

6. Определите вторую строку и третий столбец произведения матриц из задачи 5 по формулам:

$$c_{(i)} = \sum_{s=1}^{n} a_{is} b_{(s)}; \quad c^{(j)} = \sum_{s=1}^{n} a^{(s)} b_{sj}.$$

7. Найти произведение матрицы A на матрицу B слева и справа алгебраическим суммированием строк и столбцов матрицы A, если

$$A=\begin{bmatrix} 2 & 0 & 1 & 0 \\ -1 & 4 & 0 & -2 \\ 1 & 0 & 3 & 1 \\ 0 & -2 & 5 & 3 \end{bmatrix}; \quad B=\begin{bmatrix} 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

8. Каким операциям над строками и столбцами квадратной матрицы третьего порядка соответствует умножение ее слева и справа на матрицы:

$$\text{а) } \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \text{б) } \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \text{в) } \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$$\text{г) } \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \text{д) } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}; \quad \text{е) } \begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

9. Найти внутренние и внешние произведения векторов x и y:

а) $x = (2, 0, -3, 5)$; $y = (4, 1, 2, -1)$;

б) $x = (2 - 3i, 1 + 2i, 2)$; $y = (1 + i, 1 - 2i, 3 + 2i)$.

10. Вычислить степени следующих матриц:

$$\text{а) } \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix}^3; \quad \text{б) } \begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & 2 \\ 3 & -1 & 1 \end{bmatrix}^3; \quad \text{в) } \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}^n.$$

11. Вычислить многочлены $f(A) = A^2 + 3A + E$ и $g(A) = A^3 - 2A^2 + A - 3E$, где

$$A=\begin{bmatrix} 1 & -1 & 4 \\ 0 & 3 & 5 \\ -2 & 1 & 0 \end{bmatrix}.$$

12. Найти сумму и произведение многочленов $f(A)$ и $g(A)$ из задачи 11 двумя способами:

а) на основе результатов, полученных в задаче 11;

б) путем операций над многочленами и последующим определением результирующей матрицы.

13. Найти произведение квадратных матриц n-го порядка $W_1, W_2, ..., W_n$, если матрица W_k $(k = 1,2, ..., n)$ имеет вид:

$$W_k = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & a_{kk} & & \\ & & \vdots & \ddots & \\ & & a_{kk} & & 1 \end{bmatrix}$$

14. Представить матрицу

10	—1			
4	2			
		—3		
			5	
			1	—21

в квазидиагональной форме и выразить ее через прямую сумму диагональных блоков. Единственно ли такое решение?

15. Определить кронекерово произведение $A \otimes B$ матриц

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & 1 & -2 \end{bmatrix}; \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

16. Выполнить дифференцирование и интегрирование матриц:

$$\text{а) } \frac{d}{dt}\begin{bmatrix} 2t-2 & t+2 & -2 \\ 3t^2+2t-1 & t^3 & 2t \end{bmatrix};$$

$$\text{б) } \int_0^t \begin{bmatrix} \cos t & t^2 \\ 1 & \operatorname{th} t \end{bmatrix} dt.$$

17. Найти производную по t произведения матриц X и Y

$$X = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \cos t & \sin t \end{bmatrix}; \quad Y = \begin{bmatrix} \cos t & \cos\alpha \\ \sin t & \sin\alpha \end{bmatrix}:$$

а) по формуле дифференцирования произведения матриц;

б) путем перемножения матриц с последующим дифференцированием результата.

Микромодуль 3
Квадратные и псевдообратные матрицы

1.3. Квадратные матрицы

В этом разделе мы напомним некоторые положения о матрицах, которые были изложены раньше.

1. Как мы уже говорили, квадратную матрицу n-го порядка, у которой на главной диагонали стоят единицы, а все другие элементы равны нулю, называют *единичной матрицей* и обозначают через $E^{(n)}$ или просто E. Название «единичная матрица» связана со следующим свойством матрицы E: для любой прямоугольной матрицы

$$A = ||a_{ik}|| \; (i=1,2, ... m; \; k = 1,2, ... n)$$

имеют место равенства

$$E^{(m)}A = AE^{(n)} = A.$$

Очевидно,

$$E^{(n)} = \| \delta_{ik} \|^n_1$$

Пусть $A = \| a_{ik} \|^n_1$ — квадратная матрица. Тогда *степень матрицы* определяется обычным образом:

$$A^p = \underbrace{AA...A}_{p \text{ раз}} \quad (p = 1, 2, ...); \qquad A^0 = Э.$$

Из сочетательного свойства умножения матриц следует:

$$A^p A^q = A^{p+q}.$$

Здесь p, q — произвольные целые неотрицательные числа.

Рассмотрим многочлен (целую рациональную функцию) с коэффициентами из поля K:

$$f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + ... + \alpha_m.$$

Тогда под $f(A)$ будем понимать матрицу

$$f(A) = \alpha_0 A^m + \alpha_1 A^{m-1} + ... + \alpha_m E.$$

Так определяется *многочлен от матрицы.*

Пусть многочлен $f(t)$ равен произведению многочленов $h(t)$и $g(t)$:

$$f(t) = h(t)\, g(t).$$

Многочлен $f(t)$ получается из $h(t)$ и $g(t)$ путем почленного перемножения и приведения подобных членов. При этом используется правило перемножения степеней: $t^p t^q = t^{p+q}$. Так как все эти действия правомерны и при замене скалярной величины t на матрицу A, то

$$f(A) = h(A)g(A).$$

Отсюда, в частности,

$$h(A)g(A) = g(A)h(A),$$

т.е. *два многочлена от одной и той же матрицы всегда перестановочны между собой.* (Так как каждое из этих произведений равно одному и тому же $f(A)$, поскольку и $g(t)h(t)=f(t)$. Следует отметить, что подстановка матриц в алгебраическое тождество с несколькими переменными неправомерна. Впрочем, перестановочные между собой матрицы можно подставлять и в этом случае).

Примеры.

Условимся p-й *наддиагональю (поддиагональю)* в прямоугольной матрице $A = \| a_{ik} \|$ называть ряд элементов a_{ik}, у которых $k — i = p$ (соответственно $i — k = p$). Обозначим через H квадратную матрицу n-го порядка, у которой элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Тогда

$$H = \begin{Vmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & 0 & \dots & 0 \end{Vmatrix}, \qquad H^2 = \begin{Vmatrix} 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & 1 \\ \cdot & & & \cdot & 0 \\ 0 & 0 & 0 & \dots & 0 \end{Vmatrix} \text{ и т. д.;}$$

$$H^p = 0 \qquad (p \geq n).$$

В силу этих равенств, если

$$f(t)= a_0 + a_1t + a_2t^2 + \dots + a_{n-1}t^{n-1} + \dots$$

- многочлен относительно t, то

$$f(H) = a_0E + a_1H + a_2H^2 + \dots = \begin{Vmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & a_0 & a_1 & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & a_2 \\ \cdot & & & \cdot & a_1 \\ 0 & 0 & 0 & \dots & a_0 \end{Vmatrix}.$$

Аналогично, если F— квадратная матрица n-го порядка, у которой все элементы первой поддоагонали равны единице, а все остальные элементы равны нулю, то

$$f(F) = a_0E + a_1F + a_2F^2 + \dots = \begin{Vmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & 0 \\ a_{n-1} & \dots & a_1 & a_0 \end{Vmatrix}.$$

2. Напомним, что квадратную матрицу будем называть *особенной,* если $|A|=0$. В противном случае квадратная матрица A называется *неособенной.*

Пусть $A=\|a_{ik}\|^n_1$ — неособенная матрица ($|A|\neq 0$). Рассмотрим линейное преобразование с матрицей коэффициентов A

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \quad (i = 1, 2, \ldots, n). \tag{1.24}$$

Рассматривая равенства (1.24) как уравнение относительно $x_1, x_2, \ldots, x_n$ и замечая, что определитель системы уравнений (1.24) по условию отличен от нуля, мы можем однозначно по известным формулам выразить $x_1, x_2, \ldots, x_n$ через $y_1, y_2, \ldots, y_n$:

$$x_i = \frac{1}{|A|} \begin{vmatrix} a_{11} & \ldots & a_{1,\, i-1} & y_1 & a_{1,\, i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,\, i-1} & y_2 & a_{2,\, i+1} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \ldots & a_{n,\, i-1} & y_n & a_{n,\, i+1} & \ldots & a_{nn} \end{vmatrix} \equiv \sum_{k=1}^{n} a_{ik}^{(-1)} y_k \; (i = 1, 2, \ldots, n). \tag{1.25}$$

Мы получили «обратное» преобразование для (1.24). Матрицу коэффициентов этого преобразования

$$A^{-1} = \| a_{ik}^{(-1)} \|_1^n$$

называют, как мы знаем, *обратной матрицей* для матрицы A. Из (1.25) легко увидеть, что

$$a_{ik}^{(-1)} = \frac{A_{ki}}{|A|} \quad (i, k = 1, 2, \ldots, n), \tag{1.26}$$

где A_{ki} — алгебраическое дополнение (адъюнкта) элемента a_{ki} в определителе $|A|$ ($i, k = 1,2, \ldots, n$). Так, например, если

$$A = \begin{Vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{Vmatrix} \text{ и } |A| \neq 0,$$

то

$$A^{-1} = \frac{1}{|A|} \begin{Vmatrix} b_2c_3 - b_3c_2 & a_3c_2 - a_2c_3 & a_2b_3 - a_3b_2 \\ b_3c_1 - b_1c_3 & a_1c_3 - a_3c_1 & a_3b_1 - a_1b_3 \\ b_1c_2 - b_2c_1 & a_2c_1 - a_1c_2 & a_1b_2 - a_2b_1 \end{Vmatrix}.$$

Образуя составное преобразование из данного преобразования (1.24) и обратного (1.25) в одном и в другом порядке, мы в обеих случаях получаем тождественное преобразование (с единичной матрицей коэффициентов); поэтому

$$A\,A^{-1} = A^{-1}A = E. \qquad (1.27)$$

В справедливости равенств (1.27) можно убедиться и непосредственным перемножением матриц A и A^{-1}. Действительно, в силу (1.26) (здесь мы используем известное свойство определителя, согласно которому сумма произведений элементов какого-нибудь столбца на их адъюнкты равна величине определителя, а сумма произведений элементов столбца на адъюнкты соответствующих элементов другого столбца равна нулю)

$$[AA^{-1}]_{ij} = \sum_{k=1}^{n} a_{ik} a_{kj}^{(-1)} = \frac{1}{|A|} \sum_{k=1}^{n} a_{ik} A_{jk} = \delta_{ij} \quad (i, j = 1, 2, \ldots, n).$$

Аналогично

$$[A^{-1}A]_{ij} = \sum_{k=1}^{n} a_{ij}^{(-1)} a_{kj} = \frac{1}{|A|} \sum_{k=1}^{n} A_{ki} a_{kj} = \delta_{ij} \quad (i, j = 1, 2, \ldots, n).$$

Легко видеть, что матричные уравнения

$$AX = E \quad \text{и} \quad XA = E \quad (|A| \neq 0) \qquad (1.28)$$

никаких других решений, кроме решения $X=A^{-1}$, не имеют. Действительно, перемножая обе части первого уравнения слева, а второго — справа на A^{-1} и используя сочетательное свойство произведения матриц, а также равенства (1.27), мы в обеих случаях получим:

$$X = A^{-1}$$

(Если A — особенная матрица, то уравнение (1.28) не имеют решений. Действительно, если бы какое-нибудь из этих уравнений имело решение $X=\|a_{ik}\|^{n}_{1}$, то тогда было бы в силу теоремы об умножении определителей [см. формулу (1.17)] $|A|\ |X|=|E|=1$, что невозможно при $|A| = 0$).

Этим же способом устанавливается, что каждое из матричных уравнений

$$AX = B, \quad XA = B \quad (|A| \neq 0), \qquad (1.29)$$

где X и B — прямоугольные матрицы равных размеров, A — квадратная матрица соответствующего размера, имеет одно и только одно решение:

$$X = A^{-1}B \text{ и соответственно } \quad X = BA^{-1} \qquad (1.30)$$

Матрицы (1.30) являются как бы «левым» и «правым» частными от «деления» матрицы B на матрицу A. Из (1.29) и (1.30) следует соответственно $r_B \leq r_X$ и $r_X \leq r_B$, т.е. $r_X = r_B$. Сопоставляя с (1.29), имеем:

При умножении прямоугольной матрицы слева или справа на неособенную матрицу ранг исходной матрицы не меняется.

Заметим еще, что из (1.27) следует $|A||A^{-1}| = 1$, т.е.

$$|A^{-1}| = \frac{1}{|A|}$$

Для произведения двух неособенных матриц имеем:

$$(AB)^{-1} = B^{-1}A^{-1} \tag{1.31}$$

3. Все матрицы n-го порядка образуют кольцо с единичным элементом E. (*Кольцом* называется совокупность элементов, в которой определены и всегда однозначно осуществимы две операции: «сложение» двух элементов (с переместительным и сочетательным свойствами) и «умножение» двух элементов (с сочетательным и распределительным относительно сложения свойствами), причем сложение оборотимо).

Поскольку в этом кольце определена операция умножения на число из поля K и существует базис из n^2 линейно независимых матриц, через которые линейно выражаются все матрицы n-го порядка, то кольцо матриц n-го порядка является *алгеброй*. (Действительно, произвольная матрица $A=\|a_{ik}\|^n_1$ с элементами из K представима в виде

$$A = \sum_{i,k=1}^{n} a_{ik}E_{ik},$$

где E_{ik} — матрица n-го порядка, у которой на пересечении i-й строки и k-го столбца стоит 1, а все остальные элементы равны 0).

Все квадратные матрицы n-го порядка образуют коммутативную группу относительно операции сложения. (*Группой* называется всякая совокупность объектов, в которой установлена операция, которая соотносит любым двум элементам a и b совокупности определенный третий элемент $a{*}b$ той же совокупности, если 1) операция обладает сочетательным свойство $[(a{*}b){*}c = a{*}(b{*}c)]$, 2) существует в совокупности единичный элемент $e(a{*}e=e{*}a=a)$ и 3) для любого элемента a в совокупности существует обратный элемента a^{-1} $(a{*}a^{-1}=a^{-1}{*}a=e)$. Группа называется *коммутативной* или *абелевой,* если групповая операция обладает переместительным свойством).

Все неособенные матрицы n-го порядка образуют (некоммутативную) группу относительно операции умножения.

Квадратная матрица $A=\|a_{ik}\|^n_1$ называется *верхней треугольной (нижней треугольной),* если равны нулю все элементы матрицы, которые расположены под главной диагональю (над главной диагональю):

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & & \cdot \\ 0 & 0 & \dots & a_{nn} \end{Vmatrix}, \quad A = \begin{Vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \cdot & \cdot & & \\ \cdot & & \cdot & \\ \cdot & & & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}.$$

(1) (2)

Диагональная матрица является частным случаем как верхней, так и нижней треугольной матрицы.

Так как определитель треугольной матрицы равен произведению ее диагональных элементов, то треугольная (и, в частности, диагональная) матрица является неособенной только тогда, когда все ее диагональные элементы отличны от нуля.

Легко проверить, что сумма и произведение двух диагональных (верхних треугольных, нижних треугольных) матриц является диагональная (соответственно верхняя треугольная, нижняя треугольная) матрица и что обратная матрица для неособенной диагональной (верхней треугольной, нижней треугольной) матрицы является матрицей того же типа. Поэтому

1° Все диагональные, все верхние треугольные, все нижние треугольные матрицы n-го порядка образуют три коммутативные группы относительно операции сложения.

2° Все неособенные диагональные матрицы образуют коммутативную группу относительно умножения.

3° Все неособенные верхние (нижние) треугольные матрицы представляют группу (некоммутативную) относительно умножения.

4. Напомним две важные операции над матрицами — *транспонирование* матрицы и переход к *сопряженной* матрице.

Если $A=\|a_{ik}\|$ $(i = 1, 2, ..., m;\ k = 1, 2, ..., n)$, то транспонированная матрица A' определяется равенством $A' = \|a_{ki}'\|$, где $a_{ki}' = a_{ik}$ $(i = 1, 2, ..., m;\ k = 1, 2, ..., n)$. Сопряженная же матрица $A^* =\| a^*_{ki} \|$, где $a^*_{ki} = \overline{a}'_{ki} = \overline{a}_{ik}$ $(i = 1, 2, ..., m;\ k = 1, 2, ..., n)$ (черточкой обозначается переход к комплексно-сопряженной величине).

Если матрица A имеет размер $m\times n$, то матрицы A' и A^* имеют размеры $n\times m$.

Легко проверяются свойства:

1° $(A + B)' = A' + B'$, $(A + B)^* = A^* + B^*$.

2° $(\alpha A)' = \alpha A'$, $(\alpha A)^* = \overline{\alpha} A^*$.

3° $(AB)' = B'A'$, $(AB)^* = B^*A^*$.
4° $(A^{-1})' = (A')^{-1}$ $(A^{-1})^* = (A^*)^{-1}$.
5° $(A')' = A$, $(A^*)^* = A$.

В формулах 1°, 2°, 3° и 5° A, B — произвольные прямоугольные матрицы, для которых соответствующие операции выполнимы, а α — произвольное комплексное число. В формуле 4° A — произвольная квадратная неособенная матрица.

Как мы уже говорили, если квадратная матрица $S=\|s_{ik}\|^n_1$ совпадает со своей транспонированной ($S'=S$), то такая матрица называется *симметрической.* Если же квадратная матрица $H=\| h_{ik} \|$ совпадает со своей сопряженной ($H^* = H$), то она называется *эрмитовой.* В симметрической матрице элементы, которые симметрично расположены относительно главной диагонали, равны, а в эрмитовой они комплексно-сопряжены между собой. Диагональные элементы эрмитовой матрицы всегда действительны. Заметим, что произведение двух симметрических (эрмитовых) матриц, вообще говоря, не является симметрической (эрмитовой) матрицей. В силу 3° это имеет место только в том случае, когда данные две симметричные или эрмитовы матрицы перестановочны между собой.

Если A — *вещественная* матрица, т.е. матрица с вещественными элементами, то $A^*=A'$. Эрмитова действительная матрица всегда является симметрической.

С каждой прямоугольной матрицей $A=\|a_{ik}\|$ размером $m\times n$ связаны две эрмитовы матрицы AA^* и A^*A размером $m\times m$ и $n\times n$. Любое из равенств $AA^* = 0$ или $A^*A = 0$ влечет за собой равенство $A = 0$.
Это следует из того, что сумма диагональных элементов в каждой из матриц AA^* и A^*A равна

$$\sum_{i=1}^{m}\sum_{k=1}^{n} | a_{ik} |^2.$$

Если квадратная матрица $K=\|k_{ij}\|^n_1$ отличается множителем —1 от своей транспонированной ($K'=-K$), то такая матрица называется *кососимметрической.* В кососимметрической матрицы любые два элемента, которые расположены симметрично относительно главной диагонали, отличаются друг от друга множителем —1, а диагональные элементы равны нулю. С 3° следует, что произведение двух перестановочних между собой кососимметрических матриц является симметрической матрицей.

1.4. Ассоциированные матрицы. Миноры обратной матрицы

Пусть дана матрица $A=\|a_{ik}\|^n_1$. Рассмотрим различные миноры p-го ($1 \le p \le n$) порядка матрицы A:

$$A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \left(1 \leqslant \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \leqslant n\right). \tag{1.32}$$

Число таких миноров равно N^2, где $N=C^p_n$ — число сочетаний из n по p. Для того чтобы расположить миноры (1.32) в квадратную таблицу, занумеруем в определеном (например, лексикографическом) порядке все сочетания по p з n индексов 1, 2, ..., n.

Если сочетания индексов $i_1 < i_2 < \dots < i_p$ и $k_1<k_2< \dots < k_p$ при такой нумерации будут иметь номера α и β, то минор (1.32) будем обозначать и так:

$$\mathfrak{a}_{\alpha\beta} = A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}.$$

Придавая α и β независимо друг от друга все значения от 1 до N, мы охватим все миноры p-го порядка матрицы $A=\|a_{ik}\|^n_1$. Квадратная матрица N -го порядка

$$\mathfrak{A}_p = \|\mathfrak{a}_{\alpha\beta}\|_1^N$$

называется p-й *ассоциированной матрицей* для матрицы $A=\|a_{ik}\|^n_1$|; p может принимать значение 1, 2, ..., n. При этом $\mathfrak{A}_1 = A$, а матрица $\mathfrak{A}_n$ состоит из одного элемента, равного $|A|$.

Замечание. Порядок нумерации сочетаний индексов фиксируется раз и навсегда и не связан с выбором матрицы A.

Пример. Пусть

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix}.$$

Перенумеруем все сочетания из четырех индексов 1, 2, 3, 4 по два, расположив их в следующем порядке:

(1 2) (1 3) (1 4) (2 3) (2 4) (3 4).

Тогда

$$\mathfrak{A}_2 = \left\| \begin{matrix} A\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix} & A\begin{pmatrix}1 & 2\\ 1 & 3\end{pmatrix} & A\begin{pmatrix}1 & 2\\ 1 & 4\end{pmatrix} & A\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix} & A\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix} & A\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} \\ A\begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix} & A\begin{pmatrix}1 & 3\\ 1 & 3\end{pmatrix} & A\begin{pmatrix}1 & 3\\ 1 & 4\end{pmatrix} & A\begin{pmatrix}1 & 3\\ 2 & 3\end{pmatrix} & A\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix} & A\begin{pmatrix}1 & 3\\ 3 & 4\end{pmatrix} \\ A\begin{pmatrix}1 & 4\\ 1 & 2\end{pmatrix} & A\begin{pmatrix}1 & 4\\ 1 & 3\end{pmatrix} & A\begin{pmatrix}1 & 4\\ 1 & 4\end{pmatrix} & A\begin{pmatrix}1 & 4\\ 2 & 3\end{pmatrix} & A\begin{pmatrix}1 & 4\\ 2 & 4\end{pmatrix} & A\begin{pmatrix}1 & 4\\ 3 & 4\end{pmatrix} \\ A\begin{pmatrix}2 & 3\\ 1 & 2\end{pmatrix} & A\begin{pmatrix}2 & 3\\ 1 & 3\end{pmatrix} & A\begin{pmatrix}2 & 3\\ 1 & 4\end{pmatrix} & A\begin{pmatrix}2 & 3\\ 2 & 3\end{pmatrix} & A\begin{pmatrix}2 & 3\\ 2 & 4\end{pmatrix} & A\begin{pmatrix}2 & 3\\ 3 & 4\end{pmatrix} \\ A\begin{pmatrix}2 & 4\\ 1 & 2\end{pmatrix} & A\begin{pmatrix}2 & 4\\ 1 & 3\end{pmatrix} & A\begin{pmatrix}2 & 4\\ 1 & 4\end{pmatrix} & A\begin{pmatrix}2 & 4\\ 2 & 3\end{pmatrix} & A\begin{pmatrix}2 & 4\\ 2 & 4\end{pmatrix} & A\begin{pmatrix}2 & 4\\ 3 & 4\end{pmatrix} \\ A\begin{pmatrix}3 & 4\\ 1 & 2\end{pmatrix} & A\begin{pmatrix}3 & 4\\ 1 & 3\end{pmatrix} & A\begin{pmatrix}3 & 4\\ 1 & 4\end{pmatrix} & A\begin{pmatrix}3 & 4\\ 2 & 3\end{pmatrix} & A\begin{pmatrix}3 & 4\\ 2 & 4\end{pmatrix} & A\begin{pmatrix}3 & 4\\ 3 & 4\end{pmatrix} \end{matrix} \right\| .$$

Отметим некоторые свойства ассоциированных матриц:

1° *Из* $C = AB$ *следует* $\mathfrak{C}_p = \mathfrak{A}_p\mathfrak{B}_p$ *(*$p = 1, 2, ..., n$*).*

Действительно, выражая миноры p-го порядка ($1 \le p \le n$) матрицы-произведения C через миноры того же порядка матриц-сомножителей по формуле (1.18), будем иметь:

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \sum_{1 \le l_1 < l_2 < \dots < l_p \le n} A\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ l_1 & l_2 & \dots & l_p \end{pmatrix} B\begin{pmatrix} l_1 & l_2 & \dots & l_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}$$

$$\left(1 \le \begin{matrix} i_1 < i_2 < \dots < i_p \\ k_1 < k_2 < \dots < k_p \end{matrix} \le n\right). \qquad (1.33)$$

Очевидно, в обозначениях этого раздела, равенство (1.33) может быть записано так:

$$\mathfrak{c}_{\alpha\beta} = \sum_{\lambda=1}^{N} \mathfrak{a}_{\alpha\lambda}\mathfrak{b}_{\lambda\beta} \quad (\alpha, \beta = 1, 2, \dots, N)$$

(здесь α, β, λ — номера сочетаний индексов $i_1 < i_2 < ... < i_p$; $k_1 < k_2 < ... < k_p$; $l_1 < l_2 < ... < l_p$). Отсюда

$$\mathfrak{C}_p = \mathfrak{A}_p\mathfrak{B}_p \quad (p = 1, 2, \dots, n).$$

2° *Из* $B = A^{-1}$ *следует:* $\mathfrak{B}_p = \mathfrak{A}_p^{-1}$ *(*$p = 1,2, \dots, n$*).*

Это предложение непосредственно вытекает из предыдущего, если там положить $C = E$ и обратить внимание на то, что $\mathfrak{E}_p$ — единичная матрица порядка $N = C^p_n$.

Из предложения 2° следует важная формула, которая выражает миноры обратной матрицы через миноры данной матрицы:

Если $B = A^{-1}$ то при любых

$$(1 \leqslant) \begin{matrix} i_1 < i_2 < \ldots < i_p \\ k_1 < k_2 < \ldots < k_p \end{matrix} (\leqslant n)$$

$$B\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} = \frac{(-1)^{\sum_{\nu=1}^{p} i_\nu + \sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k'_1 k'_2 & \ldots & k'_{n-p} \\ i'_1 i'_2 & \ldots & i'_{n-p} \end{pmatrix}}{A\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}}, \tag{1.34}$$

где $i_1 < i_2 < \ldots < i_p$ вместе с $i'_1 < i'_2 < \ldots < i'_{n-p}$, а $k_1 < k_2 < \ldots < k_p$ вместе с $k'_1 < k'_2 < \ldots < k'_{n-p}$ представляют полную систему индексов 1, 2, ..., n.

Действительно, из $AB = E$ следует:

$$\mathfrak{A}_p \mathfrak{B}_p = \mathfrak{E}_p$$

или в более подробной записи:

$$\sum_{\alpha=1}^{N} \mathfrak{a}_{\gamma\alpha} \mathfrak{b}_{\alpha\beta} = \delta_{\gamma\beta} = \begin{cases} 1 & (\gamma = \beta), \\ 0 & (\gamma \neq \beta). \end{cases} \tag{1.35}$$

Равенства (1.35) могут быть записаны еще так:

$$\sum_{1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n} A\begin{pmatrix} j_1 j_2 & \ldots & j_p \\ i_1 i_2 & \ldots & i_p \end{pmatrix} \cdot B\begin{pmatrix} i_1 i_2 & \ldots & i_p \\ k_1 k_2 & \ldots & k_p \end{pmatrix} = \begin{cases} 1, \text{ если } \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 = 0, \\ 0, \text{ если } \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 > 0 \end{cases}$$

$$\left(1 \leqslant \begin{matrix} j_1 < j_2 < \ldots < j_p \\ k_1 < k_2 < \ldots < k_p \end{matrix} \leqslant n\right). \tag{1.36}$$

С другой стороны, применяя к определителю $|A|$ известные разложения Лапласа, получаем:

$$\sum_{1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n} A\begin{pmatrix} j_1 j_2 & \ldots & j_p \\ i_1 i_2 & \ldots & i_p \end{pmatrix} \cdot (-1)^{\sum_{\nu=1}^{p} i_\nu + \sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k'_1 k'_2 & \ldots & k'_{n-p} \\ i'_1 i'_2 & \ldots & i'_{n-p} \end{pmatrix} =$$

$$= \begin{cases} |A|, & \text{если } \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 = 0, \\ 0, & \text{если } \sum_{\nu=1}^{p} (j_\nu - k_\nu)^2 > 0, \end{cases} \tag{1.37}$$

где $i'_1 < i'_2 < \ldots < i'_{n-p}$ вместе с $i_1 < i_2 < \ldots < i_p$, а $k'_1 < k'_2 < \ldots < k'_{n-p}$ вместе из $k_1 < k_2 < \ldots < k_p$ образуют полную систему индексов 1, 2, ..., n.

Сопоставление (1.37) с (1.36') и (1.35) показывает, что равенства (1.35) удовлетворятся, если вместо $b_{\alpha\beta}$ взять не

$$B\begin{pmatrix} i_1 i_2 \dots i_p \\ k_1 k_2 \dots k_p \end{pmatrix},$$

а

$$\frac{(-1)^{\sum_{\nu=1}^{p} i_\nu + \sum_{\nu=1}^{p} k_\nu} A\begin{pmatrix} k'_1 k'_2 \dots k'_{n-p} \\ i'_1 i'_2 \dots i'_{n-p} \end{pmatrix}}{A\begin{pmatrix} 1 & 2 \dots n \\ 1 & 2 \dots n \end{pmatrix}}.$$

Так как из системы (1.35) элементы $b_{\alpha\beta}$ обратной матрицы для $\mathfrak{A}_p$ определяются однозначно, то имеет место равенство (1.34).

1.5. Обращение матриц.

Значение обратной матрицы настолько велико в различных приложениях, что она заслуживает более подробного рассмотрения, чем это сделано в начале модуля. Если матрица A имеет обратную, то эта обратная матрица A^{-1} единственна. Действительно, если предположить, что существует другая такая матрица B, то по определению должно выполняться соотношение $AB=BA=1$. Умножая равенство $AB=1$ слева на A^{-1}, получаем $A^{-1}AB=A^{-1}$. Но так как $A^{-1}A = 1$, то $B = A^{-1}$, т.е. A^{-1}— единственна.

Напомним, что обратная матрица существует только для неособенной матрицы, определитель которой не равен нулю. Из изложенной ранее процедуры вычисления обратной матрицы следуют соотношения:

$$A^{-1}=\frac{A^{(+)}}{|A|};\quad A^{(+)} = |A|A^{-1};\quad A^{(+)}A = |A|E_n,$$

где $A^{(+)}$-*присоединенная матрица,* элементы которой получаются замещением элементов A^t их алгебраическими дополнениями, т.е.

$$A^{(+)} = \begin{vmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{vmatrix}.$$

Если матрица A симметрична, то присоединенная к ней матрица $A^{(+)}$ и обратная A^{-1} также симметричные. Присоединенная к матрице A обозначается также через Adj$A$.

2. Свойства обратных матриц. Подытожим изложенные ранее свойства обратных матриц и приведем некоторые другие их свойства,

1. Если существует A^{-1} и B^{-1}, то существует и $(AB)^{-1} = B^{-1} A^{-1}$.
2. $(A^{-1})^*$ «= $(A^*)^{-1}$.
3. $(A^{-1})^t = (A^t)^{-1}$.
4. $(A^{-1})^{-1} = A$.
5. $\det(A^{-1}) = \dfrac{1}{\det A}$.

6. Если A и B - неособенные матрицы, то $(A \oplus B)^{-1} = A^{-1} \oplus B^{-1}$, или в другой записи:

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}.$$

Последнее свойство проверяется по правилу умножения блочных матриц:

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix} \quad \begin{bmatrix} AA^{-1} & 0 \\ 0 & BB^{-1} \end{bmatrix} = \begin{bmatrix} E_n & 0 \\ 0 & E_n \end{bmatrix} = E_{2n}.$$

В частности, матрица, которая обратная диагональной матрице D=diag $[d_1,d_2 \ \dots \ , d_n]$, также диагональная с элементами, обратными элементам исходной, т.е.

$$D^{-1} = \mathrm{diag}\left[\frac{1}{d_1}, \frac{1}{d_2}, \ \dots, \ \frac{1}{d_n}\right].$$

Матрица A, равная своей обратной, называется *инволютивной (взаимно обратной),* т.е. для нее выполняется условие $A^{-1}=A$, или $AA=A^2=1$. В частности, единичная матрица является инволютивной, так как $E_n= E_n^{-1}$. Из соотношения $|A||A|=|A|^2=1$ следует, что определитель инволютивной матрицы равен ±1.

3. Обращение матрицы. Вычислив определитель матрицы и алгебраические дополнения всех ее элементов, можно получить обратную матрицу по формуле $A^{-1}= \dfrac{A^{(+)}}{|A|}$, где $A^{(+)}$— присоединенная матрица. Однако такой путь слишком громоздкий, так как он требует вычисления определителя n-го порядка и n определителей $(n-1)$-го порядка. Для решения эта задача разработано много практических алгоритмов, один из которых основан на соотношении $AA^{-1}=1$.

Обозначив $A^{-1}=X$, запишем матричное уравнение $AX=1$ в развернутом виде:

$$A[x^{(1)}, x^{(2)}.....x^{(n)}] = [e^{(1)}, e^{(2)}, \ldots, e^{(n)}].$$

где $x^{(1)}, x^{(2)}.....x^{(n)}$— столбцы искомой обратной матрицы;
$e^{(1)}, e^{(2)},...,e^{(n)}$ - столбцы единичной матрицы. Это соотношение равносильно n системам уравнений с n неизвестными (элементами столбцов обратной матрицы):

$$Ax^{(1)}= e^{(1)}, Ax^{(2)}= e^{(2)},\ldots,Ax^{(n)}= e^{(n)}$$

решение которых и дает

$$X = [x^{(1)}, x^{(2)}.....x^{(n)}] = A^{-1}.$$

Как видно, объем вычислений оказывается существенно большим, чем при решении одной системы уравнений с n неизвестными. Поэтому использование обратной матрицы для решения уравнения $Ax=b$ как $x=A^{-1}b$ нецелесообразно и лучше в таких случаях пользоваться другими методами, которые будут рассмотрены дальше.

Однако во многих практических задачах нужно многократно решать уравнения $Ax = b^{(i)}$ $(i = 1, 2, ..., m)$ с различными свободными членами $b^{(i)}$. Это соответствует матричному уравнению $AY = B$, где Y и B — прямоугольные матрицы размера $n\times m$, причем $B=[b^{(1)},b^{(2)} ..., b^{(m)}]$. Для получения его решения $Y=A^{-1}B$ достаточно один раз вычислить A^{-1} и умножить ее на матрицу Y свободных членов. В подобных случаях вычисления обратной матрицы может оказаться целесообразным. Кроме того, иногда по условию задачи необходимо получить обратную матрицу в чистом виде. Все это говорит в пользу алгоритмов обращения матриц.

4. Метод исключения. Уравнение $AX=1$ можно решить относительно $X = A^{-1}$ преобразованием матрицы A к единичной при условии соблюдению равенства его левой и правой частей. Воспользуемся для этого процедурой, которая немного напоминает схему единственного деления. Разделим элементы первой строки матрицы A на a_{11} и прибавим к остальным строкам эту строку, умноженную на $-a_{i1}$ $(i = 2, 3, ..., n)$. В результате получим $a'_{11} = 1$, а остальные элементы первого столбца обатятся в нуль. Далее вторую строку делим на новое значение a'_{22} и прибавляем к остальным строкам эту строку, умноженную на новые значения a'_{i2} $(i = 1, 3, ..., n)$. В результате получим $a''_{22}=1$, а остальные элементы второго столбца равны нулю. Через n таких шагов матрица A преобразуется в единичную матрицу.

На k-м шаге строки матрицы A', полученной на предыдущем шаге, преобразуются следующим образом:

$$a''_{(k)} = \frac{1}{a'_{kk}} a'_{(k)}; \quad a''_{(i)} = a'_{(i)} - \frac{a'_{ik}}{a'_{kk}} a'_{(k)} \ (i = 1, \ \ldots, \ n; \ i \neq k),$$

что можно представить как умножение A' слева на некоторую матрицу V_k того же порядка, т.е. $A'' = V_k A'$. Так как

$$a''_{(i)} = \sum_{s=1}^{n} v_{is} a'_{(s)},$$

то сравнивая с приведенными выше соотношениями, находим

$$v_{kk} = \frac{1}{a'_{kk}}, \quad v_{ik} = -\frac{a'_{ik}}{a'_{kk}}, \ v_{ii} = 1 \ (i = 1, 2, \ \ldots, \ n; \ i \neq k),$$

а остальные элементы матрицы V_k равны нулю. Отсюда следует, что матрица V_k, которая соответствует преобразованию матрицы A на k-м шаге, имеет вид:

$$V_k = \begin{bmatrix} 1 & \cdots & -\frac{a'_{1k}}{a'_{kk}} & & \\ & \ddots & \cdots & & \\ & & \frac{1}{a'_{kk}} & & \\ & & \cdots & \ddots & \\ & & -\frac{a'_{nk}}{a'_{kk}} & \cdots & 1 \end{bmatrix}.$$

Очевидно, произведение n таких матриц для $k = 1, 2, \ldots, n$, $V = V_1 V_2 \ldots V_k = A^{-1}$ и осуществляет преобразование A к единичной матрице. Чтобы равенство $AX = 1$ не нарушилась при умножении A на V слева, необходимо правую часть также умножить на V, т.е. $VAX = VE_n$. А это значит, что над строками единичной матрицы в правой части уравнения в процессе его преобразования необходимо выполнить те же операции, что и над строками матрицы A. Это удобно реализовать, оперируя над строками расширенной матрицы $[A, 1]$ и выбирая в кчестве опорных элементов диагональные элементы матрицы A.

Проиллюстрируем метод исключения на примере обращения матрицы

$$A = \begin{bmatrix} 2 & 1 & 0 \\ -3 & 0 & 7 \\ -5 & 4 & -1 \end{bmatrix};$$

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 0 & 1 & 0 & 0 \\ -3 & 0 & 7 & 0 & 1 & 0 \\ -5 & 4 & -1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{2} & 7 & \frac{3}{2} & 1 & 0 \\ 0 & \frac{13}{2} & -1 & \frac{5}{2} & 0 & 1 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & -\frac{7}{3} & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{14}{3} & 1 & \frac{2}{3} & 0 \\ 0 & 0 & -\frac{94}{3} & -4 & -\frac{13}{3} & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{14}{47} & -\frac{1}{94} & -\frac{7}{94} \\ 0 & 1 & 0 & \frac{19}{47} & \frac{1}{47} & \frac{7}{47} \\ 0 & 0 & 1 & \frac{6}{47} & \frac{13}{94} & -\frac{3}{94} \end{array}\right].$$

В итоге получаем обратную матрицу, которая расположена в трех последних столбцах. Как побочный результат, имеем также определитель матрицы *A,* который равен произведению тех значений опорных элементов, которые они принимают на соответствующих шагах преобразования матрицы [*A,* 1] (эти элементы выделены жирным шрифтом):

$$\det A = 2 \cdot \frac{3}{2} \cdot \left(-\frac{94}{3}\right) = -94.$$

Таким образом, одновременно получаем и присоединенную матрицу:

$$A^{+} = |A|\,A^{-1} = \begin{bmatrix} -28 & 1 & 7 \\ -38 & -2 & -14 \\ -12 & -13 & 3 \end{bmatrix}.$$

5. Выбор опорных элементов. При реализации метода исключения диагональный элемент на очередном шаге может оказаться равным нулю и его нельзя принять в качестве опорного. Кроме того, очередной диагональный элемент может быть нежелательным в качестве опорного по другим соображениям (например, если он слишком мал, что может привести к снижению надежности результата). Ничто не мешает в подобных случаях организовать процедуру исключения по любой совокупности опорных элементов с единственным ограничением: все они должны находиться в различных строках и различных столбцах матрицы. Иначе говоря, совокупность первых индексов $\alpha_1, \alpha_2, ..., \alpha_n$ и вторых индексов $\beta_1, \beta_2, \ldots, \beta_n$ опорных элементов $a_{\alpha_i \beta_i}$ образует подстановку из чисел 1, 2,..., *n:*

$$\sigma = \begin{pmatrix} \alpha_1, & \alpha_2, & \ldots, & \alpha_n \\ \beta_1, & \beta_2, & \ldots, & \beta_n \end{pmatrix}.$$

По окончании преобразования матрицы [*A*,1] на месте *A* получим не единичную матрицу, а некоторую матрицу *S*, а на месте единичной — матрицу *V*, т.е. [*S*, *V*]. В матрице *S*, называемой *матрицей подстановки,* (α_i, β_i)-элементы равны единице $(i = 1, 2, \ldots, n)$, а другие равны нулю. Она получается из единичной путем перестановки строк (или столбцов) в соответствии с подстановкой σ. Обратно, из *S* получим единичную матрицу перестановкой строк (или столбцов) в соответствии с обратной подстановкой

$$\sigma^{-1} = \begin{pmatrix} \beta_1, & \beta_2, & \ldots, & \beta_n \\ \alpha_1, & \alpha_2, & \ldots, & \alpha_n \end{pmatrix},$$

которой соответствует матрица S^{-1}. Чтобы получить обратную матрицу, необходимо преобразованное уравнение $SX = V$ умножить слева на S^{-1}, в результате получим $X=S^{-1}V=A^{-1}$. Практически это сводится к перестановке строк (или столбцов) матрицы *V* в соответствии с матрицей *S*, так чтобы последняя преобразовалась к единичной матрице.

Ниже приводится процедура обращения той же матрицы, что и в п.(4), но с произвольным выбором опорных элементов:

$$\left[\begin{array}{rrr|rrr} 2 & \mathbf{1} & 0 & 1 & 0 & 0 \\ -3 & 0 & 7 & 0 & 1 & 0 \\ -5 & 4 & -1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 2 & 1 & 0 & 1 & 0 & 0 \\ -3 & 0 & 7 & 0 & 1 & 0 \\ -13 & 0 & \mathbf{-1} & -4 & 0 & 1 \end{array}\right] \to$$

$$\to \left[\begin{array}{rrr|rrr} 2 & 1 & 0 & 1 & 0 & 0 \\ \mathbf{-94} & 0 & 0 & -28 & 1 & 7 \\ 13 & 0 & 1 & 4 & 0 & -1 \end{array}\right] \to \left[\begin{array}{rrr|rrr} 0 & 1 & 0 & \frac{19}{47} & \frac{1}{47} & \frac{7}{47} \\ 1 & 0 & 0 & \frac{14}{47} & -\frac{1}{94} & -\frac{7}{94} \\ 0 & 0 & 1 & \frac{6}{47} & \frac{13}{94} & -\frac{3}{94} \end{array}\right].$$

Обратная матрица получается из *V*, которая размещена в трех последних столбцах, перестановкой первой и второй строк (или столбцов) в соответствии с матрицей *S*, которая расположена в первых трех столбцах.

6. Разбивка на блоки. Установим связь между блоками матриц *S* и S^{-1} *n*-го порядка при разбиении

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix}; \qquad S^{-1} = \begin{bmatrix} K & L \\ M & N \end{bmatrix},$$

где A и K — квадратные матрицы порядка p; D и N — квадратные матрицы порядка q, причем $p+q=n$. Согласно определению $SS^{-1}=1$, что после перемножения матриц приводит к уравнениям:

$$AK + BM=1; \quad AL + BN = 0;$$
$$CK + DM = 0; \quad CL + DN =1.$$

Умножив первую пару уравнений на A^{-1} и решив их относительно K и L, найдем:

$$K= A^{-1}- A^{-1}BM; \; L = - A^{-1}BN.$$

Подставим полученные выражения в остальные два уравнения:.

$$CA^{-1} — CA^{-1}BM + DM = 0; — CA^{-1}BM + DN=1,$$

или

$$(D — CA^{-1}B) M= — CA^{-1}; \quad (D — CA^{-1}B)N = 1,$$

откуда имеем:

$$M= — (D — CA^{-1}B)^{-1}CA^{-1}; \quad N = (D — CA^{-1}B)^{-1}.$$

Подставив значение M и N в выражения для K и L, получим соотношение, известное как *формула Фробениуса*,

$$S^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BNCA^{-1} & -A^{-1}BN \\ -NCA^{-1} & N \end{bmatrix},$$

где $N= (D — CA^{-1}B)^{-1}$.

Как видим, обращение матрицы n-го порядка требует обращения двух матриц порядков p и q (предполагается, что A — неособенная матрица). На практике удобно выполнять вычисления в следующем порядке: A^{-1}, $A^{-1}B$, CA^{-1}, $CA^{-1}B$ (контроль $C(A^{-1}B)=(C A^{-1})B)$, $D — C A^{-1}B$, после чего находим

$$N=(D — C A^{-1}B)^{-1}, \quad M=—N(CA^{-1}),$$
$$L=—(A^{-1}B)N, \quad K = A^{-1}— (A^{-1}B)M.$$

Аналогично рассуждая, можно получить другу формулу Фробениуса

$$S^{-1} = \begin{bmatrix} K & -KBD^{-1} \\ -D^{-1}CK & D^{-1} + D^{-1}CKBD^{-1} \end{bmatrix},$$

где $K = (A — BD^{-1}C)^{-1}$. Здесь предполагается, что неособенной есть матрица D.

7. Перестановка строк и столбцов. Обращение матрицы разбиением на блоки особенно удобно в тех случаях, когда матрицы A или D легко обращаются (например, диагональная или треугольная), а также когда хотя бы одна из матриц B или C нулевая. Для получения наиболее удобного разбиения можно переставить строки и столбцы исходной матрицы. Но при этом в матрице, полученной по формуле Фробениуса, необходимо переставить соответствующие столбцы и строки, после чего она и будет обратной к исходной. Справедливость такого требования легко понять, вспомнив, что процесс обращения включает в себя транспонирование матрицы.

Пусть, например, нужно найти обратную для матрицы

$$Q = \begin{bmatrix} -1 & 0 & 1 & 1 \\ 2 & -1 & -1 & 0 \\ 2 & 0 & 0 & 1 \\ 0 & 1 & 2 & 2 \end{bmatrix}.$$

Переставив взаимно первый и четвертый столбцы и разбив на блоки, получим:

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 0 & -1 & -1 & 2 \\ \hline 1 & 0 & 0 & 2 \\ 2 & 1 & 2 & 0 \end{array}\right] = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Так как матрица A диагональная, то воспользуемся первой схемой:

$$A^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \quad A^{-1}B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix};$$

$$CA^{-1} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix};$$

$$CA^{-1}B = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 3 & -4 \end{bmatrix};$$

$$D - CA^{-1}B = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ -1 & 4 \end{bmatrix};$$

$$N = (D - CA^{-1}B)^{-1} = \begin{bmatrix} -4 & 3 \\ -1 & 1 \end{bmatrix};$$

$$M = -NCA^{-1} = -\begin{bmatrix} -4 & 3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 3 \\ -1 & 1 \end{bmatrix};$$

$$L = -A^{-1}BN = -\begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} -4 & 3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix};$$

$$K = A^{-1} - A^{-1}BM = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} -2 & 3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 0 & -2 \end{bmatrix}.$$

Таким образом, получаем матрицу, которая после взаимной перестановки первой и четвертой строк является обратной для исходной

$$\left[\begin{array}{rr|rr} 2 & -2 & 3 & -2 \\ 0 & -2 & 2 & -1 \\ \hline -2 & 3 & -4 & 3 \\ -1 & 1 & -1 & 1 \end{array}\right] \to \begin{bmatrix} -1 & 1 & -1 & 1 \\ 0 & -2 & 2 & -1 \\ -2 & 3 & -4 & 3 \\ 2 & -2 & 3 & -2 \end{bmatrix} = Q^{-1}.$$

При $C = 0$ или $B = 0$ из формул Фробениуса имеем:

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{bmatrix}; \quad \begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = \begin{bmatrix} A^{-1} & 0 \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix},$$

откуда, в частичности, следует, что при таких блочных структурах матриц обратные к ним существуют лишь при условии неособенности матриц A и D. Заметим также, что приведенные выше формулы можно получить одну за одругой транспонированием и взаимной заменой матриц B и C.

8. Метод окаймления. Обращение матрицы, разбитой на блоки, сводится к обращению матриц более низкого порядка, что требует каких-либо других методов или последующего разбиения на блоки. Рассмотрим частный случай разбиения исходной матрицы

$$A = \left[\begin{array}{c|c} A_{n-1} & u_n \\ \hline v_n & a_{nn} \end{array}\right],$$

где A_{n-1} — матрица $(n$ —1)-го порядка; a_{nn}— последний диагональный элемент исходной матрицы A; u_n и v_n — соответственно последние столбец и строка без элемента a_{nn}.

Такое разбиение можно рассматривать как результат *окаймления матрицы* A_{n-1} добавлением столбца справа и строки снизу. По первой формуле Фробениуса (п. 6) имеем

$$A^{-1} = \begin{bmatrix} A_{n-1}^{-1} + \frac{1}{\alpha_n}(A_{n-1}^{-1}u_n v_n A_{n-1}^{-1}) & -\frac{1}{\alpha_n}A_{n-1}^{-1}u_n \\ -\frac{1}{\alpha_n}v_n A_{n-1}^{-1} & \frac{1}{\alpha_n} \end{bmatrix},$$

где $\alpha_n = a_{nn} - v_n A^{-1}_{n-1} u_n$ является скаляром, так как произведение строки на матрицу даст строка, а произведение строки на столбец — скаляр.

Как видно, обращение матрицы A n-го порядка требует обращения матрицы A_{n-1} $(n$ — 1)-го порядка, которое можно выполнить по той же схеме. Этот процесс продолжается до тех пор, пока не получим матрицу второго порядка, обратная которой находится непосредственно

$$A_2^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Определив A_2^{-1} по приведенной выше формуле, найдем A_3^{-1}, затем по A_3^{-1} вычислим A_4^{-1} и т.д. до $A_n^{-1} = A^{-1}$. Вычисления на k-м шаге удобно располагать по схеме, показанной на рис. 1.6 (последовательность операций указана цифрами в скобках, причем осуществляется контроль $z_k = y_k u_k = v_k x_k$).

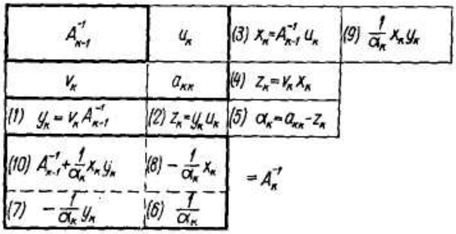


Рис. 1.6. Схемы обращения матрицы методом окаймления (k-й шаг)

Если в процессе вычисления на k-м шаге окажется, что α_k — 0, то оставшиеся строки (или столбцы) матрицы A переставляются. При этом в результирующей матрице A^{-1} осуществляется обратная перестановка столбцов (или строк). Проиллюстрируем метод окаймления на примере из (п. 7):

$$A_2^{-1} = \begin{array}{|cc|c|c|cc|}
\hline
-1 & 0 & 1 & -1 & 1 & 0 \\
-2 & -1 & -1 & -1 & 1 & 0 \\
\hline
2 & 0 & 0 & -2 & & \\
\hline
2 & 0 & -2 & 2 & & \\
\hline
\end{array}$$

$$A_3^{-1} = \begin{array}{|cc|c|c|c|ccc|}
\hline
0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & -1 & 1 & -\frac{3}{2} \\
-1 & -1 & \frac{1}{2} & 0 & -\frac{1}{2} & 1 & -1 & \frac{3}{2} \\
\hdashline
1 & 0 & \frac{1}{2} & 1 & \frac{3}{2} & -3 & 3 & -\frac{9}{2} \\
\hline
0 & 1 & 2 & 2 & \frac{5}{2} & & & \\
\hline
1 & -1 & \frac{3}{2} & \frac{5}{2} & -\frac{1}{2} & & & \\
\hline
\end{array}.$$

$$A_4^{-1} \begin{array}{|ccc|c|}
\hline
-1 & 1 & -1 & 1 \\
0 & -2 & 2 & -1 \\
-2 & 3 & -4 & 3 \\
\hdashline
2 & -2 & 3 & -2 \\
\hline
\end{array} = A^{-1}$$

9. Обращение симметричных матриц. Так как матрица, обратная симметричной, также симметрична, то процесс обращения в таких случаях существенно упрощается. При использовании любого метода достаточно получить элементы обратной матрицы, которые расположены на главной диагонали и выше (или ниже) от нее. Остальные элементы получаются из условия симметрии.

В методе исключения (п. 4) достаточно преобразовать матрицу A к нижней (или верхней) треугольной матрицы с единичными элементами по главной диагонали. При разбиении на блоки (п.6) для симметричной матрицы $A=A^t$; $C=B^t$; $D=D^t$. Соответствующие блоки обратной матрицы характеризуются аналогичными соотношениями: $K=K^t$; $L = M^t$; $N=N^t$.

Первая формула Фробениуса приводится к виду

$$S^{-1} = \begin{bmatrix} A^{-1} + HNH^t & -HN \\ -(HN)^t & N \end{bmatrix},$$

где $N = (D — B^tA^{-1} B)^{-1}$ и $H = A^{-1}B$.

Соответствующие упрощения используются и при обращении симметричных матриц A^{-1} и (D-$B^tA^{-1}B$). В методе окаймления (п.8) для симметричной матрицы A имеют место соотношения $A_{k-1} = A^t_{k-1}$ и $v_k = u^t_k$. В связи с этим $y_k = x^t_k$, и схема вычислений соответственно упрощается.

Если обращение симметричной матрицы выполняется по стандартной процедуре, то создаются дополнительные возможности контроля полученного результата по условиям симметрии обратной матрицы. Поэтому, может быть, не следует слишком категорически настаивать на использовании упрощений, которые возникают за счет специальных свойств обращаемой матрицы.

Вследствие погрешностей вычисления матрица V, полученная в результате процедуры обращения, может недостаточно точно совпадать с обратной матрицей A^{-1}. Степень отклонения вычисленной обратной матрицы от ее точного значения определяется разностью между произведением AV и единичной матрицей, которой это произведение должно равняться для точного значения $V=A^{-1}$, т.е. $R=E—AV$. Приближенное значение обратной матрицы уточняется с помощью итерационных методов.

10. Миноры обратной матрицы. Пусть A — квадратная матрица n-го порядка и $B = (A^{-1})^t$— обратная и транспонированная к ней. Тогда минор k-го порядка M^B_k матрицы B выражается через дополнительный минор M^A_k (или алгебраическое дополнение A^A_k взаимно соответственного минора матрицы A) соотношением

$$M^B_k = \frac{1}{\Delta}(-1)^\sigma \overline{M}^A_k = \frac{1}{\Delta} A^A_k,$$

где $\Delta = \det A = \frac{1}{\det B}$; σ — сумма номеров строк и столбцов, которые участуют в образовании минора M^A_k (или минора $\overline{M}^A_k$).

Предполагается, что строки и столбцы в минорах расположены в их естественном порядке (как и в исходных матрицах). Легко показать, что B — это матрица алгебраических дополнений Δ_{ij} $(i, j = 1, 2, ..., n)$ элементов матрицы A, умноженная на обратное значение определителя $1/\Delta$. Минор k-го порядка матрицы B, образованный строками с номерами $\alpha_1, \alpha_2,.., \alpha_k$ и столбцами с номерами $\beta_1, \beta_2,.., \beta_k$, запишем в следующим образом:

$$M_k^B = \begin{vmatrix} \frac{\Delta_{\alpha_1\beta_1}}{\Delta} & \frac{\Delta_{\alpha_1\beta_2}}{\Delta} & \dots & \frac{\Delta_{\alpha_1\beta_k}}{\Delta} \\ \frac{\Delta_{\alpha_2\beta_1}}{\Delta} & \frac{\Delta_{\alpha_2\beta_2}}{\Delta} & \dots & \frac{\Delta_{\alpha_2\beta_k}}{\Delta} \\ \dots & \dots & \dots & \dots \\ \frac{\Delta_{\alpha_k\beta_1}}{\Delta} & \frac{\Delta_{\alpha_k\beta_2}}{\Delta} & \dots & \frac{\Delta_{\alpha_k\beta_k}}{\Delta} \end{vmatrix} = \frac{1}{\Delta^k} \begin{vmatrix} \Delta_{\alpha_1\beta_1} & \Delta_{\alpha_1\beta_2} & \dots & \Delta_{\alpha_1\beta_k} \\ \Delta_{\alpha_2\beta_1} & \Delta_{\alpha_2\beta_2} & \dots & \Delta_{\alpha_2\beta_k} \\ \dots & \dots & \dots & \dots \\ \Delta_{\alpha_k\beta_1} & \Delta_{\alpha_k\beta_2} & \dots & \Delta_{\alpha_k\beta_k} \end{vmatrix}.$$

Алгебраическое дополнение $A^A{}_k$ получается как минор матрицы *А,* образованный удалением строк с номерами $\alpha_1, \alpha_2, .., \alpha_k$ и столбцов с номерами $\beta_1, \beta_2, .., \beta_k$, умноженный на $(-1)^\sigma$. По аналогии с обычным алгебраическим дополнением $A^A{}_k$ называется *k-кратным алгебраическим дополнением* и обозначается через

$$\Delta_{\alpha_1\beta_1,\alpha_2\beta_2,\dots,\alpha_k\beta_k}.$$

Тогда получаем важное соотношение для миноров матрицы алгебраических дополнений элементов матрицы *А,* известное как *теорема Якоби:*

$$\begin{vmatrix} \Delta_{\alpha_1\beta_1} & \dots & \Delta_{\alpha_1\beta_k} \\ \dots & \dots & \dots \\ \Delta_{\alpha_k\beta_1} & \dots & \Delta_{\alpha_k\beta_k} \end{vmatrix} = \Delta^{k-1}\Delta_{\alpha_1\beta_1,\dots,\alpha_k\beta_k}.$$

Чаще всего используется частичный случай этой теоремы для минора второго порядка, образованного элементами пересечения строк *p* и *q* $(p < q)$ и столбцов r, s $(r < s)$:

$$\begin{vmatrix} \Delta_{pr} & \Delta_{ps} \\ \Delta_{qr} & \Delta_{qs} \end{vmatrix} = \Delta_{pr}\Delta_{qs} - \Delta_{ps}\Delta_{qr} = \Delta\Delta_{pr,\,qs}.$$

Если минор второй порядок образован строками и столбцами с одинаковыми номерами ($p = r$, $q = s$), то

$$\Delta_{rr}\Delta_{ss} - \Delta_{rs}\Delta_{sr} = \Delta\Delta_{rr,\,ss}.$$

Например, для матрицы Q из (п.7), определитель которой $\Delta = -1$, имеем

$$\begin{vmatrix} \Delta_{22} & \Delta_{23} \\ \Delta_{42} & \Delta_{43} \end{vmatrix} = \Delta_{22}\Delta_{43} - \Delta_{23}\Delta_{42} = \Delta\Delta_{22,23} =$$

$$= \Delta\Delta_{22,\,43} = \Delta\,(-1)^{2+4+2+3} \begin{vmatrix} q_{11} & q_{14} \\ q_{31} & q_{34} \end{vmatrix} = -1\,(-1)^{11} \begin{vmatrix} -1 & 1 \\ 2 & 1 \end{vmatrix} = -3.$$

Приведенные соотношения используются, например, при различных преобразованиях систем линейных уравнений.

1.6. Псевдообратная матрица

Если A — квадратная и неособенная матрица, то для нее существует обратная матрица A^{-1}. Если же A — не квадратная, а прямоугольная $m\times n$-матрица $(m\neq n)$ или квадратная, но особенная, то матрица A не имеет обратной и символ A^{-1} не имеет смысла. Однако, как будет показано дальше, для произвольной прямоугольной матрицы A существует «псевдообратная» матрица A^{+}, которая обладает некоторыми свойствами обратной матрицы и имеет важные применения при решении системы линейных уравнений. В случае, когда A — квадратная неособенная матрица, псевдообратная матрица A^{+} совпадает с обратной A^{-1}.

1. Скелетное разложение матрицы. В дальнейшем мы будем пользоваться представлением произвольной прямоугольной $m\times n$-матрицы $A = \| a_{ik} \|$ ранга r в виде произведения двух матриц B и C, имеющих соответственно размеры $m\times r$ и $r\times n$:

$$A = BC = \begin{Vmatrix} b_{11} & \dots & b_{1r} \\ b_{21} & \dots & b_{2r} \\ \dots & \dots & \dots \\ b_{m1} & \dots & b_{mr} \end{Vmatrix} \begin{Vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \dots & \dots & \dots & \dots \\ c_{r1} & c_{r2} & \dots & c_{rn} \end{Vmatrix} \quad (r = r_A). \tag{1.38}$$

Здесь ранги сомножителей B и C обязательно равны рангу произведения A, $r_B = r_C = r$. Действительно, $r \leq r_B, r_C$. Но ранги r_B и r_C не могут превосходить r, так как r — один из размеров матриц B и C. Поэтому $r_B = r_C = r$.

Для того чтобы получить разложение (1.38), достаточно в качестве столбцов матрицы B взять любые r линейно независимых столбцов матрицы A, либо любые r линейно независимых столбцов, через которые линейно выражаются столбцы матрицы A. (Мы исходим из известного положения: в матрице A ранга r есть r линейно независимых столбцов, через которые линейно (т.е. в виде линейных комбинаций с числовыми коэффициентами из данного поля) выражаются все остальные столбцы. Аналогичное утверждение имеет место и для строк. Подробнее об это будет изложено дальше). Тогда любой j-й столбец матрицы A будет линейной комбинацией столбцов матрицы B с коэффициентами c_{1j}, c_{2j}, ..., c_{rj}; эти коэффициенты и образуют j -й столбец матрицы C (j = 1, ..., n,) (совершенно так же строками матрицы C могут быть любые r строк, через которые выражаются в виде линейных комбинаций все строки матрицы A.

Тогда коэффициенты этих линейных комбинаций образуют строки матрицы B).

Поскольку матрицы B и C имеют максимально возможный ранг r, то квадратные матрицы B^*B и CC^* являются неособенными:

$$|B^*B| \neq 0, \quad |CC^*| \neq 0. \tag{1.39}$$

Действительно, пусть столбец x — произвольное решение уравнения

$$B^*Bx = 0. \tag{1.40}$$

Помножим это уравнение слева на строку x^*. Тогда $x^*B^*Bx = (Bx)^*\, Bx = 0$. Отсюда следует $Bx = 0$ и (поскольку Bx — линейная комбинация линейно независимых столбцов матрицы B; ср. с формулой (1.8)) $x = 0$. Из того, что уравнение (1.40) имеет только нулевое решение $x = 0$, вытекает, что $|B^*B| \neq 0$. Аналогично устанавливается второе неравенство (1.39). (Неравенства (1.39) также непосредственно следуют из формулы Бине-Коши. Согласно этой формуле определитель $|B^*B|$ ($|CC^*|$) равен сумме квадратов модулей всех миноров r-го порядка матрицы B (соответственно C).

Разложение (1.38) будем называть *скелетным* разложением матрицы A.

2. Существование и единственность псевдообратной матрицы. Рассмотрим матричное уравнение

$$AXA = A. \tag{1.41}$$

Если A — квадратная неособенная матрица, то это уравнение имеет единственное решение $X=A^{-1}$. Если же A — произвольная прямоугольная $m\times n$-матрица, то искомое решение X имеет размеры $n\times m$, но не определяется однозначно. В общем случае уравнения (1.41) имеет бесчисленное множество решений. Ниже будет показано, что среди этих решений есть только одно, обладающее тем свойством, что его строки и столбцы являются линейными комбинациями соответственно строк и столбцов сопряженной матрицы A^*. Именно это решение мы будем называть псевдообратной матрицей для A и обозначать через A^+.

Определение. Матрица A^+ размеров $n\times m$ называется *псевдообратной* для $m\times n$-матрицы A, если выполняются равенства

$$AA^+A=A, \tag{1.42}$$

$$A^+= UA^*=A^*V, \tag{1.43}$$

где U и V — некоторые матрицы.

Условия (1.43) означают, что строки (столбцы) матрицы A^+ являются линейными комбинациями строк (столбцов) матрицы A^*. Условия (1.43) могут быть заменены одним условием $A^+ = A^*WA^*$, где W — некоторая матрица.

Докажем сначала, что для данной матрицы A не может существовать двух различных псевдообратных матриц A^+_1 и A^+_2. Действительно, из равенств
$A A^+_1 A = A A^+_2 A = A$, $A^+_1 = U_1 A^* = A^* V_1$, $A^+_2 = U_2 A^* = A^* V_2$, полагая $D = A^+_2 — A^+_1$, $U = U_2 — U_1$, $V = V_2 — V_1$ найдем:

$$ADA = 0, \quad D = UA^* = A^*V.$$

Отсюда

$$(DA)^* DA = A^*D^*DA = A^*V^*ADA = 0$$

и, следовательно,

$$DA = 0.$$

Но тогда $DD^* = DAU^* = 0$, т.е. $D = A^+_2 — A^+_1 = 0$.

Для того чтобы установить существование матрицы A^+, мы воспользуемся скелетным разложением (1.38) и будем искать сначала псевдообратные матрицы B^+ и C^+ (из приведенного выше определения сразу следует, что если A=0, то и A^+=0. Поэтому в дальнейшем предполагается, что $A \neq 0$, и потому $r — r_A > 0$). Так как по определению должны иметь место равенства

$$BB^+B = B, \quad B^+ = \tilde{U} B^*, \tag{1.44}$$

где $\tilde{U}$ — некоторая матрица, то

$$B \tilde{U} B^* B = B.$$

Умножая слева на B^* и замечая, что B^*B — неособенная квадратная матрица, найдем:

$$\tilde{U} = (B^*B)^{-1}.$$

Но тогда второе из равенств (1.44) дает искомое выражение для B^+:

$$B^+ = (B^*B)^{-1}B^*. \tag{1.45}$$

Аналогично найдем:

$$C^+ = C^*(CC^*)^{-1} \tag{1.46}$$

Покажем теперь, что матрица

$$A^+ = C^+B^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^* \tag{1.47}$$

удовлетворяет условиям (1.42), (1.43) и, следовательно, есть псевдообратной матрицей для A.

В самом деле,

$$AA^+A = BCC^* (CC^*)^{-1} (B^*B)^{-1}B^*BC = BC = A.$$

С другой стороны, из равенств (1.45), (1.46) и (1.47) с учетом равенства $A^* = C^*B^*$, полагая $K = (CC^*)^{-1}(B^*B)^{-1}$ находим

$$A^+ = C^*KB^* = C^*K(CC^*)^{-1}CC^*B = UC^*B^* = UA^*,$$
$$A^+ = C^*KB^* = C^*B^*B(B^*B)^{-1}KB^* = C^*B^*V = A^*V,$$

где

$$U = C^*K\,(CC)^{-1}C, \quad V = B\,(B^*B)^{-1}\,KB^*.$$

Таким образом доказано, что для произвольной прямоугольной матрицы A существует одна и только одна псевдообратная матрица A^+, которая определяется формулой (1.47), где B и C — сомножители в скелетном разложении $A = BP$ матрицы A. Из самого определения псевдооберненої матрицы непосредственно следует, что в случае квадратной неособенной матрицы A псевдообратная матрица A^+ совпадает из обратной A^{-1}.

Пример. Пусть

$$A = \begin{Vmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}.$$

Здесь $r = 2$. Примем в качестве столбцов матрицы B первые два столбца матрицы A. Тогда

$$A = BC = \begin{Vmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{Vmatrix} \begin{Vmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}$$

и

$$B^*B = \begin{Vmatrix} 2 & -3 \\ -3 & 6 \end{Vmatrix}, \quad (B^*B)^{-1} = \begin{Vmatrix} 2 & 1 \\ 1 & {}^2/_3 \end{Vmatrix},$$

$$CC^* = \begin{Vmatrix} 3 & 0 \\ 0 & 3 \end{Vmatrix}, \quad (CC^*)^{-1} = \begin{Vmatrix} {}^1/_3 & 0 \\ 0 & {}^1/_3 \end{Vmatrix} = \frac{1}{3}E.$$

Поэтому согласно формуле (1.47)

$$A^+ = \frac{1}{3}\begin{Vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & 1 \end{Vmatrix} \begin{Vmatrix} 2 & 1 \\ 1 & {}^2/_3 \end{Vmatrix} \cdot \begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{Vmatrix} = \begin{Vmatrix} {}^1/_3 & 0 & {}^1/_3 \\ {}^1/_9 & {}^1/_9 & {}^2/_9 \\ {}^2/_9 & -{}^1/_9 & {}^1/_9 \\ {}^4/_9 & {}^1/_9 & {}^5/_9 \end{Vmatrix}.$$

3. Свойства псевдообратной матрицы. Отметим следующие свойства псевдообратной матрицы:

1° $(A^*)^+=(A^+)^*$;

2° $(A^+)^+=A$;

3° $(AA^+)^* =AA^+$, $(AA^+)^2 = AA^+$;

4° $(A^+A)^* = A^+A$, $(A^+A)^2 = A^+A$.

Первое свойство означает, что операции перехода к сопряженной и к псевдообратной матрицы перестановочны между собой. Равенство 2° выражает собой взаимность понятия псевдообратной матрицы, так как согласно 2° псевдообратной матрицей для A^+ является исходная матрица A. Согласно равенствам 3° и 4° матрицы AA^+ и A^+A являются эрмитовыми и *инволютивными* (квадрат каждой из этих матриц равен самой матрице).

Для вывода равенства 1° воспользуемся скелетным разложением (1.38): $A=BC$. Тогда равенство $A^*=C^*B^*$ дает скелетное разложение матрицы A^*. Поэтому, заменяя в формуле (1.47) матрицу B на C^*, а матрицу C на B^*, получим:

$$(A^*)^+ = B(B^*B)^{-1}(CC^*)^{-1}C = [C^*(CC^*)^{-1}(B^*B)^{-1}B^*]^* = (A^+)^*.$$

Равенства $A^+=C^+B^+$, $B^+=(B^*B)^{-1}CB^*$, $C^+=C^*(CC^*)^{-1}$ являются скелетными разложениями. Следовательно,

$$(A^+)^+ = (B^+)^+ (C^+)^+ = (B^*)^+ B^*BCC^* (C^*)^+.$$

Используя свойство 1°, а также выражения для B^+ и C^+, найдем:

$$(A^+)^+ = B(B^*B)^{-1} B^*BCC^* (CC^*)^{-1} C = BC = A.$$

Справедливость равенств 3° и 4° проверяется непосредственно путем подстановки в эти равенства вместо A^+ соответствующего выражения из формулы (1.47). :

Заметим, что в общем случае, когда разложение $A = BC$ не является скелетным, не всегда имеет место равенство $A^+=C^+B^+$. Так, например,

$$A = \|1\| = \|0 \quad 1\| \begin{Vmatrix} 1 \\ 1 \end{Vmatrix} = BC.$$

Здесь

$$A^+ = A^{-1} = \|1\|, \quad B^+ = \|0 \quad 1\|^+ = (\|1\| \cdot \|0 \quad 1\|)^+ = \begin{Vmatrix} 1 \\ 0 \end{Vmatrix} \|1\| = \begin{Vmatrix} 1 \\ 0 \end{Vmatrix},$$

$$C^+ = \begin{Vmatrix} 1 \\ 1 \end{Vmatrix}^+ = \left(\begin{Vmatrix} 1 \\ 1 \end{Vmatrix} \cdot \|1\|\right)^+ = \|1\| \cdot \|2\|^{-1} \|1 \quad 1\| = \|{}^1/_2 \quad {}^1/_2\|.$$

Поэтому

$$C^+B^+ = \|{}^1/_2 \quad {}^1/_2\| \begin{Vmatrix} 0 \\ 1 \end{Vmatrix} = \|{}^1/_2\| \neq A^+.$$

4. Наилучшее приближенное решение (по методу наименьших квадратов). Рассмотрим произвольную систему линейных уравнений

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= y_m \end{aligned} \tag{1.48}$$

или в матричной записи

$$Ax = y. \tag{1.49}$$

Здесь $y_1, y_2, \ldots, y_m$ — заданные числа, а $x_1, x_2, \ldots, x_n$ — искомые.

В общем случае система (1.48) может быть и несовместной

Столбец

$$x^0 = (x_1^0, x_2^0, \dots, x_n^0) \tag{1.50}$$

называется *наилучшим приближенным решением* системы (1.48), если при значениях $x_1=x^0_1$, $x_2=x^0_2,\dots,x_n=x^0_n$ «квадратичное отклонение»

$$|y - Ax|^2 = \sum_{i=1}^{m} \left| y_i - \sum_{k=1}^{n} a_{ik}x_k \right|^2 \tag{1.51}$$

достигает своего наименьшего значения и среди всех столбцов x, для которых это отклонение имеет минимальное значение, столбец x^0 имеет наименьшую «длину», т.е. для этого столбца величина

$$|x|^2 = x^*x = \sum_{i=1}^{n} |x_i|^2 \tag{1.52}$$

имеет наименьшее значение.

Покажем, что система (1.48) всегда имеет одно и только одно наилучшее приближенное решение и это приближенное решение определяется по формуле

$$x^0 = A^+y, \tag{1.53}$$

где A^+ — псевдообратная матрица для матрицы A.

Для этого рассмотрим произвольный столбец x и положим

$$y - Ax = u + v,$$

где

$$u = y - Ax^0 = y - AA^+y, \quad v = A(x^0-x). \tag{1.54}$$

Тогда

$$|y - Ax|^2 = (y - Ax)^*(y - Ax) = \\ = (u + v)^*(u + v) = u^*u + v^*u + u^*v + v^*v. \tag{1.55}$$

Но -

$$u^*v = (x\ 0\text{-}^x)^*A^*(y - AA^+y) = (x^0-x)^*(A^* - A^*AA^+)\,y. \tag{1.56}$$

Исходя из разложения (1.38) и формулы (1.47), найдем:

$$A^*AA^+ = C^*B^*BCC^*(CC^*)^{-1}(B^*B)^{-1}B^* = C^*B^* = A^*.$$

Поэтому из равенства (1.56) следует

$$u^*v=0, \tag{1.57}$$

но тогда и

$$u^*v = (v^*u)^* = 0. \tag{1.58}$$

Поэтому из равенства (1.55) находим:

$$|y - Ax|^2 = |u|^2 + |v|^2 = |y - Ax^0|^2 + |A(x^0 - x)|^2, \tag{1.59}$$

и, следовательно, для любого столбца x

$$|y\text{-}ax| \geq |y\text{-}ax^0|. \tag{1.60}$$

Пусть теперь

$$|y - Ax| = |y - Ax^0|;$$

тогда согласно равенству (1.59)

$$Az = 0, \qquad (1.61)$$

где

$$z = x - x^0.$$

С другой стороны,

$$|x|^2 = (x^0 + z)^*(x^0 + z) = |x^0|^2 + |z|^2 + (x^0)^* z + z^* x^0 \qquad (1.62)$$

Вспоминая, что $A^+ = A^*V$ (см. определение 5), получим в силу (1.61):

$$(x^0)^* z = (A^+y)^* z = (A^*Vy)^* z = y^*V^*Az = 0. \qquad (1.63)$$

Но тогда и

$$z^*x^0 = ((x^0)^* z)^* = 0.$$

Поэтому из равенства (1.62) находим:

$$|x|^2 = |x^0|^2 + |z|^2,$$

и, следовательно,

$$|x|^2 \geq |x^0|^2,$$

причем знак = имеет место только при z = 0, т.е. при $x = x^0$, где $x^0 = A^+y$.

Пример. Найти наилучшее приближенное решение (по методу наименьших квадратов) системы линейных уравнений:

$$\begin{aligned} x_1 - x_2 + 2x_3 \qquad &= 3, \\ -x_1 + 2x_2 - 3x_3 + x_4 &= 6, \\ x_2 - x_3 + x_4 &= 0. \end{aligned}$$

Здесь

$$A = \begin{Vmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{Vmatrix}.$$

Но тогда

$$A^+ = \begin{Vmatrix} 1/3 & 0 & 1/3 \\ 1/9 & 1/9 & 2/9 \\ 2/9 & -1/9 & 1/9 \\ 4/9 & 1/9 & 5/9 \end{Vmatrix},$$

и потому

$$x^0 = \begin{Vmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \\ x_4^0 \end{Vmatrix} = \begin{Vmatrix} 1/3 & 0 & 1/3 \\ 1/9 & 1/9 & 2/9 \\ 2/9 & -1/9 & 1/9 \\ 4/9 & 1/9 & 5/9 \end{Vmatrix} \begin{Vmatrix} 3 \\ 6 \\ 0 \end{Vmatrix}.$$

Следовательно,

$$x^0_1=1,\quad x^0_2=1,\quad x^0_3=0,\quad x^0_4=2.$$

Определим норму $\|A\|$ $m\times n$-матрицы $A=\|a_{ik}\|$ как неотрицательное число, которое задается формулой

$$\| A \|^2 = \sum_{i,k} | a_{ik} |^2. \tag{1.65}$$

При этом очевидно, что

$$\| A \|^2 = \sum_{k=1}^{n} | A_{\cdot k} |^2 = \sum_{i=1}^{n} | A_{i \cdot} |^2. \tag{1.66}$$

Рассмотрим матричное уравнение

$$AX = Y, \tag{1.67}$$

где A и Y — заданные $m\times n$- и $m\times p$-матрицы, а X — искомая $m\times p$-матрица.

Определим наилучшее приближенное решение X^0 уравнение (1.67) из условия

$$\| Y-AX^0\| = \min\|Y-AX\|,$$

причем в случае, когда

$$\|Y-AX\| = \|Y-AX^0\|,$$

требуется, чтобы

$$\| X^0\| \le \|X\|$$

Из соотношений

$$\| Y - AX \|^2 = \sum_{k=1}^{p} | Y_{\cdot k} - AX_{\cdot k} |^2, \tag{1.68}$$

$$\| X \|^2 = \sum_{k=1}^{p} | X_{\cdot k} |^2 \tag{1.69}$$

следует, что k-й столбец искомой матрицы $X^0_{\cdot k}$ должен быть наилучшим приближенным решением системы линейных уравнений

$$AX_{\cdot k} = Y_{\cdot k}.$$

Поэтому

$$X^0_{\cdot k} = A^+Y_{\cdot k}.$$

Поскольку это равенство справедливо при любом $k = 1, \ldots, p$, то

$$X^0 = A^+Y. \tag{1.70}$$

Таким образом, уравнение (1.67) всегда имеет одно и только одно наилучшее приближенное решение, определяемое формулой (1.70).

В частном случае, когда $Y = E$ — единичная матрица m-го порядка, имеем $X^0=A^+$. Следовательно, *псевдообратная матрица A^+ является наилучшим приближенным решением (по методу наименьших квадратов) матричного уравнения*

$$AX=E.$$

Это свойство псевдообратной матрицы A^+ может быть принято в качестве ее определения.

5. Метод Гревилля последовательного нахождения псевдообратной матрицы заключается в следующем. Пусть a_k — k-й столбец в $m\times n$-матрице A, $A_k = (a_1, ...,a_k)$ — матрица, которая образована первыми k столбцами матрицы A. b_k — последняя строка в матрице A^+_k ($k = 1, ...,n$, $A_1 = a_1$, $A_n= A$). Тогда (если $A_1 = a_1 = 0$, то и $A^+_1 = 0$)

$$A_1^+ = a_1^+ = \frac{a_1^*}{a_1^* a_1} \tag{1.71}$$

и для $k > 1$ имеют место рекуррентные формулы

$$A_k^+ = \begin{pmatrix} B_k \\ b_k \end{pmatrix}, \quad B_k = A_{k-1}^+ - d_k b_k, \quad d_k = A_{k-1}^+ a_k; \tag{1.72}$$

при этом, если $c_k = a_k — A_{k-1}d_k \neq 0$, то

$$b_k = c^+_k = (a_k - A_{k-1}d_k)^+; \tag{1.73}$$

если же $c_k = 0$, т.е. $a_k = A_{k-1}d_k$, то

$$b_k = (1 + d_k^* d_k)^{-1}\, d_k^* A_{k-1}^+. \tag{1.74}$$

Пример. Пусть

$$A = \begin{Vmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & -3 & -1 \\ 0 & 1 & 1 \end{Vmatrix}.$$

Заметим, что для каждой вещественной матрицы M мы можем писать M' вместо M^*. Тогда

$$A_1^+ = (A_1' A_1)^{-1} A_1' = \frac{1}{6} A_1' = \| {}^1/_6,\ -{}^1/_6,\ {}^1/_3,\ 0 \|,$$

$$d_2 = A_1^+ a_2 = -{}^3/_2, \qquad c_2 = a_2 - A_1 d_2 = \begin{Vmatrix} {}^1/_2 \\ {}^1/_2 \\ 0 \\ 1 \end{Vmatrix},$$

$$b_2 = c_2^+ = (c_2' c_2)^{-1}\, c_2' = \frac{2}{3} c_2' = \| {}^1/_3,\ {}^1/_3,\ 0,\ {}^2/_3 \|,$$

$$B_2 = A_1^+ - d_2 b_2 = \| {}^2/_3, {}^1/_3, {}^1/_3, 1 \|.$$

Таким образом

$$A_2^+ = \begin{Vmatrix} 2/3 & 1/3 & 1/3 & 1 \\ 1/3 & 1/3 & 0 & 2/3 \end{Vmatrix}.$$

Далее

$$d_3 = A_2^+ a_3 = \begin{Vmatrix} 1 \\ 1 \end{Vmatrix} \quad \text{и} \quad c_3 = a_3 - A_2 d_3 = 0.$$

Поэтому

$$b_3 = (1 + d_3' d_3)^{-1} d_3' A_2^+ = \| 1/3, \ \ 1/3 \| A_2^+ = \| 1/3 \ \ 2/9 \ \ 1/9 \ \ 5/9 \|$$

и

$$B_3 = A_2^+ - d_3 b_3 = \begin{Vmatrix} 2/3 & 1/3 & 1/3 & 1 \\ 1/3 & 1/3 & 0 & 2/3 \end{Vmatrix} - \begin{Vmatrix} 1/3 & 2/9 & 1/9 & 5/9 \\ 1/3 & 2/9 & 1/9 & 5/9 \end{Vmatrix} = \begin{Vmatrix} 1/3 & 1/9 & 2/9 & 4/9 \\ 0 & 1/9 & -1/9 & 1/9 \end{Vmatrix},$$

$$A^+ = A_3^+ = \begin{Vmatrix} 1/3 & 1/9 & 2/9 & 4/9 \\ 0 & 1/9 & -1/9 & 1/9 \\ 1/3 & 2/9 & 1/9 & 5/9 \end{Vmatrix}.$$

<u>Микромодуль 3</u>
Индивидуальные тестовые задачи

Проверить следующие свойства матриц H и F:

1. *В результате умножения произвольной $m \times n$-матрицы A слева на матрицу H (матрицу F) n-го порядка все строки матрицы A поднимаются (опускаются) на одно место вверх (вниз), первая (последняя) строка матрицы A исчезает, а последняя (первая) строка произведения заполняется нулями.* Так, например,

$$\begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{Vmatrix} \begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} = \begin{Vmatrix} b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 0 \end{Vmatrix},$$

$$\begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix} \begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{Vmatrix}.$$

2. *В результате умножения произвольной $m \times n$-матрицы A справа на матрицу $H(F)$ n-го порядка все столбцы матрицы A сдвигаются вправо (влево) на одно место, при этом последний (первый) столбец матрицы A исчезает, а первый (последний) столбец произведения заполняется нулями.* Так, например,

$$\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} \begin{Vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & b_1 & b_2 & b_3 \\ 0 & c_1 & c_2 & c_3 \end{Vmatrix},$$

$$\begin{Vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{Vmatrix} \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{Vmatrix} = \begin{Vmatrix} a_2 & a_3 & a_4 & 0 \\ b_2 & b_3 & b_4 & 0 \\ c_2 & c_3 & c_4 & 0 \end{Vmatrix}.$$

3. С помощью метода исключения найдите обратные для следующих матриц:

$$\text{а)} \begin{bmatrix} 2 & 1 & 0 \\ 3 & 0 & 5 \\ 7 & 6 & 4 \end{bmatrix}; \quad \text{б)} \begin{bmatrix} 0 & 0 & 1 & -1 \\ 0 & 3 & 1 & 4 \\ 2 & 7 & 6 & -1 \\ 1 & 2 & 2 & -1 \end{bmatrix}.$$

Вычислите определители и найдите присоединенные матрицы на основе данных, полученных и процессе исключения.

4. Для матрицы

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 0 & 5 \\ 7 & 6 & 4 \end{bmatrix}$$

образуйте матрицы V_k (k=1, 2, 3), соответствующие преобразованию A к единичной матрице, и найдите $A^{-1}=V_1V_2V_3$. Проверьте результат вычислением A^{-1} через определитель и алгебраические дополнения.

5. Найдите обратную для матрицы

$$\begin{bmatrix} 1 & 1 & 3 & 2 \\ 2 & 2 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{bmatrix}$$

а) методом исключения;

б) разбиением на блоки;

в) методом окаймления.

6. Обратите наиболее рациональным способом матрицу

$$\begin{bmatrix} 3 & 1 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ 0 & 0 & 4 & -6 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

7. Запишите выражение для матрицы, обратной треугольной третьего порядка

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix},$$

и обобщите результат на матрицу n-го порядка.

8. Покажите, что определитель $|G|$ блочной матрицы

$$G = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

можно выразить такими способами:

а) $| G | = |AD — ACCA^{-1}B | = | A | | D— CA^{-1}B|$, если $| A | \neq 0$;

б) $| G | =|AD — BD^{-1}CD | = | A — BD^{-1}C| | D |$, если $| D | \neq 0$;

в) $| G | =|AD — CB|$, если A и C перестановочны $(ACC = CA)$;

г) $| G | = | AD — BC |$, если C и D перестановочны $(CD = DC)$.

8. Дана матрица

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 3 & 4 & 2 \\ 1 & 3 & 3 & 2 \\ 4 & 6 & 7 & 1 \end{bmatrix}.$$

По теореме Якоби определите следующие миноры матрицы алгебраических дополнений:

$$\text{а) } \begin{vmatrix} \Delta_{21} & \Delta_{24} \\ \Delta_{31} & \Delta_{34} \end{vmatrix}; \quad \text{б) } \begin{vmatrix} \Delta_{11} & \Delta_{13} & \Delta_{14} \\ \Delta_{21} & \Delta_{23} & \Delta_{24} \\ \Delta_{41} & \Delta_{43} & \Delta_{44} \end{vmatrix}.$$

9. Проверить, что матрица

$$\begin{pmatrix} B_k \\ b_k \end{pmatrix}$$

является псевдообратной для матрицы $A^{+}{}_{k}$, если матрица B_k и строка b_k определяются формулами (1.65)— (1.69). Этот метод не требует вычисления детерминантов и может быть использован для вычисления обратной матрицы.

Модуль 2
Определители и уравнения

Микромодуль 4
Определители

2.1. Определители и графы матриц

1. Определитель матрицы. Понятие определителя (детерминанта) возникло в связи с решением систем линейных уравнений и благодаря ему эта задача получила компактное выражение, например, в виде правила Крамера. Представление таких систем в матричной форме $Ax=q$ естественным образом связывает квадратную матрицу A с ее определителем $\det A$ (или $|A|$). Общее выражение для определителя матрицы n-го порядка обычно дается в виде:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \Sigma (-1)^{\varepsilon} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}.$$

В правой части стоит сумма произведений вида $a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}$. Каждое такое произведение по определению должно содержать элементы матрицы a_{ij}, которые расположены в различных строках и различных столбцах. Это значит, что среди всех первых индексов, как и среди всех вторых индексов не должно быть одинаковых.

Если расположить первые индексы в порядке их возрастания, как это сделано выше, то совокупность вторых индексов образует некоторую перестановку $(\alpha_1, \alpha_2, \dots, \alpha_n)$ множества чисел от 1 до n. Иначе говоря, каждое произведение под знаком суммы определяется подстановкой n-й степени:

$$\sigma = \begin{pmatrix} 1, & 2, \dots, & n \\ \alpha_1, & \alpha_2, \dots, & \alpha_n \end{pmatrix}$$

Так как число всех подстановок n-й степени равно $n!$, то можно образовать такое же количество произведений $a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}$ из элементов данной матрицы (при нулевых элементах некоторые из них равны нулю). Определитель равен сумме всех таких произведений, взятых со знаком $(-1)^{\varepsilon}$, где ε — число инверсий перестановки $(\alpha_1, \alpha_2, ..., \alpha_n)$. Вместо множителя $(-1)^{\varepsilon}$ можно писать знак подстановки $\mathrm{sgn}\sigma$, который положительный для четной и отрицательный для нечетной подстановки σ.

Порядок определителя совпадает с порядком его матрицы. Элементы a_{ij} матрицы A называют также *элементами определителя* $|A|$, а произведение $(-1)^{\varepsilon} a_{1\alpha_1} a_{2\alpha_2} ... a_{n\alpha_n}$ - членами *определителя.*

Для определителей второго и третьего порядка получаем выражения, которые совпадают с хорошо известными схемами вычисления этих определителей (рис. 2.1):

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21};$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

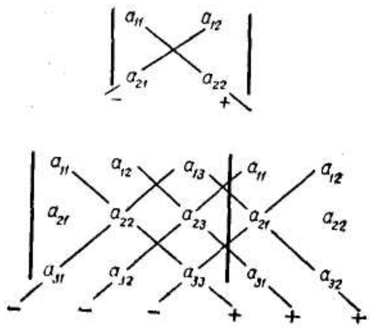


Рис. 2.1. Схемы вычисления определителя второго и третьего порядков.

Как видно, индексы столбцов всех членов определителя третьего порядка определяются перестановками (1, 2, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1), (1, 3, 2), (2, 1, 3), число инверсий которых равно соответственно 0, 2, 2, 3, 1, 1.

Общее выражение определителя n-го порядка является удобным для исследования и доказательства его свойств, но для вычисления определителей используются другие более практичные соотношения и методы.

2. Граф матрицы. Квадратной матрице n-го порядка можно поставить в соответствие направленный граф, который строится на множестве n вершин. При этом ij-элементу матрицы соответствует дуга, которая выходит из i-й вершины и входит в j-ю, и этой дуге приписывается вес, который равен значению a_{ij} элемента. Каждая пара вершин такого графа связана двумя дугами в прямом и обратном направлениях, которые отображают элементы a_{ij} и a_{ji}, а петли соответствуют элементам a_{ii} $(i = 1, 2, ..., n)$ главной диагонали (рис. 2.2, *а*).

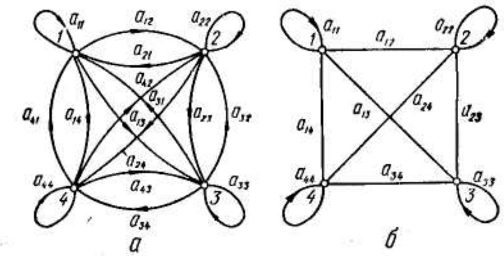


Рис. 2.2. Граф матрицы *(а)* и граф симметричной матрицы (*б*).

Граф можно упростить, если условиться симметричные и равные элементы $a_{ij}=a_{ji}$ изображать одной дугой без указания ее направления, котораяе заменяет две противоположно направленные дуги (граф симметричной матрицы показан на рис. 2.2, *б*). Дальнейшее упрощение достигается, если дуги нулевых элементов на графе не изображать (но наличие их обязательно подразумевается). Например, граф на рис. 2.3, *а* соответствует матрице:

$$A = \begin{array}{|c|c|c|c|}\hline 1 & 2 & & 3 \\ \hline & -3 & 4 & 1 \\ \hline & & & 5 \\ \hline -2 & 1 & -1 & \\ \hline \end{array}.$$

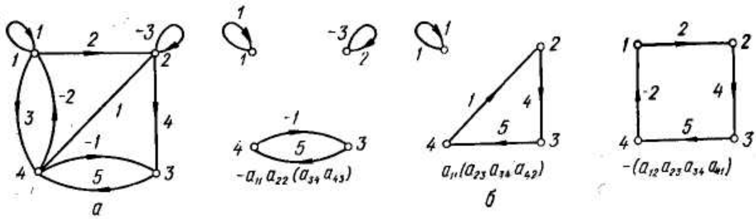


Рис. 2.3. Граф матрицы четвертого порядка (*а*) и его факторы (*б).*

Граф матрицы позволяет наглядно представить выражение для ее определителя. В рассмотренном примере из 4!=16 членов определителя $|A|$ ненулевых только три, которые определяются подстановками:

$$\sigma_1 = \begin{pmatrix} 1\,2\,3\,4 \\ 1\,2\,4\,3 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 1\,2\,3\,4 \\ 1\,3\,4\,2 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,3\,4\,1 \end{pmatrix}.$$

Подстановки σ_1 и σ_3 нечетные, а σ_2 четная, следовательно, имеем:

$$\det A = \begin{vmatrix} 1 & 2 & 0 & 3 \\ 0 & -3 & 4 & 1 \\ 0 & 0 & 0 & 5 \\ -2 & 1 & -1 & 0 \end{vmatrix} = -a_{11}a_{22}a_{34}a_{43} + a_{11}a_{23}a_{34}a_{42} - -a_{12}a_{23}a_{34}a_{41} = 85.$$

Разложив подстановку на независимые циклы и построив ее граф, замечаем, что соответствующий член определителя равен произведению весов всех дуг графа подстановки, а его знак определяется четностью декремента (разности между степенью и количеством независимых циклов подстановки). Так, для данного примера σ_1 = (1) (2)(3, 4); σ_2=(1) (2, 3, 4); σ_3 = (1, 2, 3, 4) и соответственно декременты $d_1 = 4 — 3=1$; $d_2= 4 — 2 = 2$ и $d_3 = 4 — 1=3$. Взвешенные графы подстановок изображены на рис. 2.3, *б*. Каждый из них представляет собой совокупность несоприкасающихся контуров и петель, инцидентную всем вершинам, и определяет соответствующий член определителя.

Совокупность всех ненулевых членов определителя можно получить из графа матрицы, выделяя всевозможные подграфы, которые включают все вершины и состоят исключительно из несоприкасающихся контуров и петель. Такие подграфы называют *факторами* графа. Произведение весов дуг фактора графа дает соответствующий член определителя, а четность разности между числом вершин (порядком матрицы) и числом контуров (циклов подстановки) фактора определяет знак этого члена.

3. Основные свойства определителей. Прежде всего отметим, что $\det A = \det A^t$, т.е. определитель матрицы не изменяет своего значения при взаимной замене ее строк и столбцов. Поэтому все свойства определителя, которые сформулированы для столбцов, справедливы и для строк, и обратно. Ниже приводятся основные свойства определителей, которые легко доказываются на основе общего выражения (п.1).

1. При перестановке двух столбцов определитель меняет знак *(свойство антисимметрии).*

2. Определитель равен нулю, если все элементы какого-нибудь столбца равны нулю или если один из столбцов является линейной комбинацией любых его других столбцов (в частности, определитель, у которого хотя бы два столбца одинаковы, равен нулю).

3. Умножение всех элементов какого-нибудь столбца на скаляр λ равнозначно умножению определителя на λ, (общий множитель элементов строки или столбца можно вынести за знак определителя).

4. Умножение матрицы n-го порядка на скаляр λ соответствует умножению ее определителя на λ^n, т.е. $\det(\lambda A) = \lambda^n \det A$.

5. Значение определителя не изменится, если к какому-нибудь столбцу прибавить другой столбец, умноженный на скаляр λ.

6. Если два определителя одинаковых порядков различаются между собой только элементами i-го столбца, то их сумма равна определителю, элементы i-го столбца которого равні суммам соответствующих элементов i-х столбцов исходных определителей, а другие элементы те же, что в исходных *(свойство линейности).*

4. Миноры и алгебраические дополнения. Пусть в определителе n-го порядка Δ выделені k различных строк $(k \leq n)$ с номерами $i_1, i_2, \ldots, i_k$ и столько же различных столбцов с номерами $j_1, j_2, \ldots, j_k$. Элементы, которые расположены на пересечении этих строк и столбцов, образуют определитель, который называется *минором k-го порядка* и обозначается

$$M_k = M\begin{pmatrix} i_1, & \ldots, & i_k \\ j_1, & \ldots, & j_k \end{pmatrix}.$$

Если удалить из определителя строки и столбцы, которые принимают участие в построении минора M_k, то элементы, которые остались, образуют *определитель* $(n-k)$-го порядка, который *называется дополнительным* минором к M_k и обозначается

$$\overline{M}_k = \overline{M}\begin{pmatrix} i_1, & \ldots, & i_k \\ j_1, & \ldots, & j_k \end{pmatrix}.$$

Алгебраическим дополнением минора k-го порядка M_k называют величину $A_k = (-1)^{\sigma}\overline{M}_k$, где

$$\sigma = (i_1 + i_2 + \ldots + i_k) + (j_1 + j_2 + \ldots + j_k)$$

- сумма номеров строк и столбцов, которые определяют минор M_k. В частности, миноры первого порядка совпадают с элементами определителя, поэтому алгебраические дополнения A_1 первого порядка называют также алгебраическими дополнениями элементов определителя или матрицы. При этом

$$\Delta_{ij} = A_{ij} = (-1)^{i+j}\ \overline{M}\begin{pmatrix} i \\ j \end{pmatrix}.$$

Дополнительный минор $(n-1)$-го порядка $\overline{M}\begin{pmatrix} i \\ j \end{pmatrix}$ часто называют просто *минором* и обозначают M_{ij}.

Минор нулевого порядка считается равным единицы, т.е. $M_0=1$, при этом дополнительный к нему минор и алгебраическое дополнение совпадают с определителем Δ. Минор n-го порядка M_n совпадает с определителем Δ, в то время как дополнительный к нему минор и алгебраическое дополнение считаются равными единицы. Сказанное выражается соотношениями:

$$M_n = \overline{M}_0 = A_0 = \Delta; \quad M_0 = \overline{M}_n = A_n = 1$$

Миноры, которые образованы строками и столбцами с одинаковыми номерами, называют *главными* (их диагональные элементы являются и диагональными элементами определителя). Очевидно, определитель n-го порядка имеет C^m_n главных миноров m-го порядка, а всего 2^n главных миноров всех возможных порядков (от 0 до n).

Если речь идет об определителе матрицы A, то его миноры n-го порядка обозначают:

$$M_k = A\begin{pmatrix} i_1, & \dots, & i_k \\ j_1, & \dots, & j_k \end{pmatrix}.$$

Пусть имеется две или несколько матриц $A_1, A_2, \dots, A_m$ одинаковых порядков. Будем называть миноры k-го порядка этих матриц $M^{A_1}_{k}, M^{A_2}_{k}, \dots, M^{A_m}_{k}$ *взаимно соответственными,* если они образованы из данных определителей выделением строк и столбцов с теми самыми номерами в каждой матрице. Ясно, что миноры $\bar{M}^{A_1}_{k}, \bar{M}^{A_2}_{k}, \dots, \bar{M}^{A_m}_{k}$, дополнительные к взаимно соответствующим, также есть взаимно соответствующими.

5. Разложение определителя. Определитель Δ n-го порядка выражается через элементы произвольной его строки или столбца следующим образом:

$$\Delta = \sum_{s=1}^{n} a_{is}\Delta_{is} = \sum_{s=1}^{n} a_{si}\Delta_{si} \quad (i = 1,\ 2,\ \dots,\ n),$$

где Δ_{ij} — алгебраическое дополнение элемента a_{ij}. Эти соотношения позволяют представить определитель n-го порядка через определители $(n—1)$-го порядка. При вычислениях целесообразно раскладывать определитель по строке или столбцу, которые имеют большее количество нулевых элементов. Например, раскладывая данный определитель по элементам второго столбца, имеем:

$$\Delta = \begin{vmatrix} -2 & 3 & 2 & 0 \\ 5 & 0 & -3 & 2 \\ 4 & 0 & 1 & 5 \\ 1 & 6 & 2 & 1 \end{vmatrix} = -3\begin{vmatrix} 5 & -3 & 2 \\ 4 & 1 & 5 \\ 1 & 2 & 1 \end{vmatrix} + 6\begin{vmatrix} -2 & 2 & 0 \\ 5 & -3 & 2 \\ 4 & 1 & 5 \end{vmatrix}.$$

Раскладывая определители третьего порядка, получаем:

$$\begin{vmatrix} 5 & -3 & 2 \\ 4 & 1 & 5 \\ 1 & 2 & 1 \end{vmatrix} = 5\begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} - 4\begin{vmatrix} -3 & 2 \\ 2 & 1 \end{vmatrix} + 1\begin{vmatrix} -3 & 2 \\ 1 & 5 \end{vmatrix} =$$
$$= 5(-9)(-4(-7) + 1(-17) = -34.$$

$$\begin{vmatrix} -2 & 2 & 0 \\ 5 & -3 & 2 \\ 4 & 1 & 5 \end{vmatrix} = -2\begin{vmatrix} -3 & 2 \\ 1 & 5 \end{vmatrix} - 2\begin{vmatrix} 5 & 2 \\ 4 & 5 \end{vmatrix} = -2(-17) - 2\cdot 17 = 0.$$

Второй определитель оказался равным нулю, так как его третий столбец равен сумме первого и второго (свойство 2). Таким образом, находим $\Delta = -3(-34) + 6\cdot 0 = 102$.

Обобщением изложенного метода является *разложение Лапласа* по нескольким строкам или столбцам. Пусть заданы любые k строк (или столбцов) определителя Δ. Тогда этот определитель можно представить как сумму произведений всевозможных миноров k-го порядка, расположенных в этих строках (или столбцах) на их алгебраические дополнения, т.е.

$$\Delta = \Sigma M_k A_k = \Sigma (-1)^\sigma M_k \bar{M}_k,$$

где σ — сумма номеров строк и столбцов, которые принимают участие в формировании минора M_k (или $\bar{M}_k$). Очевидно, число слагаемых в этой сумме равно C_n^k (некоторые из них могут равняться нулю). Например, раскладывая приведенный выше определитель по первому и второму столбцам, имеем:

$$\begin{vmatrix} -2 & 3 & 2 & 0 \\ 5 & 0 & -3 & 2 \\ 4 & 0 & 1 & 5 \\ 1 & 6 & 2 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 3 \\ 5 & 0 \end{vmatrix}\begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} -2 & 3 \\ 4 & 0 \end{vmatrix}\begin{vmatrix} -3 & 2 \\ 2 & 1 \end{vmatrix} + \begin{vmatrix} -2 & 3 \\ 1 & 6 \end{vmatrix}\begin{vmatrix} -3 & 2 \\ 1 & 5 \end{vmatrix} +$$

$$+ \begin{vmatrix} 5 & 0 \\ 4 & 0 \end{vmatrix}\begin{vmatrix} 2 & 0 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} 5 & 0 \\ 1 & 6 \end{vmatrix}\begin{vmatrix} 2 & 0 \\ 1 & 5 \end{vmatrix} + \begin{vmatrix} 4 & 0 \\ 1 & 6 \end{vmatrix}\begin{vmatrix} 2 & 0 \\ -3 & 2 \end{vmatrix} = -15(-9) -$$

$$-(-12)-(-7) + (-15)(-17) + 0\cdot 2 - 30\cdot 10 + 24\cdot 4 = 102.$$

Приведенные соотношения редко используются непосредственно для вычисления определителей, но они чрезвычайно полезны при обосновании различных методов. Приведем также выражение, которое обобщает разложение по элементам строки или столбца:

$$\sum_{s=1}^{n} a_{is}\Delta_{js} = \sum_{s=1}^{n} a_{si}\Delta_{sj} = \begin{cases} \Delta & \text{при } i = j; \\ 0 & \text{при } i \neq j. \end{cases}$$

6. Вычисление определителей. Разложение определителя по элементам строки или столбца проще всего выполнять, когда в этой строке или столбце имеется единственный ненулевой элемент. Тогда определитель равен произведению этого элемента на его алгебраическое дополнение. К такому виду можно преобразовать определитель путем операций над его строками или столбцами, используя основные свойства (3).

Процесс вычисления иллюстрируется следующим примером:

$$\begin{vmatrix} -2 & 3 & 2 & 0 \\ 5 & 0 & -3 & 2 \\ 4 & 0 & 1 & 5 \\ 1 & 6 & 2 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 3 & 2 & 0 \\ 5 & 0 & -3 & 2 \\ 4 & 0 & 1 & 5 \\ 5 & 0 & -2 & 1 \end{vmatrix} = -3 \begin{vmatrix} 5 & -3 & 2 \\ 4 & 1 & 5 \\ 5 & -2 & 1 \end{vmatrix} =$$

$$= -3 \begin{vmatrix} 5 & -3 & 2 \\ 0 & \frac{17}{5} & \frac{17}{5} \\ 0 & 1 & -1 \end{vmatrix} = -3 \cdot 5 \begin{vmatrix} \frac{17}{5} & \frac{17}{5} \\ 1 & -1 \end{vmatrix} = -3 \cdot 5 \cdot \frac{17}{5} \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 102.$$

Здесь сначала первая строка, умноженная на (-2), прибавляется к последней строке, в результате чего второй столбец содержит только один ненулевой элемент. Разложение по этому столбцу приводит к определителю третьего порядка. Прибавляя ко второй и третьей строкам первую, умноженную соответственно на -4/5 и -1, получаем столбец с единственным ненулевым элементом. Теперь раскладываем определитель по первому столбцу и сводим его к определителю второго порядка. Так как элементы первой строки оказались равными, выносим за знак определителя множитель 17/5 и, раскрывая определитель второго порядка, получаем окончательный результат $\Delta = 102$.

Наиболее просто вычисляется определитель треугольной или диагональной матрицы: он равен произведению диагональных элементов. Это следует из разложения по элементам столбцов (строк) определителя верхней (нижней) треугольной матрицы (в случае диагональной матрицы разложение можно выполнять по элементам строк или столбцов). Значение определителей треугольных матриц не зависят от элементов, расположенных вне главной диагонали. Примеры:

$$\begin{vmatrix} 2 & 3 & -7 & 4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 2 \end{vmatrix} = 2 \cdot 1 \cdot 5 \cdot 2 = 20; \quad \begin{vmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 2 \end{vmatrix} = 2 \cdot 1 \cdot 5 \cdot 2 = 20.$$

С помощью разложения Лапласа можно также показать, что определитель квазидиагональной матрицы равен произведению определителей квадратных матриц, расположенных вдоль главной диагонали:

$$\begin{vmatrix} A_1 & & & \\ & A_2 & & \\ & & \ldots & \\ & & & A_m \end{vmatrix} = |A_1|\,|A_2| \ \ldots \ |A_m|.$$

Действительно, разлагая определитель по элементам строк матрицы A_1, получаем единственный ненулевой минор, который совпадает с $|A_1|$, и т.д. Пример:

$$\left|\begin{array}{ccc|cc} 4 & 0 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right| = \begin{vmatrix} 4 & 0 & 1 \\ 2 & 1 & -2 \\ 3 & 2 & 0 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ 0 & -2 \end{vmatrix} = 17(-2) = -34.$$

7. Схема единственного деления. При ручном счете можно использовать особенности конкретного определителя для упрощения процесса вычисления. Для вычислительных машин более подходящими являются стандартные процедуры. Одна из них *(схема единственного деления)* основана на последовательном преобразовании элементов определителя по *ведущим (опорным) элементам.* На первом шаге ведущий элемент a_{11} выносится в качестве общего множителя из первой строки (все элементы первой строки делятся на a_{11}). Потом из каждой строки вычитается первая строка, умноженная на первый элемент данной строки. В результате получаем

$$\Delta = a_{11} \begin{vmatrix} 1 & \frac{a_{12}}{a_{11}} & \ldots & \frac{a_{1n}}{a_{11}} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} 1 & \frac{a_{12}}{a_{11}} & \ldots & \frac{a_{1n}}{a_{11}} \\ 0 & a_{22}-a_{21}\frac{a_{12}}{a_{11}} & \ldots & a_{2n}-a_{21}\frac{a_{1n}}{a_{11}} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{n2}-a_{n1}\frac{a_{12}}{a_{11}} & \ldots & a_{nn}-a_{n1}\frac{a_{1n}}{a_{11}} \end{vmatrix}.$$

Разлагая по элементам первого столбца, получаем произведение a_{11} на определитель $(n — 1)$-го порядка, элементы которого выражаются соотношением:

$$a'_{ij} = a_{ij} - \frac{a_{i1}\,a_{1j}}{a_{11}} \ (i, j = 2,\ 3,\ \ldots,\ n).$$

Затем в полученном определителе выбираем в качестве ведущего элемент a'_{22} и поступаем аналогично. На k-м шаге образуется произведение $a_{11}a'_{22}\ldots a^{(k-1)}_{kk}$ и определитель $(n-k)$-го порядка с элементами

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)} a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} \quad (i,\ j = k+1,\ \ldots,\ n).$$

Процедура заканчивается за n шагов, причем искомое значение определителя равняется произведению ведущих элементов: $\Delta = a_{11}a'_{22}a''_{33}\ldots a^{(n-1)}_{nn}$. Рассмотренная схема иллюстрируется примером:

$$\Delta = \begin{vmatrix} -2 & 3 & 2 & 0 \\ 5 & 0 & -3 & 2 \\ 4 & 0 & 1 & 5 \\ 1 & 6 & 2 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & -\frac{3}{2} & -1 & 0 \\ 0 & \frac{15}{2} & 2 & 2 \\ 0 & 6 & 5 & 5 \\ 0 & \frac{15}{2} & 3 & 1 \end{vmatrix} = -2 \begin{vmatrix} \frac{15}{2} & 2 & 2 \\ 6 & 5 & 5 \\ \frac{15}{2} & 3 & 1 \end{vmatrix} =$$

$$= -2 \cdot \frac{15}{2} \begin{vmatrix} 1 & \frac{4}{15} & \frac{4}{15} \\ 0 & \frac{17}{5} & \frac{17}{5} \\ 0 & 1 & -1 \end{vmatrix} = -2 \cdot \frac{15}{2} \begin{vmatrix} \frac{17}{5} & \frac{17}{5} \\ 1 & -1 \end{vmatrix} = -2 \cdot \frac{15}{2} \cdot \frac{17}{5} \begin{vmatrix} 1 & 1 \\ 0 & -2 \end{vmatrix} =$$

$$= -2 \cdot \frac{15}{2} \cdot \frac{17}{5} \cdot (-2) = 102.$$

Схему единственного деления в равной мере можно реализовать путем соответствующих операций над столбцами, что равносильно работе с транспонированным определителем. Если на каком-либо шаге $a^{(k-1)}_{kk}=0$ (или близко к нулю), то перестановкой строк и столбцов можно избежать этой ситуации, которая препятствует осуществлению схемы (или снижающей точность вычислений). Вообще рекомендуется на каждом шаге переводить в левый верхний угол (перестановкой строк и столбцов) наибольший по абсолютной величине элемент, выбирая его в качестве ведущего.

8. Метод опорного элемента. Одной из разновидностей схемы единственного деления есть *метод опорного элемента,* который позволяет свести определитель n-го порядка к определителю (n — 1)-го порядка, элементы которого выражены как миноры второго порядка. Пусть в качестве опорного избран элемент a_{rs}. Делением элементов r-й строки на a_{rs} преобразуем (rs)-элемент к единице. Далее операциями над строками определителя, как в схеме единственного деления, образуем все элементы s-го столбца, кроме (rs)-элемента, в нули. При этом элементы определителя (за исключением элементов r-й строки и r-го столбца) выражаются следующим образом:

$$a'_{ij} = a_{ij} - \frac{a_{rj}\,a_{is}}{a_{rs}} = \frac{1}{a_{rs}}\begin{vmatrix} a_{rs} & a_{rj} \\ a_{is} & a_{ij} \end{vmatrix} = \frac{1}{a_{rs}} M\begin{pmatrix} r, & i \\ s, & j \end{pmatrix}.$$

Разлагая определитель по элементам s-го столбца, получаем $\Delta=(-1)^{r+s}M_{rs}$, где M_{rs} — дополнительный минор к (rs)- элементу, т.е. определитель (n — 1)-го порядка с элементами a'_{ij} ($i \neq r$, $j \neq s$). Каждый такой элемент является минором второго порядка, который образован из элементов определителя Δ, расположенных на пересечении строк r, i и столбцов s, j. При $i>r$ и $j > s$ сохраняется естественный порядок элементов минора. Для сохранения такого же порядка при $r>i$ и $s>j$ необходимо поменять знаки всех миноров, которые лежат выше r-й строки и левее s-го столбца (рис. 2.4, *а).*

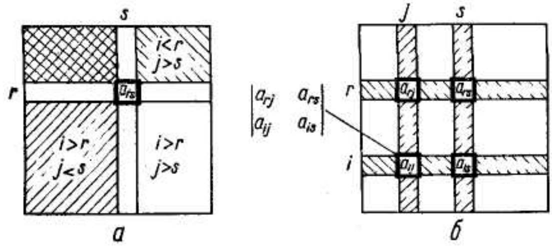


Рис. 2.4. Метод опорного элемента:
а — опорный элемент; *б* — замещение элемента определителя минором второго порядка.

Это значит, что определитель меняет знак (r - 1)+(s - 1) раз, что должно быть скомпенсировано множителем

$$(-1)^{r+s-2}=(-1)^{r+s}.$$

Таким образом, при сохранении естественного расположения элементов определителя Δ в минорах второго порядка $\Delta = a_{rs}M_{rs}$. Вынесем из каждой строки M_{rs} общий элемент $1/a_{rs}$. Тогда за счет всех (n — 1) строк перед определителем (n — 1)-го порядка появится множитель $\frac{1}{(a_{rs})^{n-1}}$, и в результате получаем разложение определителя Δ относительно опорного элемента a_{rs}:

$$\Delta = \frac{1}{(a_{rs})^{n-2}}\,\Delta'.$$

Здесь Δ'— определитель (n — 1)-го порядка, элементами которого являются миноры второго порядка. В соответствии с изложенным выше Δ' образуется из Δ замещением в последнем каждого элемента a_{ij} ($i \neq r$, $j \neq s$) минором, образованным из элементов пересечения строк r, i

и столбцов *s*, *j*, сохраняя порядок их следования в исходном определителе (рис. 2.4, *б).* Так, для нашего примера, выбирая $a_{41} = 1$ в качестве опорного элемента, имеем:

$$\Delta = \begin{vmatrix} -2 & 3 & 2 & 0 \\ 5 & 0 & -3 & 2 \\ 4 & 0 & 1 & 5 \\ 1 & 6 & 2 & 1 \end{vmatrix} = \frac{1}{1^2} \begin{Vmatrix} \begin{vmatrix} -2 & 3 \\ 1 & 6 \end{vmatrix} & \begin{vmatrix} -2 & 2 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} -2 & 0 \\ 1 & 1 \end{vmatrix} \\ \begin{vmatrix} 5 & 0 \\ 1 & 6 \end{vmatrix} & \begin{vmatrix} 5 & -3 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 5 & 2 \\ 1 & 1 \end{vmatrix} \\ \begin{vmatrix} 4 & 0 \\ 1 & 6 \end{vmatrix} & \begin{vmatrix} 4 & 1 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 4 & 5 \\ 1 & 1 \end{vmatrix} \end{Vmatrix} =$$

$$= \begin{vmatrix} -15 & -6 & -2 \\ 30 & 13 & 3 \\ 24 & 7 & -1 \end{vmatrix}.$$

Приняв за опорный элемент в полученном определителе $a_{33} = -1$, найдем:

$$\Delta = \begin{vmatrix} -15 & -6 & -2 \\ 30 & 13 & 3 \\ 24 & 7 & -1 \end{vmatrix} = \frac{1}{-1} \begin{Vmatrix} \begin{vmatrix} -15 & -2 \\ 24 & -1 \end{vmatrix} & \begin{vmatrix} -6 & -2 \\ 7 & -1 \end{vmatrix} \\ \begin{vmatrix} 30 & 3 \\ 24 & -1 \end{vmatrix} & \begin{vmatrix} 13 & 3 \\ 7 & -1 \end{vmatrix} \end{Vmatrix} = -\begin{vmatrix} 63 & 20 \\ -102 & -34 \end{vmatrix} =$$

$$= 102.$$

9. Определитель суммы матриц. Пусть требуется найти определитель суммы $C=A+B$ двух квадратных матриц *n*-го порядка Представим определитель этой суммы через столбцы слагаемых матриц

$$\Delta = | c^{(1)}, c^{(2)}, \ldots, c^{(n)} | = | a^{(1)} + b^{(1)}, a^{(2)} + b^{(2)}, \ldots, a^{(n)} + b^{(n)} |.$$

В соответствии со свойством линейности определителя относительно столбцов (3) запишем

$$\Delta = | a^{(1)} + b^{(1)}, c^{(2)}, \ldots, c^{(n)} | = | a^{(1)}, c^{(2)}, \ldots, c^{(n)} | + \\ + | b^{(1)}, c^{(2)}, \ldots, c^{(n)} |.$$

Применяя это свойство относительно вторых столбцов полученных определителей, имеем

$$\Delta = |a^{(1)}, a^{(2)}, c^{(3)}, \ldots, c^{(n)} | + | a^{(1)}, b^{(2)}, c^{(3)}, \ldots, c^{(n)} | + \\ + | b^{(1)}, a^{(2)}, c^{(3)}, \ldots, c^{(n)} | + | b^{(1)}, b^{(2)}, c^{(3)}, \ldots, c^{(n)} |.$$

Продолжая этот процесс до последних столбцов включительно, получаем разложение в сумму, которая содержит определители слагаемых матриц det *A* и det *B*, а также определители, образованные из столбцов матриц *A* и *B* всеми возможными сочетаниями, причем столбцы в таких определителях занимают те же места, которые они занимали в матрицах A и B. Это можно выразить соотношением

$$\det (A + B) = \det A + \sum\Delta (1) + \sum\Delta (2) + \ldots + \sum\Delta (n - 1) + \det B,$$

где $\Delta(s)$ — определитель, полученный замещением s столбцов определителя первой матрицы соответствующими столбцами второй матрицы. Знаки сумм означают, что суммируются определители для свевозможных сочетаний s столбцов, которые замещаются. Так как $\det A=\Delta(0)$ и $\det B=\Delta(n)$, то можно предложить более краткую запись

$$\det (A + B) = \det A + \sum_{s=1}^{n-1} \sum \Delta (s) + \det B = \sum_{s=0}^{n} \sum\Delta (s).$$

Воспользовавшись разложением Лапласа (5) для определителей $\Delta(s)$ по s замещенным столбцам, получим другое выражение для определителя суммы двух матриц

$$\det (A + B) = \sum_{s=0}^{n}\sum (-1)^{\sigma} M_s^B \overline{M}_s^A.$$

В силу коммутативности сложения матриц безразлично, какую из матриц A и B считать первой и какую — второй. Полученные разложения из-за своей сложности непригодны для практических вычислений определителей без применения компьютеров. Они могут быть полезны при доказательствах различных соотношений. В частности, они позволяют выразить вещественную и мнимую части определителя комплексной матрицы

$$\det (A) = \det (A' + iA'') = \Delta' + i\Delta'';$$

$$\Delta' = \sum_{k=0}^{m} (-1)^k \sum\Delta (2k); \quad \Delta'' = \sum_{k=0}^{m} (-1)^k \sum\Delta (2k + 1),$$

где $m = \frac{1}{2}\,n$ — для четных *n;* $m = \frac{1}{2}\,(n - 1)$ - для не четных *n.*

Применим эти формулы для вычисления определителя комплексной матрицы

$$A = \begin{bmatrix} 2i & 5 & 3-4i \\ 4+2i & 1+i & i \\ 3 & -1+2i & 2-i \end{bmatrix} = \begin{bmatrix} 0 & 5 & 3 \\ 4 & 1 & 0 \\ 3 & -1 & 2 \end{bmatrix} + i \begin{bmatrix} 2 & 0 & -4 \\ 2 & 1 & 1 \\ 0 & 2 & -1 \end{bmatrix}.$$

Для вещественной и мнимой частей определителя $\det A$ имеем:

$$\Delta' = \begin{bmatrix} 0 & 5 & 3 \\ 4 & 1 & 0 \\ 3 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 3 \\ 2 & 1 & 0 \\ 0 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 5 & -4 \\ 2 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & -4 \\ 4 & 1 & 1 \\ 3 & 2 & -1 \end{bmatrix} =$$
$$= -61 - 16 - 18 + 20 = -75;$$
$$\Delta'' = \begin{bmatrix} 2 & 5 & 3 \\ 2 & 1 & 0 \\ 0 & -1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 3 \\ 4 & 1 & 0 \\ 3 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 5 & -4 \\ 4 & 1 & 1 \\ 3 & -1 & -1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & -4 \\ 2 & 1 & 1 \\ 0 & 2 & -1 \end{bmatrix} =$$
$$= -22 + 15 + 63 + 22 = 78.$$

Таким образом, $\det A = -75 + 78i$, что можно проверить непосредственным вычислением определителя.

Разложение определителя суммы двух матриц можно обобщить для любого количества квадратных матриц одного и того же порядка. Так, для трех матриц имеем:

$$\det(A + B + C) = \sum_{s=0}^{n} \sum_{k=0}^{n-s} \sum \Delta(s, k),$$

где через $\Delta(s, k)$ обозначені определители, образованные всеми возможными замещениями столбцов определителя первой матрицы s столбцами второй матрицы и k столбцами третьей матрицы.

10. Определитель произведения матриц. Можно показать, что определитель произведения двух квадратных матриц A и B одинаковых порядков равен n произведению их определителей, т.е.
$\det(AB) = \det A \det B$. Для этого рассмотрим матрицу порядка $2n$

$$D = \begin{bmatrix} A & 0 \\ -E_n & B \end{bmatrix},$$

где E_n - единичная матрица.

Применяя разложение Лапласа по первым n строкам определителя этой матрицы, имеем $|D|=|A||B|$. Представим определитель $|D|$ в виде:

$$|D| = \left|\begin{array}{cccc|c} a^{(1)}, & a^{(2)}, & \cdots, & a^{(n)} & 0 \\ \hline -1 & & \cdots & & b_{(1)} \\ & -1 & \cdots & & b_{(2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & \cdots & -1 & b_{(n)} \end{array}\right|,$$

где $a^{(1)}, a^{(2)}, \ldots, a^{(n)}$ — столбцы матрицы A; $b_{(1)}, b_{(2)}, \ldots, b_{(n)}$ — строки матрицы B; 0 — нулевая матрица n-го порядка.

Преобразуем определитель $|D|$ к такому виду, чтобы на месте элементов b_{ij} ($i, j=1,2,\ldots, n$) были нули. Для этого первый столбец умножим на элементы строки $b_{(1)}$ и прибавим его к соответствующим *(n*+1, *n*+2, ..., 2*n)* столбцов. Аналогично поступим со вторым, третьим и т.д. до *n*-го включительно столбцами. В результате получим:

$$|D| = \left|\begin{array}{cccc|c} a^{(1)}, & a^{(2)}, & \cdots, & a^{(n)} & 0 \\ \hline -1 & & \cdots & & b_{(1)} \\ & -1 & \cdots & & b_{(2)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & \cdots & -1 & b_{(n)} \end{array}\right|,$$

где сумма в правом верхнем углу заменена произведением *AB* в соответствии с (п. 1.4).

На основе разложения Лапласа по первым *n* строкам находим, что $|D|=|AB|$. Таким образом, $|AB| =|A||B|$ или $\det(AB) = \det A \det B$, что и требовалось доказать.

Естественным обобщением этого результата является *теорема Бине-Коши* об определителе произведения *AB* двух прямоугольных матриц размера $(m\times n)$ и $(n \times m)$:

$$\det(AB) = \sum A\begin{pmatrix} 1, & 2, & \ldots, & m \\ \alpha_1, & \alpha_2, & \ldots, & \alpha_m \end{pmatrix} B\begin{pmatrix} \alpha_1, & \alpha_2, & \ldots, & \alpha_m \\ 1, & 2, & \ldots, & m \end{pmatrix},$$

где сумма означает, что суммируются произведения всевозможных миноров *m*-го порядка матрицы *A,* которые образованы *m* ее столбцами с номерами $\alpha_1, \alpha_2, \ldots, \alpha_m$, на миноры матрицы *B,* которые образованы ее строками с теми же номерами. В других обозначениях эту теорему можно записать следующим образом:

$$|AB| = \sum |a^{(\alpha_1)}, \ a^{(\alpha_2)}, \ \ldots, \ a^{(\alpha_m)}| \begin{vmatrix} b_{(\alpha_1)} \\ b_{(\alpha_2)} \\ \cdots \\ b_{(\alpha_m)} \end{vmatrix},$$

где $\alpha_1, \alpha_2, \ldots, \alpha_m$ — всевозможные сочетания из *n* номеров, которые расположены в порядке их следования.

При $m>n$ полагают $|AB|=0$, а при $m=n$ имеем рассмотренный выше частный случай произведения квадратных матриц.

Из соотношения $\det(AB)= \det A \det B$ следует, что определители можно умножатб по правилам умножения матриц.

Пример:

$$A=\begin{vmatrix}2&0&1\\1&3&2\\0&1&4\end{vmatrix};\qquad B=\begin{vmatrix}1&2&4\\3&1&0\\0&5&1\end{vmatrix};$$

$$AB=\begin{vmatrix}2\cdot1+0\cdot3+1\cdot0 & 2\cdot2+0\cdot1+1\cdot5 & 2\cdot4+0\cdot0+1\cdot1\\ 1\cdot1+3\cdot3+2\cdot0 & 1\cdot2+3\cdot1+2\cdot5 & 1\cdot4+3\cdot0+2\cdot1\\ 0\cdot1+1\cdot3+4\cdot0 & 0\cdot2+1\cdot1+4\cdot5 & 0\cdot4+1\cdot0+4\cdot1\end{vmatrix}=$$

$$=\begin{vmatrix}2&9&9\\10&15&6\\3&21&4\end{vmatrix}=1155;$$

$$|A|=21;\quad |B|=55;\quad |AB|=|A|\,|B|=1155.$$

На заключение отметим, что $|A\oplus B| = |A||B|$ и $|A\otimes B| = — |A|^q|B|^p$, где q — порядок матрицы B и p — порядок матрицы A.

11. Дифференцирование определителей. Если элементы определителя представляют собой некоторые функции переменной x, то такой определитель можно дифференцировать по этой же переменной. При этом каждый член в разложении (1) запишется как сумма n слагаемых, которые получаются заменой в данном члене одного из элементов его производной. Сгруппировав члены, которые содержат производные первых, вторых и, вообще, j-х элементов, получим n групп слагаемых.

Каждая из этих групп соответствует определителю, который получается из исходного заменой в нем элементов j-го столбца производными этих элементов. Сумма n таких определителей при j = 1, 2, ..., n и будет равна производной данного определителя. Ясно, что тот же результат можно сформулировать для строк:

$$\Delta'(x)=\begin{vmatrix}a'_{11}(x)\,a'_{12}(x) & \cdots & a'_{1n}(x)\\ a_{21}(x)\,a_{22}(x) & \cdots & a_{2n}(x)\\ \cdots\ \ \cdots & \cdots & \cdots\\ a_{n1}(x)\,a_{n2}(x) & \cdots & a_{nn}(x)\end{vmatrix}+\begin{vmatrix}a_{11}(x)\,a_{12}(x) & \cdots & a_{1n}(x)\\ a'_{21}(x)\,a'_{22}(x) & \cdots & a'_{2n}(x)\\ \cdots\ \ \cdots & \cdots & \cdots\\ a_{n1}(x)\,a_{n2}(x) & \cdots & a_{nn}(x)\end{vmatrix}+\ldots$$

$$\ldots+\begin{vmatrix}a_{11}(x)\,a_{12}(x) & \cdots & a_{1n}(x)\\ a_{21}(x)\,a_{22}(x) & \cdots & a_{2n}(x)\\ \cdots\ \ \cdots & \cdots & \cdots\\ a'_{n1}(x)\,a'_{n2}(x) & \cdots & a'_{nn}(x)\end{vmatrix}.$$

Продифференцируем, например, определитель

$$\Delta(x)=\begin{vmatrix}x^2 & \sin x & 2\\ x+2 & \cos x & 3x^2\\ 1 & e^{2x} & 1+e^x\end{vmatrix}.$$

В соответствии с изложеным правилом получаем

$$\Delta'(x)=\begin{vmatrix}2x & \sin x & 2\\ 1 & \cos x & 3x\\ 0 & e^{2x} & 1+e^{x}\end{vmatrix}+\begin{vmatrix}x^2 & \cos x & 2\\ x+2 & -\sin x & 3x^2\\ 1 & 2e^{2x} & 1+e^{x}\end{vmatrix}+\begin{vmatrix}x^2 & \sin x & 0\\ x+2 & \cos x & 6x\\ 1 & e^{2x} & e^{x}\end{vmatrix}.$$

<u>Микромодуль 4</u>
Примеры решения типовых задач

Вычислите определители:

1. $a)\ \begin{vmatrix}0 & -2\\ 8 & 7\end{vmatrix};\quad б)\ \begin{vmatrix}\sin\alpha & \cos\alpha\\ \sin\beta & \cos\beta\end{vmatrix}$

Решение

$$a)\ \begin{vmatrix}0 & -2\\ 8 & 7\end{vmatrix}=0\cdot 7-8\cdot(-2)=16$$

$$б)\ \begin{vmatrix}\sin\alpha & \cos\alpha\\ \sin\beta & \cos\beta\end{vmatrix}=\sin\alpha\cos\beta-\sin\beta\cos\alpha=\sin(\alpha-\beta)$$

2. $\begin{vmatrix}3 & 2 & -4\\ 3 & -1 & 1\\ 1 & 1 & 3\end{vmatrix}$

Решение. Первый способ. По правилу треугольников имеем

$$\Delta=\begin{vmatrix}3 & 2 & -4\\ 3 & -1 & 1\\ 1 & 1 & 3\end{vmatrix}=3\cdot(-1)\cdot 3+2\cdot 1\cdot 1+(-4)\cdot 3\cdot 1-(1\cdot(-1)\cdot(-4)+2\cdot 3\cdot 3+3\cdot 1\cdot 1)=$$

$$=-9+2-12-(4+18+3)=-19-25=-44$$

Второй способ. Используя свойства определителя, получаем
$\Delta=$

$$=\begin{vmatrix}3 & 2 & -4\\ 3 & -1 & 1\\ 1 & 1 & 3\end{vmatrix}=\begin{vmatrix}0 & -1 & -13\\ 0 & -4 & -8\\ 1 & 1 & 3\end{vmatrix}=a_{31}A_{31}=1\cdot(-1)^{3+1}M_{31}=\begin{vmatrix}-1 & -13\\ -4 & -8\end{vmatrix}=8-52=-44$$

Третий способ. Разложим определитель по первой строке: Δ=

$$=\begin{vmatrix}3 & 2 & -4\\ 3 & -1 & 1\\ 1 & 1 & 3\end{vmatrix}=(-1)^{1+1}\cdot 3\cdot\begin{vmatrix}-1 & 1\\ 1 & 3\end{vmatrix}+(-1)^{1+2}\cdot 2\cdot\begin{vmatrix}3 & 1\\ 1 & 3\end{vmatrix}+(-1)^{1+3}\cdot(-4)\cdot\begin{vmatrix}3 & -1\\ 1 & 1\end{vmatrix}=3\cdot(-4)-2\cdot 8-4\cdot 4=-44$$

3. $\begin{vmatrix}1 & 0 & -1 & 2\\ 1 & 2 & -2 & 0\\ -1 & 3 & 0 & 2\\ 2 & 1 & -1 & 3\end{vmatrix}$

Решение. В определителе есть несколько нулевых элементов, тем не менее удобно, когда нулевые элементы содержатся в одной строке или столбце. Сделаем, например, нулевыми все элементы первой строки, кроме первого элемента. Для этого прибавим к третьему столбику первый, после чего умножим элементы первого столбика на -2 и прибавим их к соответствующим элементам четвертого столбика. Получим

$$\Delta=\begin{vmatrix}1 & 0 & -1 & 2\\ 1 & 2 & -2 & 0\\ -1 & 3 & 0 & 2\\ 2 & 1 & -1 & 3\end{vmatrix}=\begin{vmatrix}1 & 0 & -1+1 & 2-2\\ 1 & 2 & -2+1 & 0-2\\ -1 & 3 & 0-1 & 2+2\\ 2 & 1 & -1+2 & 3-4\end{vmatrix}=\begin{vmatrix}1 & 0 & 0 & 0\\ 1 & 2 & -1 & -2\\ -1 & 3 & -1 & 4\\ 2 & 1 & 1 & -1\end{vmatrix}$$

Теперь разложим определитель по элементам первой строки:

$$\Delta=a_{11}A_{11}=1\cdot(-1)^{1+1}M_{11}=\begin{vmatrix}2 & -1 & -2\\ 3 & -1 & 4\\ 1 & 1 & -1\end{vmatrix}$$

Создадим два нуля во втором столбике. Для этого к первому и второму строкам по-очереди прибавляем третью строку. Тогда

$$\Delta=\begin{vmatrix}5 & 0 & -3\\ 4 & 0 & 3\\ 1 & 1 & -1\end{vmatrix}$$

Разложим определитель по элементам второго столбца:

$$\Delta=a_{32}A_{32}=1\cdot(-1)^{3+21}M_{32}=-\begin{vmatrix}5 & -3\\ 4 & 3\end{vmatrix}=-(15+12)=-27$$

4. Решите уравнение $\begin{vmatrix} -2-k & 2 & 0 \\ 2 & 4-k & 6 \\ 1 & 2 & 3-k \end{vmatrix} = 0$

Решение. Разложим определитель по элементам первой строки:

$$\begin{vmatrix} -2-k & 2 & 0 \\ 2 & 4-k & 6 \\ 1 & 2 & 3-k \end{vmatrix} = (-2-k)\begin{vmatrix} 4-k & 6 \\ 2 & 3-k \end{vmatrix} - 2\begin{vmatrix} 2 & 6 \\ 1 & 3-k \end{vmatrix} + 0 =$$

$$= -(2+k)((4-k)(3-k)-12) - 2(6-2k-6) = -(2+k)(k^2-7k)+4k =$$

$$= -(k^3+2k^2-14k)+4k = -k^3+5k^2+18k$$

Следовательно, данное уравнение равносильно уравнению

$$-k^3+5k^2+18k$$

Далее имеем

$$-k(k^2-5k-18=0, \quad k=0 \text{ либо } k^2-5k-18=0 \text{ откуда } k=\frac{5\pm\sqrt{97}}{2}$$

Микромодуль 4.
Индивидуальные тестовые задачи.

1. Вычислите по схемам (см. рис. 2.1) следующие определители:

а) $\begin{vmatrix} 2 & -1 \\ 3 & 5 \end{vmatrix}$; б) $\begin{vmatrix} \sin\alpha & -\cos\alpha \\ \cos\alpha & \sin\alpha \end{vmatrix}$; в) $\begin{vmatrix} x-1 & x^3 \\ x & x^2+2x-1 \end{vmatrix}$;

г) $\begin{vmatrix} 0 & -3 & 4 \\ 2 & 1 & -2 \\ 1 & 5 & 1 \end{vmatrix}$; д) $\begin{vmatrix} a+x & a & a \\ a & a+x & a \\ a & a & a+x \end{vmatrix}$.

2. Вычислите определитель матрицы *А,* пользуясь только его определением, приведенным в (п.1):

$$A = \begin{bmatrix} -2 & 4 & 0 & 1 \\ 3 & 1 & 1 & -2 \\ 6 & -3 & 7 & 3 \\ 0 & 5 & 4 & -1 \end{bmatrix}.$$

3. Постройте граф матрицы *А* из задачи 2 и все его факторы. Установите соответствие между факторами и членами определителя, полученными в предыдущей задаче.

4. Вычислите определитель матрицы *А* из задачи 2, пользуясь общими свойствами, приведенными в (п. 3).

5. Найдите алгебраические дополнения всех элементов матрицы A из задачи 2 и вычислите ее определитель разложением по элементам какой-либо строки и какого-либо столбца. Покажите, что разложение по другим строкам или столбцам приводят к тому же результату.

6. Вычислите определитель матрицы A из задачи 2 с помощью разложения Лапласа по первой и четвертой строкам.

7. Вычислите определитель матрицы A из задачи 2 следующими способами:

а) по схеме единственного деления;

б) методом опорного элемента.

8. Найдите наиболее простой способ вычисления следующего определителя:

$$\begin{vmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & -4 & 0 & 0 & 0 & 0 & 0 \\ 7 & 1 & 5 & 0 & 0 & 0 & 0 \\ 5 & 2 & 3 & 1 & 2 & 0 & 0 \\ 4 & 3 & 1 & -3 & 5 & 0 & 0 \\ 2 & -3 & 4 & 6 & 5 & 3 & -1 \\ 1 & 4 & 3 & 1 & 6 & 2 & 1 \end{vmatrix}.$$

9. Как изменится определитель n-го порядка, если

а) из его первой строки вычесть вторую, из второй — третью и из третьей строки — первоначальную первую строку $(n \geq 3)$;

б) первые его k строк расположить в обратном порядке, а также в обратном порядке записать следующие $(n — k)$ строк;

в) каждый его элемент a_{ij} умножить на число ij.

10. С помощью формулы для определителя суммы двух матриц вычислите действительную и мнимую части определителя

$$\begin{vmatrix} 4+5i & 0 & 1 & 3-2i \\ 3 & 1-i & 6 & 7+5i \\ 0 & 7 & 4 & 0 \\ 3+2i & 2 & 2+i & 1 \end{vmatrix}$$

и проверьте результат непосредственным вычислением.

11. Представьте в виде многочлена от p с помощью формулы для определителя суммы двух матриц следующий определитель:

$$\begin{vmatrix} p+5 & 0 & -p & 0 \\ 10 & 2p+4 & -3p & -1 \\ -p & -3p & 5p+7 & -2 \\ 0 & -1 & -2 & 4 \end{vmatrix}.$$

12. Покажите, что det (AB) = det A det B = det (BA) на примере матриц:

$$A = \begin{vmatrix} 1 & 0 & 2 \\ 2 & -1 & -1 \\ -1 & 3 & 5 \end{vmatrix}; \quad B = \begin{vmatrix} 1 & -3 & 4 \\ 2 & 1 & 0 \\ 4 & 0 & 2 \end{vmatrix}.$$

13. Найдите определитель произведения матриц det (AB), где

$$A = \begin{bmatrix} 1 & 2 & 4 & 0 \\ 3 & -1 & 2 & 5 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ -1 & 5 \\ 4 & 0 \end{bmatrix};$$

а) с помощью теоремы Бине-Коши;

б) умножением матриц и последующим вычислением определителя результирующей матрицы.

14. Найдите производную по x определителя

$$\begin{vmatrix} x^2+3 & x+2 & 4 \\ x^3 & x^4 & 3x-2 \\ x-1 & x^3 & x \end{vmatrix}.$$

15. Покажите справедливость приведенного ниже выражения для определителя n-го порядка

$$\begin{vmatrix} n & -1 & \dots & -1 \\ -1 & n & \dots & -1 \\ \dots & \dots & \dots & \dots \\ -1 & -1 & \dots & n \end{vmatrix} = (n+1)^{n-1}.$$

16. Покажите, что определитель k-го порядка вида

$$D_k = \begin{vmatrix} 3 & 2 & 0 & 0 & \dots & 0 & 0 \\ 1 & 3 & 2 & 0 & \dots & 0 & 0 \\ 0 & 1 & 3 & 2 & \dots & 0 & 0 \\ 0 & 0 & 1 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & 3 \end{vmatrix}$$

можно представить рекуррентным соотношением $D_k = 3D_{k-1} - 2D_{k-2}$.

Приняв $D_1 = 3$, вычислите D_2, D_3 и D_4.

Методом индукции докажите, что $D_k = 2^{k+1} - 1$.

Микромодуль 5. Алгебраические уравнения

2.2. Линейные уравнения

1. Общая характеристика. Решение систем линейных уравнений - это одна из центральных задач вычислительной математики, наиболее часто встречающаяся в различных областях практической деятельности. К этой задаче сводятся или ею сопровождаются процедуры анализа и синтеза многих систем различной природы: электрических, механических, гидравлических и т.п. Она играет важную роль в прикладных методах математической статистики и экономики, в теории оптимального кодирования при передаче информации и во многих других разделах современной науки и техники. Даже если исследуемая система нелинейна, то типичный путь ее числового анализа лежит через линеаризацию и сводится к решению систем линеаризованных уравнений.

В общем случае число уравнений m может отличаться от числа неизвестных. Такая система

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\}$$

в матричной форме $Ax=b$ характеризуется прямоугольной матрицей коэффициентов A размера $m \times n$ и столбцом свободных членов b. Если все свободные члены $b_1, b_2, \ldots, b_m$ равны нулю, то система называется *однородной,* а если среди них имеется хотя бы один ненулевой член, то система называется *неоднородной.*

Среди уравнений системы могут быть линейно-зависимые, т.е. такие, которые можно представить как результат сложения других уравнений, умноженных на какие-либо числа (при умножении уравнения на число его левая и правая части умножаются на это число). Ясно, что зависимые уравнения не содержат никакой дополнительной информации, влияющей на значения искомых величин. Исключение зависимых уравнений приводит к эквивалентной системе, решение которой совпадает с решением исходной системы.

2. Ранг системы. Число независимых уравнений определяет *ранг системы.* Соответственно число независимых строк (либо столбцов) матрицы называют *рангом матрицы.* Если ранг матрицы равен r, то хотя бы один ее минор r-го порядка отличен от нуля, а все остальные миноры, выше r-го порядка, равны нулю.

Система n уравнений из n неизвестными $Ax=b$, матрица которой A неособенная, т.е. det $A\neq 0$, имеет единственное решение $x=A^{-1}b$ и поэтому называется *определенной системой n-го порядка.* Подстановка решения в уравнение превращает его в тождество $A(A^{-1}b)=b$ или $b=b$. Ранг неособенной матрицы A равен ее порядку $(r — n)$. Если среди n уравнений имеются зависимые, то A — особенная матрица и ее ранг меньше порядка $(r<n)$. Разность $d = n — r$ называют *дефектом матрицы,* а самую матрицу — *d-кратно вырожденной* (при $d=1$ матрица называется *просто вырожденной).* После исключения d зависимых уравнений получим эквивалентную систему r уравнений с n неизвестными.

Система m уравнений с n неизвестными при $m\neq n$ может иметь решение, а может и не иметь их вовсе. Если система имеет хотя бы одно решение, ее называют *совместной,* а систему, для которой решение не существует, называют *несовместной.* Совместная система при $m<n$ всегда имеет бесконечное множество решений и называется *неопределенной.* При $m > n$ система является или несовместной, либо сводится к эквивалентной ей совместной системе, которая может быть определенной $(r=n)$ или не определенной $(r<n)$.

Применение компьютеров позволяет решать системы уравнений с большим числом неизвестных. Для этой цели разработаны высокоэффективные алгоритмы и программы, а также целые программные системы. Многие из них основаны на идее исключения, которая уже использовались для вычисления определителей и обращения матриц.

3. Алгоритм Гаусса. Для решения неоднородных систем линейных уравнений n-го порядка идея исключения нашла одно из своих первых воплощений в *алгоритме Гаусса.*

Пусть дана система n линейных уравнений из n неизвестными $x_1, x_2, ..., x_n$ и правыми частями $y_1, y_2, ..., y_n$:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = y_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = y_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = y_n. \end{array}\right\} \quad (2.1)$$

В матричной форме эта система, как мы знаем, может быть записана так:

$$Ax = y. \tag{2.2}$$

Здесь $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ - столбцы и $A = \| a_{ik} \|^n_1$ - квадратная матрица коэффициентов.

Если A — неособенная матрица, то можно написать:

$$x = A^{-1}y \tag{2.3}$$

или в развернутом виде:

$$x_i = \sum_{k=1}^{n} a_{ik}^{(-1)} y_k \quad (i = 1, 2, \ldots, n). \tag{2.4}$$

Таким образом, задача вычисления элементов обратной матрицы $A^{-1} = \| a^{(-1)}_{ik} \|^n_1$ эквивалентна задаче решения системы уравнений (2.1) при любых правых частях $y_1, y_2, \ldots, y_n$. Элементы обратной матрицы определяются формулами (1.26) модуля 1. Однако фактическое вычисление элементов матрицы A^{-1} по этим формулам при большом n весьма затруднительно. Поэтому большое практическое значение имеют эффективные методы вычисления элементов обратной матрицы и, следовательно, решения системы линейных уравнений.

Ниже будут изложены теоретические основы некоторых из этих методов, которые представляют собой разновидности метода исключения Гаусса, знакомство с которым у читателя началось еще в курсе алгебры средней школы.

Главнейший сводится к последовательному исключению неизвестных, в результате чего данная система уравнений преобразуется к эквивалентной системе с верхней треугольной матрицей, решение которой не составляет труда.

Пусть в системе уравнений (2.1) $a_{11} \neq 0$. Мы исключим x_1 из всех уравнений, начиная со 2-го, для чего ко второму уравнению почленно прибавим первое, умноженное на $-\frac{a_{21}}{a_{11}}$, к третий почленно прибавим первое, умноженное на $-\frac{a_{31}}{a_{11}}$, и т.д. После этого система уравнений (2.1) заменится эквивалентной системой

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n2}^{(1)}x_2 + \ldots + a_{nn}^{(1)}x_n &= y_n^{(1)}. \end{aligned}\right\} \tag{2.5}$$

Коэффициенты при неизвестных и свободные члены в последних (n - 1) уравнениях определяются формулами

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}, \quad y_i^{(1)} = y_i - \frac{a_{i1}}{a_{11}} y_1 \quad (i, j = 2, \ldots, n).$$

(2.6)

Пусть $a^{(1)}_{22} \neq 0$. Тогда таким же образом мы исключим x_2 из последних n — 2 уравнений системы (2.5) и получим систему уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= y_3^{(2)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{n3}^{(1)}x_3 + \ldots + a_{nn}^{(2)}x_n &= y_n^{(2)}. \end{aligned}\right\}$$

(2.7)

При этом новые коэффициенты и правые части связаны с предыдущими формулами:

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} a_{2j}^{(1)}, \quad y_i^{(2)} = y_i^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} y_2^{(1)} \quad (i, j = 3, \ldots, n).$$

(2.8)

Продолжая этот алгоритм дальше, мы на (n - 1)-в этапе приведем исходную систему (2.1) к треугольной рекуррентной системе

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= y_1, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= y_2^{(1)}, \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= y_3^{(2)}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{nn}^{(n-1)}x_n &= y_n^{(n-1)}. \end{aligned}\right\}$$

(2.9)

Это приведение осуществимо в том и только в том случае, если в процессе приведения все числа $a_{11}, a^{(1)}_{22}, a^{(2)}_{33},\ldots, a^{(n-2)}_{(n-1)(n-1)}$ оказываются отличными от нуля.

Изложенный алгоритм Гаусса состоит из однотипных операций, которые легко выполняются на компьютерах.

Выразим коэффициенты и правые части приведенной системы через коэффициенты и правые части исходной системы (2.1). При этом мы не будем предполагать, что в процессе приведения все числа $a_{11}, a^{(1)}_{22}, a^{(2)}_{33},\ldots, a^{(n-2)}_{(n-1)(n-1)}$ оказываются отличными от нуля, а рассмотрим общий случай, когда первые p из этих чисел отличны от нуля:

$$a_{11}\neq 0,\ a^{(1)}_{22}\neq 0, \dots, a^{(p-1)}_{nn}\neq 0 \quad (p\leq n-1), \tag{2.10}$$

что дает возможность (на p-м этапе приведения) привести исходную систему уравнений к виду

$$\left.\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \dots\dots\dots\dots + a_{1n}x_n = y_1, \\
\quad a^{(1)}_{22}x_2 + \dots\dots\dots\dots + a^{(1)}_{2n}x_n = y^{(1)}_2, \\
\quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\quad a^{(p-1)}_{pp}x_p + \dots\dots + a^{(p-1)}_{pn}x_n = y^{(p-1)}_p, \\
\quad a^{(p)}_{p+1,\,p+1}x_{p+1} + \dots + a^{(p)}_{p+1,\,n}x_n = y^{(p)}_{p+1}, \\
\quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\quad a^{(p)}_{n,\,p+1}x_{p+1} + \dots + a^{(p)}_{nn}x_n = y^{(p)}_n.
\end{array}\right\} \tag{2.11}$$

Матрицу коэффициентов этой системы уравнений обозначим через G_p:

$$G_p = \left\|\begin{array}{llllll}
a_{11} & a_{12} & \dots\ a_{1p} & a_{1,\,p+1} & \dots\ a_{1n} \\
0 & a^{(1)}_{22} & \dots\ a^{(1)}_{2p} & a^{(1)}_{2,\,p+1} & \dots\ a^{(1)}_{2n} \\
\dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots\ a^{(p-1)}_{pp} & a^{(p-1)}_{p,\,p+1} & \dots\ a^{(p-1)}_{pn} \\
0 & 0 & \dots\ 0 & a^{(p)}_{p+1,\,p+1} & \dots\ a^{(p)}_{p+1,\,n} \\
\dots & \dots & \dots & \dots & \dots \\
0 & 0 & \dots\ 0 & a^{(p)}_{n,\,p+1} & \dots\ a^{(p)}_{nn}
\end{array}\right\|. \tag{2.12}$$

Переход от матрицы A к матрице G_p осуществляется следующим образом: к каждой строке матрицы A, начиная со 2-й и кончая n-ой, последовательно прибавлялись какие-то предыдущие строки (из числа первых p), помноженные на некоторые коэффициенты. Поэтому у матриц A и G_p одинаковы все миноры p-го порядка, которые содержатся в первых p строках, а также все миноры $(p + 1)$-го порядка, которые содержатся в строках с номерами $1, 2, \dots, p, i$ $(i > p)$:

$$\left.\begin{array}{c}
A\begin{pmatrix}1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p\end{pmatrix} = G_p\begin{pmatrix}1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p\end{pmatrix} \\
(1 \leq k_1 < k_2 < \dots < k_p \leq n), \\
A\begin{pmatrix}1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1}\end{pmatrix} = G_p\begin{pmatrix}1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1}\end{pmatrix} \\
(1 \leq k_1 < k_2 < \dots < k_{p+1} \leq n).
\end{array}\right\} \tag{2.13}$$

Из этих формул, учитывая структуру (2.12) матрицы G_p, найдем:

$$A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix} = a_{11}a_{22}^{(1)} \dots a_{pp}^{(p-1)},$$

$$A\begin{pmatrix}1 & 2 & \dots & p & i\\ 1 & 2 & \dots & p & k\end{pmatrix} = a_{11}a_{22}^{(1)} \dots a_{pp}^{(p-1)}a_{ik}^{(p)} \quad (i, k = p+1, \dots, n).$$

(2.14)

Деля почленно второе из этих равенств на первое, получим основные формулы

$$a_{ik}^{(p)} = \frac{A\begin{pmatrix}1 & 2 & \dots & p & i\\ 1 & 2 & \dots & p & k\end{pmatrix}}{A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}} \quad (i, k = p+1, \dots, n).$$

(2.15)

Если условия (2.10) выполнены для данного значения *p,* то такие же условия выполнены для любого меньшего значения *p.* Поэтому формулы (2.15) имеют место не только для данного значения *p,* но и для всех меньших значений *p.* То же можно сказать и о формуле (2.14). Поэтому вместо этой формулы можно написать равенства

$$A\begin{pmatrix}1\\ 1\end{pmatrix} = a_{11}, \quad A\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix} = a_{11}a_{22}^{(1)}, \quad A\begin{pmatrix}1 & 2 & 3\\ 1 & 2 & 3\end{pmatrix} = a_{11}a_{22}^{(1)}a_{33}^{(2)}, \dots$$

(2.16)

Таким образом, условия (2.10), т.е. необходимое и достаточное условия выполнения первых *p* этапов алгоритма Гаусса, могут быть записаны в виде следующих неравенств:

$$A\begin{pmatrix}1\\ 1\end{pmatrix} \neq 0, \quad A\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix} \neq 0, \quad \dots, \quad A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix} \neq 0.$$

(2.17)

Тогда из (2.16) находим:

$$a_{11} = A\begin{pmatrix}1\\ 1\end{pmatrix}, \quad a_{22}^{(1)} = \frac{A\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix}}{A\begin{pmatrix}1\\ 1\end{pmatrix}},$$

$$a_{33}^{(2)} = \frac{A\begin{pmatrix}1 & 2 & 3\\ 1 & 2 & 3\end{pmatrix}}{A\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix}}, \dots, \quad a_{pp}^{(p-1)} = \frac{A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}}{A\begin{pmatrix}1 & 2 & \dots & p-1\\ 1 & 2 & \dots & p-1\end{pmatrix}}.$$

(2.18)

Для того чтобы в алгоритме исключения Гаусса можно было последовательно исключить $x_1, x_2, ..., x_p$, нужно, чтобы все величины (2.18) были отличны от нуля, т.е. чтобы выполнялись неравенства (2.17). В то же время формулы для $a^{(p)}_{ik}$ имеют смысл, если выполняется только последнее из условий (2.17).

Пусть матрица коэффициентов в системе уравнений (2.1) имеет ранг r. Тогда соответствующей перестановкой уравнений и изменением нумерации неизвестных можно добиться выполнения неравенств

$$A\begin{pmatrix}1 & 2 & \dots & j\\ 1 & 2 & \dots & j\end{pmatrix} \neq 0 \qquad (j = 1, 2, \dots, r).$$

(2.19)

Это позволяет последовательно исключить $x_1, x_2, ..., x_r$ и получить систему уравнений

$$\left.\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \dots\dots\dots\dots + a_{1n}x_n = y_1,\\
\qquad a_{22}^{(1)}x_2 + \dots\dots\dots\dots + a_{2n}^{(1)}x_n = y_2^{(1)},\\
\qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\\
\qquad a_{rr}^{(r-1)}x_r + \dots\dots + a_{rn}^{(r-1)}x_n = y_r^{(r-1)},\\
\qquad a_{r+1,\ r+1}^{(r)}x_{r+1} + \dots + a_{r+1,\ n}^{(r)}x_n = y_{r+1}^{(r)},\\
\qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\\
\qquad a_{n,\ r+1}^{(r)}x_{r+1} + \dots + a_{nn}^{(r)}x_n = y_n^{(r)}.
\end{array}\right\}$$

(2.20)

Здесь коэффициенты определяются по формулам (2.15). Из этих формул, поскольку ранг матрицы $A = \| a_{ik} \|^n_1$ равен r, следует, что

$$a_{ik}^{(r)} = 0 \qquad (i,k = r+1, ...,n)$$

и матрица G_r, получающаяся из матрицы $A=\|a_{ik}\|^n_1$ после применения r-этапного алгоритма исключения Гаусса, имеет вид

$$G_r = \left\|\begin{array}{llllll}
a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots\ a_{1n}\\
0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots\ a_{2r}^{(1)}\\
\dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,\ r+1}^{(r-1)} & \dots\ a_{rn}^{(r-1)}\\
0 & 0 & \dots & 0 & 0 & \dots\ 0\\
\dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 0 & 0 & \dots\ 0
\end{array}\right\|.$$

(2.21)

Последние n — r уравнений (2.20) сводятся к условиям совместимости

$$y_i^{(r)} = 0 \quad (i=r+1,..., n). \tag{2.22}$$

Заметим, что столбец свободных членов при алгоритме исключения подвергается таким же преобразованиям, как и любой столбец коэффициентов. Поэтому, дополняя матрицу $A=\|a_{ik}\|^n_1$ $(n + 1)$-м столбцом из свободных членов, мы получим:

$$y_i^{(p)} = \frac{A\begin{pmatrix}1 \dots p & i \\ 1 \dots p & n+1\end{pmatrix}}{A\begin{pmatrix}1 \dots p \\ 1 \dots p\end{pmatrix}} \quad (i = 1, 2, \dots, n;\ p = 1, 2, \dots, r). \tag{2.23}$$

В частности, условия совместимости (2.22) сводятся к известным условиям

$$A\begin{pmatrix}1 \dots r & r+j \\ 1 \dots r & n+1\end{pmatrix} = 0 \quad (j = 1, 2, \dots, n-r). \tag{2.24}$$

Если $r = n$, т.е. матрица $A=\|a_{ik}\|^n_1$ неособенная, и

$$A\begin{pmatrix}1 & 2 \dots j \\ 1 & 2 \dots j\end{pmatrix} \neq 0 \quad (j = 1, 2, \dots, n),$$

то при помощи алгоритма Гаусса можно последовательно исключить $x_1, x_2, \dots x_{n-1}$ и привести систему уравнений к виду (2.9).

4. Механическая интерпретация алгоритма Гаусса. Рассмотрим произвольную упругую статическую систему S, закрепленную на краях (например, струну, стрежень, многопролетный стрежень, мембрану, пластину или дискретную систему), и возьмем на ней n точек (1), (2), ...,(n).

Мы будем рассматривать перемещение (прогибы) $y_1, y_2, \dots, y_n$ точек (1), (2), ..., (n) системы S под действием сил $F_1, F_2, \dots, F_n$, приложенных в этих же точках. Мы будем предполагать, что силы и перемещения параллельны одному и тому же направлению и потому определяются своими алгебраическими величинами (рис. 2.1).

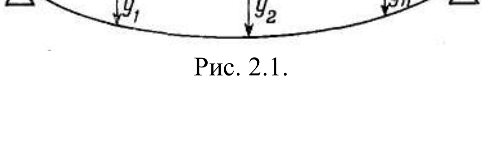


Рис. 2.1.

Кроме того, мы примем, что имеет место принцип линейного наложения сил:
1° *При суммарном наложении двух систем сил соответствующие прогибы складываются.*
2° При умножении величин всех сил на одно и то же вещественное число все прогибы умножаются на это число.

Обозначим через a_{ik} коэффициент влияния точки *(k)* на точку (*i*), т.е. прогиб в точке (*i)* под действием единичной силы, приложенной в точке (*k*) *(i,k*= 1,2, ..., *n)* (рис. 2.2).

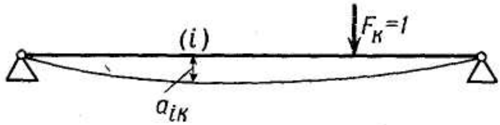


Рис. 2.2.

Тогда при совместном действии сил $F_1, F_2, ..., F_n$ прогибы $y_1, y_2, \ldots, y_n$ определятся по формулам

$$\sum_{k=1}^{n} a_{ik}F_k = y_i \quad (i = 1, 2, \ldots, n). \tag{2.25}$$

Сопоставляя (2.25) с исходной системой (2.1), мы задачу отыскания решения системы уравнений (2.1) можем интерпретировать так:

Даны прогибы y_1, y_2, ..., y_n. Ищутся соответствующие силы $F_1, F_2,..., F_n$.

Обозначим через S_p статическую систему, получающуюся из *S* введением *p* неподвижных шарнирных опор в точках (1), (2),..., (*p)* *(p≤n*). Коэффициенты влияния для оставшихся подвижных точек *(p* + 1),…,(*n*) системы S_p обозначим через $a^{(p)}_{ik}$ *(i,k*= *p*+1, ..., *n)* (см. рис. 2.3 для *p* — 1).

Коэффициент $a^{(p)}_{ik}$ можно рассматривать как прогиб в точке (*i*) системы *S* при действии единичной силы в точке *(k)* и сил реакций $R_1, R_2, ..., R_p$ в закрепленных точках (1), (2), ...,(*p)*.
Поэтому

$$a_{ik}^{(p)} = R_1a_{i1} + \ldots + R_pa_{ip} + a_{ik}. \tag{2.26}$$

С другой стороны, при этих же силах прогибы системы *S* в точках (1), (2), ..., (*p)* равны нулю:

$$\left.\begin{array}{l} R_1 a_{11} + \ldots + R_p a_{1p} + a_{1k} = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ R_1 a_{p1} + \ldots + R_p a_{pp} + a_{pk} = 0. \end{array}\right\}$$

(2.27)

Если

$$A\begin{pmatrix} 1 & 2 \ldots p \\ 1 & 2 \ldots p \end{pmatrix} \neq 0.$$

то мы можем из (2.27) определить R_1, R_2, ..., R_p и полученные выражения подставить в (2.24). Это исключение R_1, R_2, ..., R_p можно сделать и так. К системе равенств (2.27) прибавим равенство (2.26), записанное в виде

$$R_1 a_{i1} + \ldots + R_p a_{ip} + a_{ik} - a_{ik}^{(p)} = 0.$$

(2.28)

Рассматривая (2.27) и (2.28) как систему p+1 однородных уравнений, которая имеет ненулевое решение R_1, R_2, ..., R_p, R_{p+1}= 1, получаем, что определитель этой системы равен нулю:

$$\begin{vmatrix} a_{11} \ldots a_{1p} & a_{1k} \\ \dots\dots\dots\dots & \\ a_{p1} \ldots a_{pp} & a_{pk} \\ a_{i1} \ldots a_{ip} & a_{ik} - a_{ik}^{(p)} \end{vmatrix} = 0.$$

Отсюда

$$a_{ik}^{(p)} = \frac{A\begin{pmatrix} 1 & 2 \ldots p & i \\ 1 & 2 \ldots p & k \end{pmatrix}}{A\begin{pmatrix} 1 & 2 \ldots p \\ 1 & 2 \ldots p \end{pmatrix}} \quad (i, k = p + 1, \ldots, n).$$

(2.29)

По этим формулам коэффициенты влияния «опорной» системы S_p выражаются через коэффициенты влияния исходной системы S.

Но формулы (2.29) совпадают с формулами (2.16). Поэтому *для любого* $p(\leq n — 1)$ *коэффициенты* $a^{(p)}_{ik}$ (i, k = p+ 1, ..., n) *в алгоритме Гаусса есть коэффициентами влияния опорной системы* S_p.

В справедливости этого основного положения можно убедиться из чисто механических соображений, не опираясь на алгебраический вывод формул (2.16). Для этого рассмотрим сначала частный случай одной опоры: p = 1 (рис. 2.3).

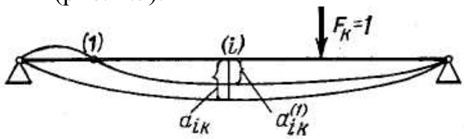


Рис. 2.3.

В этом случае коэффициенты влияния системы S_1 определятся по формулам [полагаем p= 1 в (2.29)]:

$$a_{ik}^{(1)} = \frac{A\begin{pmatrix}1 & i\\ 1 & k\end{pmatrix}}{A\begin{pmatrix}1\\ 1\end{pmatrix}} = a_{ik} - \frac{a_{i1}}{a_{11}} a_{1k} \quad (i, k = 1, 2, \ldots, n).$$

Эти формулы совпадают с формулами (2.6).

Таким образом, если коэффициенты a_{ik} $(i,k$= 1,2, ..., $n)$ в системе уравнений (2.1) являются коэффициентами влияния статической системы S, то коэффициенты $a^{(1)}_{ik}$ $(i,k$=2, ..., $n)$ в алгоритме Гаусса являются коэффициентами влияния системы S_1. Применяя эти же соображения к системе S_1 и вводя в ней вторую опору в точке (2), получим, что коэффициенты $a^{(2)}_{ik}$ $(i,k$= 3, ..., $n)$ в системе уравнений (2.7) являются коэффициентами влияния опорной системы S_2, и вообще для любого $p(\leq n$ — 1) коэффициенты $a^{(p)}_{ik}$ $(i,k = p + 1, ..., n)$ в алгоритме Гаусса являются коэффициентами влияния опорной системы S_n.

Из механических соображений очевидно, что *последовательное введение р опор равносильно одновременному введению этих опор.*

Замечание. Обращаем внимание на то, что при механической интерпретации алгоритма исключения не было необходимости предполагать, что точки, в которых рассматриваются прогибы, совпадают с точками приложения сил $F_1, F_2, ..., F_n$. Можно считать, что $y_1, y_2, ..., y_n$ — прогибы точек (1), (2), ..., (n), а силы F_1, F_2, ..., F_n приложены в точках (1'), (2'), ..., (n'). Тогда a_{ik} — коэффициент влияния точки *(k')* на точку (i). В этом случае вместо опоры в точке *(j)* следует рассматривать обобщенную опору в точках *(j), (j'),* при которой прогиб в точке *(j)* поддерживается все время равному нулю за счет надлежащим образом выбранной вспомогательной силы R_j в точке *(j').* Условие возможности введения p обобщенных опор в точках (1), (1');

(2), (2'); …; (*p*), (*p*'), т.е. возможность удовлетворить условиям $y_1 = 0$, $y_2 = 0$, ..., $y_p = 0$ при любых F_{p+1}, ..., F_n за счет соответствующих $R_1 = F_t$, ...,$R_p = F_p$, выражается неравенством

$$A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix} \neq 0.$$

5. Детерминантное тождество Сильвестра. В п.1 путем сопоставления матриц A и G_p мы пришли к равенствам (2.13) и (2.14).

Эти равенства позволяют сразу получить важное детерминантное тождество Сильвестра. Действительно, из (2.13) и (2.14).
находим:

$$|A| = A\begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix} = A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}\begin{vmatrix} a^{(p)}_{p+1,\,p+1} & \dots & a^{(p)}_{p+1,\,n}\\ \dots & \dots & \dots\\ a^{(p)}_{n,\,p+1} & \dots & a^{(p)}_{nn}\end{vmatrix}.$$

(2.30)

Введем в рассмотрение окаймляющий минор

$$A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}$$

определители:

$$b_{ik} = A\begin{pmatrix}1 & 2 & \dots & p & i\\ 1 & 2 & \dots & p & k\end{pmatrix} \quad (i, k = p+1, \dots, n).$$

Матрицу, составленную из этих определителей, обозначим через

$$B = ||b_{ik}||^{n}_{p+1}.$$

Тогда согласно формулам (2.15)

$$\begin{vmatrix} a^{(p)}_{p+1,\,p+1} & \dots & a^{(p)}_{p+1,\,n}\\ \dots & \dots & \dots\\ \dots & \dots & \dots\\ a^{(p)}_{n,\,p+1} & \dots & a^{(p)}_{nn}\end{vmatrix} = \frac{\begin{vmatrix} b_{p+1,\,p+1} & \dots & b_{p+1,\,n}\\ \dots & \dots & \dots\\ \dots & \dots & \dots\\ b_{n,\,p+1} & \dots & b_{nn}\end{vmatrix}}{\left[A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}\right]^{n-p}} = \frac{|B|}{\left[A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}\right]^{n-p}}.$$

Поэтому равенство (2.30) может быть записано так:

$$|B| = \left[A\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & p\end{pmatrix}\right]^{n-p-1}|A|.$$

(2.31)

Это и есть детерминантное тождество Сильвестра. Оно выражает определитель $|B|$, составленный из окаймляющих определителей, через исходный определитель и окаймляемый минор.

Равенство (2.31) было установлено для матриц $A=\|a_{ik}\|_1^n$, элементы которых удовлетворяют неравенствам

$$A\begin{pmatrix}1 & 2 & \dots & j\\ 1 & 2 & \dots & j\end{pmatrix} \neq 0 \qquad (j=1, 2, \dots, p). \tag{2.32}$$

Однако из «соображений непрерывности» следует, что эти ограничения можно отвергнуть и что тождество Сильвестра справедливо для любой матрицы $A = \| a_{ik} \|_1^n$. В самом деле, пусть неравенства (2.32) не выполняются. Введем матрицу

$$A_\varepsilon = A + \varepsilon E.$$

Очевидно, $\lim\limits_{\varepsilon\to 0} A_\varepsilon = A.$. С другой стороны, миноры

$$A_\varepsilon\begin{pmatrix}1 & 2 & \dots & j\\ 1 & 2 & \dots & j\end{pmatrix} = \varepsilon^j + \dots \qquad (j=1, 2, \dots, p)$$

представляют собой p не равных тождественно нулю многочленов относительно ε. Поэтому можно выбрать такую последовательность $\varepsilon_m \to 0$, что

$$A_{\varepsilon m}\begin{pmatrix}1 & 2 & \dots & j\\ 1 & 2 & \dots & j\end{pmatrix} \neq 0 \qquad (j=1, 2, \dots, p;\ m=1, 2, \dots).$$

Для матрицы $A_{\varepsilon m}$ мы можем написать тождество (2.31). Переходя в обеих частях этого тождества к пределу при $m\to\infty$, мы получим тождество Сильвестра для предельной матрицы

$$A = \lim_{m\to\infty} A_{\varepsilon m}$$

Под *пределом* (при $p\to\infty$) *последовательности матриц* $B_p = \| b^{(p)}_{ik} \|$ понимают матрицу $B = \| b_{ik} \|_1^n$, где

$$b_{ik} = \lim_{p\to\infty} b_{ik}^{(p)} \ (i, k = 1, 2, \dots, n).$$

Если мы тождество (2.31) применим к определителю

$$A\begin{pmatrix}1 & 2 & \dots & p & i_1 & i_2 & \dots & i_q\\ 1 & 2 & \dots & p & k_1 & k_2 & \dots & k_q\end{pmatrix} \qquad \left(p < \begin{matrix} i_1 < i_2 < \dots < i_q \\ k_1 < k_2 < \dots < k_q\end{matrix} \leq n\right),$$

то мы получим удобный для применений вид тождество Сильвестра:

$$B\begin{pmatrix} i_1 & i_2 & \dots & i_q \\ k_1 & k_2 & \dots & k_q \end{pmatrix} = \left[A\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}\right]^{q-1} A\begin{pmatrix} 1 & 2 & \dots & p & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & k_2 & \dots & k_q \end{pmatrix}. \tag{2.33}$$

2.3. *LU*- разложение квадратных матриц

1. *LU*- разложение. Перед тем как перейти к рассмотрению вопросов *LU*- разложение, напомним этапы решения неоднородних систем линейных уравнений с помощью алгоритма Гаусса.

Подобно методу исключения при обращении матрицы, это достигается соответствующими операциями над строками расширенной матрицы системы $[A, b]$ размера $n\times(n+1)$. Расхождение заключается в том, что в нули преобразуются лишь те элементы матрицы A, которые расположены ниже ее главной диагонали. В результате $[A, b]$ приводится к матрице $[U, y]$, где U — верхняя треугольная матрица с единичными элементами на главной диагонали; y - преобразованный столбец свободных членов:

$$U = \begin{bmatrix} 1 & u_{12} & u_{13} & \cdots & u_{1n} \\ & 1 & u_{23} & \cdots & u_{2n} \\ & & \cdots & \cdots & \cdots \\ & & & \cdots & 1 \end{bmatrix}; \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}.$$

Преобразованная система имеет вид:

$$\left.\begin{aligned} x_1 + u_{12}x_2 + u_{13}x_3 + \cdots + u_{1n}x_n &= y_1 \\ x_2 + u_{23}x_3 + \cdots + u_{2n}x_n &= y_2 \\ \cdots\cdots\cdots\cdots\cdots & \\ x_n &= y_n \end{aligned}\right\},$$

откуда на основании $x_n = y_n$ находим последовательно $x_{n-1},\dots, x_1$ по формуле:

$$x_k = y_k - \sum_{s=k+1}^{n} u_{ks}x_s \quad (k = n-1, \dots, 1).$$

Итак, алгоритм Гаусса содержит два этапа: 1) построение вспомогательной системы с треугольной матрицей U *(прямой ход);*

2) получение решения системы x *(обратный ход).* Проиллюстрируем его на следующем примере:

$$\left.\begin{aligned} 2x_1 + x_2 + 4x_3 &= 16 \\ 3x_1 + 2x_2 + x_3 &= 10 \\ x_1 + 3x_2 + 3x_3 &= 16 \end{aligned}\right\}.$$

Прямой ход включает следующие преобразования расширенной матрицы:

$$\left[\begin{array}{ccc|c} 2 & 1 & 4 & 16 \\ 3 & 2 & 1 & 10 \\ 1 & 3 & 3 & 16 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & 2 & 8 \\ 0 & \frac{1}{2} & -5 & -14 \\ 0 & \frac{5}{2} & 1 & 8 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & 2 & 8 \\ 0 & 1 & -10 & -28 \\ 0 & 0 & 26 & 78 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & 2 & 8 \\ 0 & 1 & -10 & -28 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

Преобразованная система уравнений имеет вид:

$$\left.\begin{aligned} x_1 + \frac{1}{2}x_2 + 2x_3 &= 8 \\ x_2 - 10x_3 &= -28 \\ x_3 &= 3 \end{aligned}\right\},$$

откуда находим: $x_3 = 3$; $x_2 = -28 + 10x_3 = 2$;

$$x_1 = 8 - \frac{1}{2}x_2 - 2x_3 = 1.$$

Если исключение выполнить так же, как при обращении матрицы, то в результате A преобразуется в единичную, а b — в столбец, элементы которого равны значениям искомых величин $x_1, x_2, \ldots, x_n$, т.е. $[A, b]$ преобразуется в $[1, x]$. Эта разновидность метода исключения называют *алгоритмом Гаусса-Жордана.* Так, для рассмотренного примера имеем

$$\left[\begin{array}{ccc|c} 2 & 1 & 4 & 16 \\ 3 & 2 & 1 & 10 \\ 1 & 3 & 3 & 16 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & 2 & 8 \\ 0 & \frac{1}{2} & -5 & -14 \\ 0 & \frac{5}{2} & 1 & 8 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|c} 1 & 0 & 7 & 22 \\ 0 & 1 & -10 & -28 \\ 0 & 0 & 26 & 78 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right],$$

откуда сразу получаем: $x_1 = 1$, $x_2 = 2$, $x_3 = 3$.

Разработано много различных вариантов гауссова исключение, которые отличаются выбором ведущих элементов, способами уточнения решения, распределением оперативной памяти при использовании вычислительных машин. По мнению специалистов, для решения систем линейных уравнения не найдено лучших по объему вычислений и точности алгоритмов, чем методы последовательного исключения. Все их разновидности связаны по сути с разложением неособенной квадратной матрицы в произведение двух треугольных матриц — нижней L и верхней U, т.е. $A = LU$. Это разложение называют *треугольным* или *LU-разложением.*

Доказательство возможности *LU-разложения* обычно проводится методом математической индукции. Пусть имеется квадратная матрица A n-го порядка. При $n = 1$ имеем $[a_{11}]=[l_{11}][u_{11}]$, что является LU -разложением, так как всякая матрица первого порядка может рассматриваться и как треугольная. Покажем, что если это разложение имеет место для матрицы A_{n-1} $(n — 1)$-го порядка, то оно возможно и для матрицы A n-го порядка. Представив соотношение $A = LU$ в блочном виде и перемножив матрицы, получим:

$$\begin{bmatrix} A_{n-1} & w \\ v & a_{nn} \end{bmatrix} = \begin{bmatrix} L_{n-1} & 0 \\ x & l_{nn} \end{bmatrix} \begin{bmatrix} U_{n-1} & y \\ 0 & u_{nn} \end{bmatrix} = \begin{bmatrix} L_{n-1}\,U_{n-1} & L_{n-1} \quad y \\ x \quad U_{n-1} & xy + l_{nn}\,u_{nn} \end{bmatrix},$$

откуда путем сравнения соответствующих элементов находим:

$$L_{n-1}U_{n-1} = A_{n-1}; \quad L_{n-1}y = w;$$
$$xU_{n-1} = v; \quad xy + l_{nn}u_{nn} = a_{nn}.$$

Первое уравнение удовлетворяется по предположению. Второе и третье уравнение позволяют определить $x = vU^{-1}_{n-1}$ и $y = wL^{-1}_{n-1}$ если U_{n-1} и L_{n-1} — неособенные матрицы. Так как в соответствии с теоремой об определителе произведения двух матриц $|L_{n-1}||U_{n-1}|=|A_{n-1}|$, то из требования $|L_{n-1}|\neq 0$ и $|U_{n-1}|\neq 0$ вытекает условие $|A_{n-1}|\neq 0$. Это значит, что недиагональные элементы матриц L и U определяются однозначно, если главные миноры матрицы A, составленные из ее первых k строк и столбцов (k = 1, 2, ..., n — 1), не равны нулю. Наконец, из последнего уравнения видно, что диагональные элементы l_{nn} и u_{nn} матриц L и U связаны соотношением $l_{nn}u_{nn}=a_{nn}—xy$. Поэтому для однозначного определения LU - разложение следует одному з них на каждом шаге приписывать некоторое значение, отличное от нуля. Обычно полагают все диагональные элементы одной из матриц L или U равными единицы. В дальнейшем будем считать u_{ii} = 1 (i= 1, 2, ..., n). Таким образом, показана возможность LU-разложения и условия, при которых матрицы L и U определяются однозначно. Так как определитель треугольной матрицы равен произведению ее диагональных элементов, то из $A = LU$ также следует

$$|A|=|L||U|=1\cdot(l_{11}l_{22}...l_{nn})=\prod_{i=1}^{n} l_{ii}\ .$$

Разложение матрицы A в произведение LU позволяет представить систему $Ax=b$ в виде $LUx=b$, и она сводится к двум систем:

$$Ly = b; \quad Ux = y.$$

Благодаря треугольной структуре L и U, эти уравнения решаются без обращения матриц последовательной подстановкой:

$$y_i=\frac{1}{l_{ii}}\left(b_i-\sum_{j=i}^{i-n} l_{ij}y_j\right)\quad (i=1,\ 2,\ \ldots,\ n);$$

$$x_k=y_k-\sum_{s=k+1}^{n} u_{ks}x_s\quad (k=n,\ n-1,\ \ldots,\ 1).$$

Компактная схема. Если условия единственности LU-разложение для матрицы A выполняются, т.е.

$$a_{11}\neq 0,\ \begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}\neq 0,\ \ldots,\ |A|\neq 0,$$

и принято соглашение о том, что u_{ii} = 1 (i = 1, 2, ..., n), то элементы матриц L и U вычисляются по *компактной схеме* на основе рекуррентных формул:

$$l_{ij} = a_{ij} - \sum_{s=1}^{j-1} l_{is} u_{sj} \quad (i \geqslant j);$$

$$u_{ij} = \frac{1}{l_{ii}} \left(a_{ij} - \sum_{s=1}^{i-1} l_{is} u_{sj} \right) \ (i < j).$$

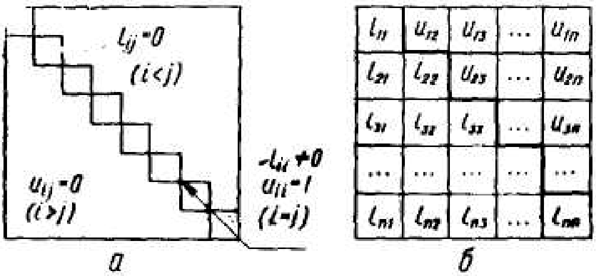


Рис. 2.4. LU -разложение неособенной матрицы:
а — структура матриц L и U ; *б* — компактная схема.

Справедливость этих соотношений непосредственно вытекает из структуры треугольных матриц L и U (рис. 2.4, *а)* и общего вида элемента произведения $A — LU$:

$$a_{ij} = \sum_{s=1}^{n} l_{is} u_{sj}.$$

Учитывая, что элементы u_{ij} матрицы U ниже главной диагонали ($i > j$) равны нулю, а на главной диагонали ($i=j$) равны единице, получаем первую из приведенных выше рекуррентных формул. Аналогично, на основании того, что элементы l_{ij} матрицы L выше главной диагонали ($i<j$) равны нулю, получаем другу формулу.

На первом шаге $l_{i1} = a_{i1}$ ($i = 1, 2, \ldots, n$) и $u_{1j}=a_{1j}/l_{11}=a_{1j}/a_{11}$ ($j=2,3,\ldots, n$). На втором шаге, используя значение l_{i1} и u_{1j}, вычисляем $l_{i2}= a_{i2} — l_{i1}u_{12}$ ($i = 2, 3.....n$) и $u_{2j} = (a_{1j}/l_{22})(a_{2j} — l_{21}u_{1j})$ ($j = 3, 4, \ldots,n$) и т.д. Процедура заканчивается на n-м шаге, причем диагональные элементы u_{ii} не вычисляются, так как по условию все они равны единице.

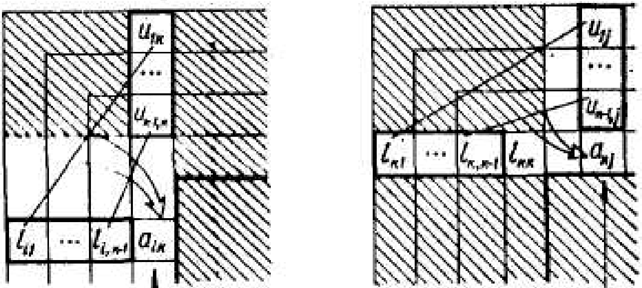
Как видно, для вычисления элементов матриц L и U соответствующий элемент матрицы A используется только один раз, и в дальнейших операциях он не принимает участие. Поэтому при реализации алгоритма на вычислительных машинах найденные на данном шаге элементы l_{ik}, и u_{kj} обычно заносятся в память на места освобождающихся элементов матрицы A. В результате LU -разложения представляется таблицей, которая изображена на рис. 2.4, *б*. При этом элементы u_{ii}=1 не хранятся, а их значения учитываются программой при соответствующих операциях. Так, для предыдущего примера имеем:

$$\begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 2 & \frac{1}{2} & 2 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 2 & \frac{1}{2} & 2 \\ 3 & \frac{1}{2} & -10 \\ 1 & \frac{5}{2} & 3 \end{bmatrix} \to \begin{bmatrix} 2 & \frac{1}{2} & 2 \\ 3 & \frac{1}{2} & -10 \\ 1 & \frac{5}{2} & 26 \end{bmatrix};$$

$$L = \begin{bmatrix} 2 & 0 & 0 \\ 3 & \frac{1}{2} & 0 \\ 1 & \frac{5}{2} & 26 \end{bmatrix}; \quad U = \begin{bmatrix} 1 & \frac{1}{2} & 2 \\ 0 & 1 & -10 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$|A| = 2 \cdot \frac{1}{2} \cdot 26 = 26.$$

Вычисления на k-м шаге по компактной схеме иллюстрируются на рис. 2.5 (заштрихованные части матриц не принимают участие в соответствующих операциях).



$$l_{ik} = a_{ik} - \sum_{s=1}^{k-1} l_{is} u_{sk} \qquad u_{kj} = \frac{1}{l_{kk}}\left(a_{kj} - \sum_{s=1}^{k-1} l_{ks} u_{sj}\right)$$

Рис. 2.5. Вычисление на k-м шаге по компактной схеме.

Исходная система уравнений сводится к двум простым системам:

$$\left.\begin{aligned} 2y_1 &= 16 \\ 3y_1 + \frac{1}{2} y_2 &= 10 \\ y_1 + \frac{5}{2} y_2 + 26 y_3 &= 16 \end{aligned}\right\}; \qquad \left.\begin{aligned} x_1 + \frac{1}{2} x_2 + 2x_3 &= y_1 \\ x_2 - 10x_3 &= y_2 \\ x_3 &= y_3 \end{aligned}\right\}.$$

Решение первого из них: y_1 =8; y_2= 2 (10-3• 8) =-28; y_3=(1/26) (16 -8+ 5/2 • 28)= 3. Из второго уравнения находим: x_3 = 3; x_2 = —28 + 10 • 3=2; x_1= 8— (1/2)• 2— 2• 3=1, что совпадает с полученным ранее результатом.

Рассмотренную схему часто называют *алгоритмом Крауша*.

Следует иметь в виду, что при использовании компактной схемы структура матриц L и U определяется фиксацией ведущих элементов по главной диагонали. Если на каком-нибудь шаге диагональный элемент l_{ii} окажется равным или близким нулю, то дальнейшее продолжение процедуры по этой схеме невозможно, а всякое изменение стратегии выбора ведущих элементов потребовало бы возвращение к ее началу. Поэтому, если нет уверенности в соблюдении условий, необходимых для LU-разложения при заранее фиксированных ведущих элементах, следует прибегать к другим методам, один из которых рассматривается ниже.

Получение LU-разложении методом исключения. Верхняя треугольная матрица U формируется в процессе гауссова исключение. Поэтому для получения LU-разложения достаточно дополнить алгоритм Гаусса так, чтобы формировалась и матрица L. В то же время можно освободить этот алгоритм от операций над элементами столбца свободных членов b, так как преобразованный столбец получается как $y = L^{-1}b$.

Выясним способ образования матрицы L в процессе исключения. Пусть на (k—1)-в шаге матрица A преобразовалась в $A^{(k-1)}$ с элементами $a_{ij}^{(k-1)}$. При этом $a_{ii}^{(k-1)}$= 1 (i = 1, 2..... k —1) и $a_{ij}^{(k-1)}$= 0 (i >j; j = 1, 2, ..., k — 1). Преобразование на k-м шаге приводит к матрице $A^{(k)}$ с элементами $a_{ij}^{(k)}$, причем

$$a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}; \qquad a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} a_{ik}^{(k-1)},$$

где $i = k + 1, ..., n; j = k, k + 1, ..., n.$

Легко понять, что это преобразование соответствует разложению $A^{(k-1)}$ в произведение $W_kA^{(k)}$, где матрица W_k имеет простую структуру:

$$W_k = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & a_{kk}^{(k-1)} & & \\ & & \cdots & & \\ & & & \ddots & \\ & & a_{nk}^{(k-1)} & \cdots & 1 \end{bmatrix}.$$

Процесс гауссова исключение преобразует A в U за n шагов, что соответствует последовательному разложению

$A = (W_1 W_2 \dots \dots W_n)U = LU$,откуда

$$L = W_1 W_2 \dots W_n = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22}^{(1)} & 0 & \cdots & 0 \\ a_{31} & a_{32}^{(1)} & a_{33}^{(2)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2}^{(1)} & a_{n3}^{(2)} & \cdots & a_{nn}^{(n-1)} \end{bmatrix}.$$

Таким образом, матрица L формируется последовательно в процессе исключения, причем ее k-й столбец состоит из тех значений элементов k-го столбца матрицы A, которые они получают на $(k - 1)$-м шаге преобразования. Практически при проведении процедуры гауссова исключение достаточно на каждом шаге по главной диагонали и ниже ее сохранять значения элементов преобразованной матрицы, вычисленных на предыдущем шаге. В результате получаем квадратную таблицу (рис. 2.4, *б),* которая и представляет LU-разложение. Например,

$$\begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 2 & \frac{1}{2} & 2 \\ 3 & \frac{1}{2} & -5 \\ 1 & \frac{5}{2} & 1 \end{bmatrix} \to \begin{bmatrix} 2 & \frac{1}{2} & 2 \\ 3 & \frac{1}{2} & -10 \\ 1 & \frac{5}{2} & \mathbf{26} \end{bmatrix},$$

что соответствует матрицам L и U, приведенным ранее.

Такой способ получения LU-разложения не связан с предарительной фиксацией ведущих элементов, и на каждом шаге ведущий элемент может быть выбран в любой строке и столбце, в которых такие элементы еще не выбирались на предыдущих шагах.

Это позволяет повышать точность вычислений путем выбора в качестве ведущего элемента наибольший по абсолютной величине элемент очередного столбца (в специальной литературе можно найти и другие рекомендации).

Если не все ведущие элементы расположены на главной диагонали, то матрицы L и U будут, вообще говоря, не треугольными. Но они легко приводятся к треугольным соответствующей перестановкой строк (при машинной реализации этого алгоритма вместо перестановки строк просто изменяется их нумерация). Вернемся к рассмотренному выше примеру и выполним LU-разложение, выбирая ведущий в каждом столбце наибольший элемент:

$$\begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 2 & \frac{1}{3} & \frac{10}{3} \\ 3 & \frac{2}{3} & \frac{1}{3} \\ 1 & \frac{7}{3} & \frac{8}{3} \end{bmatrix} \to \begin{bmatrix} 2 & -\frac{1}{3} & \frac{26}{7} \\ 3 & \frac{2}{3} & \frac{1}{3} \\ 1 & \frac{7}{3} & \frac{8}{7} \end{bmatrix},$$

что после перестановки строк соответствует LU-разложению:

$$L = \begin{bmatrix} 3 & 0 & 0 \\ 1 & \frac{7}{3} & 0 \\ 2 & -\frac{1}{3} & \frac{26}{7} \end{bmatrix}; \quad U = \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & \frac{8}{7} \\ 0 & 0 & 1 \end{bmatrix}.$$

Как видно из рассмотренного примера, при разной стратегии выбора ведущих элементов матрицы L и U получаются различными (по сути здесь L и U соответствуют не матрице A, а той матрице, которая получается из A перестановкой строк).

2.4. Разложение квадратной матрицы на треугольные множители

1. Пусть дана матрица $A = \| a_{ik} \|_1^n$ ранга r. Введем следующие обозначения для последовательных главных миноров этой матрицы:

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \quad (k = 1, 2, \dots, n).$$

Предположим, что имеют место условия для выполнения алгоритма Гаусса:

$$D_k \neq 0 \quad (k=1,2,\dots, r).$$

Обозначим через G матрицу коэффициентов системы уравнений (2.20), к которой приводится система уравнений

$$\sum_{k=1}^{n} a_{ik}x_k = y_i \qquad (i = 1, 2, \ldots, n)$$

методом исключения Гаусса. Матрица G имеет верхнюю треугольную форму, причем элементы ее первых r строк определяются формулами (2.15), а элементы последних n-r строк все равны нулю (матрица G совпадает с матрицей G_r):

$$G = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1r} & a_{1,r+1} & \ldots & a_{1n} \\ 0 & a_{22}^{(1)} & \ldots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \ldots & a_{2n}^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \ldots & a_{rn}^{(r-1)} \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{Vmatrix}.$$

Переход от матрицы A к матрице G происходил при помощи некоторого числа N операций следующего типа: к i -й строке матрицы прибавлялся j-я $(j<i)$ строка, предварительно умноженная на некоторое число α. Такая операция равносильна умножению преобразованной матрицы слева на матрицу

$$\begin{array}{c} \begin{array}{ccccccc} & & (j) & & (i) & & \end{array} \\ \begin{Vmatrix} 1 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ \vdots & & 1 & & \vdots & & \vdots \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \ldots & \alpha & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{Vmatrix} \end{array} \qquad (2.34)$$

В этой матрице на главной диагонали стоят единицы, а все другие элементы, за исключением элемента α, равны нулю.

Таким образом,

$$G=W_N \ldots W_2W_1A,$$

где каждая из матриц W_1, W_2, ..., W_N имеет вид (2.34) и, следовательно, является нижней треугольной матрицей с диагональными элементами, равными 1.

Пусть

$$W = W_N \dots W_2 W_1. \tag{2.35}$$

Тогда

$$G = WA. \tag{2.36}$$

Матрицу W будем называть *преобразующей* матрицей для матрицы A в методе исключения Гаусса. Обе матрицы, G и W, однозначно определяются заданием матрицы A. Из (2.35) следует, что W — нижняя треугольная матрица с диагональными элементами, равными 1.

Поскольку W — неособенная матрица, то из (2.36) находим:

$$A = W^{-1}W. \tag{37}$$

Мы представили матрицу A в виде произведения нижней треугольной матрицы W^{-1} на верхнюю треугольную матрицу G. Вопрос о разложении матрицы A на множители такого типа полностью выясняется следующей теоремой:

Теорема 1. *Всякую матрицу $A = \| a_{ik} \|_1^n$ ранга r, у которой первые r последовательных главных миноров отличны от нуля,*

$$D_k = A\begin{pmatrix} 1 & 2 \dots k \\ 1 & 2 \dots k \end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r), \tag{2.38}$$

можно представить в виде произведения нижней треугольной матрицы B на верхнюю треугольную матрицу C

$$A = BC = \begin{Vmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{Vmatrix} \begin{Vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{nn} \end{Vmatrix}. \tag{2.39}$$

При этом

$$b_{11}c_{11} = D_1,\ b_{22}c_{22} = \frac{D_2}{D_1},\ \dots,\ b_{rr}c_{rr} = \frac{D_r}{D_{r-1}}. \tag{2.40}$$

Первым r диагональным элементам матриц B и C можно дать произвольные значения, которые удовлетворяют условиям (2.40).

Задание первых r диагональных элементов матриц B и C определяет однозначно элементы первых r столбцов матрицы B и *первых r строк матрицы C. Для этих элементов имеют место формулы*

$$b_{gk} = b_{kk}\frac{A\begin{pmatrix}1 & 2 \dots k-1 & g\\ 1 & 2 \dots k-1 & k\end{pmatrix}}{A\begin{pmatrix}1 & 2 \dots k\\ 1 & 2 \dots k\end{pmatrix}}, \quad c_{kg} = c_{kk}\frac{A\begin{pmatrix}1 & 2 \dots k-1 & k\\ 1 & 2 \dots k-1 & g\end{pmatrix}}{A\begin{pmatrix}1 & 2 \dots k\\ 1 & 2 \dots k\end{pmatrix}}$$

$$(g = k, k+1, \dots, n;\ k = 1, 2, \dots, r).$$

(2.41)

В случае $r < n(|A|=0)$ в последних n — r столбцах матрицы B можно все элементы положить равными нулю, а в последних n — r строках матрицы C всем элементам дать произвольные значения, или наоборот, последние n — r строк матрицы C заполнить нулями, а последние n — r столбцов матрицы B взять произвольными.

Доказательство. Возможность представления матрицы, которая удовлетворяет условию (2.38), в виде произведения (2.39) была доказана выше [см. (2.37)].

Пусть теперь *B* и *C* — произвольные нижняя и верхняя треугольные матрицы, произведение которых равны *A*. Пользуясь формулой для миноров произведения двух матриц, найдем:

$$A\begin{pmatrix}1 & 2 \dots & k-1 & g\\ 1 & 2 \dots & k-1 & k\end{pmatrix} = \sum_{\alpha_1<\alpha_2<\dots<\alpha_k} B\begin{pmatrix}1 & 2 & \dots & k-1 & g\\ \alpha_1 & \alpha_2 & \dots & \alpha_{k-1} & \alpha_k\end{pmatrix} C\begin{pmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_k\\ 1 & 2 & \dots & k\end{pmatrix}$$

$$(g = k, k+1, \dots, n, k = 1, 2, \dots, r).$$

(2.42)

Поскольку *C* — верхняя треугольная матрица, то первые *k* столбцов матрицы *C* содержат только один отличный от нуля минор *k*-го порядка

$$C\begin{pmatrix}1 & 2 & \dots & k\\ 1 & 2 & \dots & k\end{pmatrix}.$$

Поэтому равенство (2.42) может быть записано так:

$$A\begin{pmatrix}1 & 2 \dots & k-1 & g\\ 1 & 2 \dots & k-1 & k\end{pmatrix} = B\begin{pmatrix}1 & 2 \dots & k-1 & g\\ 1 & 2 \dots & k-1 & k\end{pmatrix} C\begin{pmatrix}1 & 2 \dots k\\ 1 & 2 \dots k\end{pmatrix} =$$

$$= b_{11}b_{22} \dots b_{k-1,\,k-1}b_{gk}c_{11}c_{22} \dots c_{kk}$$

$$(g = k, k+1, \dots, n;\ k = 1, 2, \dots, r).$$

(2.43)

Положим сначала здесь $g = k$. Тогда получим:

$$b_{11}b_{22} \dots b_{kk}c_{11}c_{22} \dots c_{kk}=D_k\ (k=1, 2, \dots, r), \qquad (2.44)$$

откуда уже вытекают соотношения (2.40).

Не нарушая равенства (2.39), мы можем в нем умножить матрицу B справа на произвольную неособенную диагональную матрицу $M=\|\mu_i\delta_{ik}\|^n_1$, одновременно умножая матрицу C слева на $M^{-1}=\|\mu^{-1}{}_i\delta_{ik}\|\,^n{}_1$. Это равносильно умножению столбцов матрицы B соответственно на μ_1, μ_2, ..., μ_n и строк матрицы C на $\mu^{-1}{}_1$, $\mu^{-1}{}_2$, ..., $\mu^{-1}{}_n$. Поэтому диагональным элементам b_{11}, ..., b_{rr}, c_{11}, ..., c_{rr} можно придать любые значения, которые удовлетворяют условиям (2.40).

Далее, из (2.43) и (2.44) находим:

$$b_{gk}=b_{kk}\frac{A\begin{pmatrix}1 & 2\ldots & k-1 & g\\ 1 & 2\ldots & k-1 & k\end{pmatrix}}{A\begin{pmatrix}1 & 2\ldots & k\\ 1 & 2\ldots & k\end{pmatrix}}\qquad (g=k,k+1,\ldots,n;\ k=1,2,\ldots,r),$$

т.е. первые формулы (2.41). Совершенно аналогично устанавливаются вторые формулы (2.41) для элементов матрицы C.

Обратим внимание на то, что при перемножении матриц B и C элементы b_{kg} последних n — r столбцов матрицы B элементы c_{gk} последних n — r строк матрицы C перемножаются только между собой. Мы видели, что все элементы последних n — r строк матрицы C можно выбрать равными нулю (это следует из (2.37). При этом диагональным элементам b_{11},...,b_{rr}, c_{11},..., c_{rr}, как было уже показано, можно придать любые значения, которые удовлетворяют условиям (2.40) за счет надлежащих множителей μ_1, μ_2, ..., μ_r). Тогда элементы последних n — r столбцов матрицы B можно выбрать произвольными. Ясно, что произведение матриц B и C не изменится, если мы последние n — r столбцов матрицы B возьмем нулевыми, а элементы последних n — r строк матрицы C произвольными.

Теорема доказана.

Из доказанной теоремы вытекает ряд следствий.

Следствие 1. *Элементы первых r столбцов матрицы B и первых r строк матрицы C связаны с элементами матрицы A рекуррентными соотношениями:*

$$\left.\begin{aligned} b_{ik}&=\frac{a_{ik}-\sum\limits_{j=1}^{k-1}b_{ij}c_{jk}}{c_{kk}}\qquad (i\geqslant k;\ i=1,2,\ldots,n;\ k=1,2,\ldots,r),\\ c_{ik}&=\frac{a_{ik}-\sum\limits_{j=1}^{i-1}b_{ij}c_{jk}}{b_{ii}}\qquad (i\leqslant k;\ i=1,2,\ldots,r;\ k=1,2,\ldots,n).\end{aligned}\right\}$$

(2.45)

Соотношение (2.45) непосредственно следует из матричного равенства (2.39), ими удобно пользоваться для фактического вычисления элементов матриц B и C.

Следствие 2. *Если* $A = \| a_{ik} \|^n_1$ — *неособенная матрица* ($r = n$), *удовлетворяющая условию* (2.38), *то в представлении* (2.39) *матрицы* B *и* C *определяются однозначно, как только диагональные элементы этих матриц избраны в соответствии с условиями* (2.40).

Следствие 3. *Если* $S=\|s_{ik}\|^n_1$ — *симметричная матрица ранга* r *и*

$$D_k = S\begin{pmatrix}1 & 2 \dots k\\ 1 & 2 \dots k\end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r),$$

то

$$S = BB',$$

где $B = \| b_{ik} \|^n_1$ — *нижняя треугольная матрица, в которой*

$$b_{gk} = \begin{cases} \dfrac{1}{\sqrt{D_k D_{k-1}}} A\begin{pmatrix}1 & 2 \dots & k-1 & g\\ 1 & 2 \dots & k-1 & k\end{pmatrix} & (g = k, k+1, \dots, n;\ k = 1, 2, \dots, r),\\ 0 & (g = k, k+1, \dots, n;\ k = r+1, \dots, n). \end{cases}$$

(2.46)

2. Пусть в представлении (2.40) в матрице B элементы последних $n — r$ столбцов равны нулю. Тогда можно положить:

$$B = F\cdot\begin{Vmatrix} b_{11} & & & & 0\\ & \ddots & & & \\ & & b_{rr} & & \\ & & & 0 & \\ & & & & \ddots & \\ 0 & & & & & 0\end{Vmatrix}, \qquad C = \begin{Vmatrix} c_{11} & & & & 0\\ & \ddots & & & \\ & & c_{rr} & & \\ & & & 0 & \\ & & & & \ddots & \\ 0 & & & & & 0\end{Vmatrix}\cdot L,$$

(2.47)

где F — нижняя, а L — верхняя треугольная матрица; при этом первые r диагональных элементов у матриц F и L равны 1, а элементы последних $n — r$ столбцов матрицы F и последних $n — r$ строк матрицы L выбраны совершенно произвольно. Подставляя в (2.39) выражение (2.47) для B и C и используя равенства (2.40), придем к следующей теореме:

Теорема 2. *Всякая матрица* $A = \| a_{ik} \|^n_1$ *ранга* r, *у которой*

$$D_k = A\begin{pmatrix}1 & 2 \dots k\\ 1 & 2 \dots k\end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r),$$

представима в виде произведения нижней треугольной матрицы F, диагональной D и верхней треугольной L:

$$A = FDL = \begin{Vmatrix} 1 & 0 \dots 0 \\ f_{21} & 1 \dots 0 \\ \dots\dots & \dots \\ f_{n1}\, f_{n2} & \dots 1 \end{Vmatrix} \begin{Vmatrix} D_1 & & & & & \\ & \frac{D_2}{D_1} & & & & \\ & & \ddots & & & \\ & & & \frac{D_r}{D_{r-1}} & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{Vmatrix} \begin{Vmatrix} 1 & l_{12} \dots l_{1n} \\ 0 & 1 \dots l_{2n} \\ \dots\dots & \dots \\ 0 & 0 \dots 1 \end{Vmatrix}, \tag{2.48}$$

где

$$f_{gk} = \frac{A\begin{pmatrix} 1 & 2 \dots & k-1 & g \\ 1 & 2 \dots & k-1 & k \end{pmatrix}}{A\begin{pmatrix} 1 & 2 \dots & k \\ 1 & 2 \dots & k \end{pmatrix}}, \quad l_{kg} = \frac{A\begin{pmatrix} 1 & 2 \dots & k-1 & k \\ 1 & 2 \dots & k-1 & g \end{pmatrix}}{A\begin{pmatrix} 1 & 2 \dots & k \\ 1 & 2 \dots & k \end{pmatrix}}$$
$$(g = k+1, \dots, n;\ k = 1, 2, \dots, r), \tag{2.49}$$

аf_{qk}, l_{kq} *произвольные при* $g = k + 1, ..., n$; $k = r + 1, ..., n$.

3. Метод исключения Гаусса, будучи примененный к матрице $A = \| a_{ik} \|^n_1$ ранга r, для которой $D_k \neq 0$ $(k = 1, 2, .. .,r)$, дает нам две матрицы: нижнюю треугольную матрицу W с диагональными элементами 1 и верхнюю треугольную матрицу G, у которой первые r диагональных элементов равные $D_1, D_2/D_1, ..., D_r/D_{r-1}$, а последние $n — r$ строк заполнены нулями. G — гауссова форма матрицы A, W — преобразующая матрица.

Для конкретного вычисления элементов матрицы W можно рекомендовать следующий прием.

Мы получим матрицу W, если к единичной матрице E применим все те преобразования (задаваемые матрицами $W_1, ..., W_N$), которые мы в алгоритме Гаусса делали над матрицей A (в этом случае вместо произведения WA, равного G, мы будем иметь произведение WE, равное W). Поэтому к матрице A приписываем справа единичную матрицу E:

$$\left\|\begin{matrix} a_{11} \ldots a_{1n}\ 1 \ldots 0 \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ a_{n1} \ldots a_{nn}\ 0 \ldots 1 \end{matrix}\right\|. \tag{2.50}$$

Применяя к этой прямоугольной матрице все преобразования алгоритма Гаусса, получим прямоугольную матрицу, которая состоит из двух квадратных матриц G и W:

$$(G,W).$$

Таким образом, *применение алгоритма Гаусса к матрице* (2.50) *дает одновременно и матрицу G и матрицу W.*

Если A — неособенная матрица, т.е. $|A| \neq 0$, то и $|G| \neq 0$. В этом случае из (2.36) следует $A^{-1}=G^{-1}W$. Поскольку матрицы G и W определены при помощи алгоритма Гаусса, то нахождение обратной матрицы A^{-1} сводится к определению G^{-1} и умножению G^{-1} на W.

Хотя нахождение обратной матрицы G^{-1}, после того как определена матрица G, не представляет затруднений, поскольку G — треугольная матрица, тем не менее можно избежать этой операции. Для этого наряду с матрицами G и W введем аналогичные матрицы G_1 и W_1 для транспонированной матрицы A'. Тогда $A' = W_1^{-1}G_1$, т.е.

$$A = G'_1W'^{-1}_1 \tag{2.51}$$

Сопоставим между собой равенства (2.37) и (2.48):

$$A = W^{-1}G, \quad A = FDL.$$

Эти равенства могут быть рассмотрены как два различных разложения вида (2.39); при этом мы произведение DL рассматриваем как второй множитель C. Поскольку первые r диагональных элементов у первых множителей одинаковы (они равны 1), то первые r столбцов у них совпадают. Тогда, поскольку последние $n — r$ столбцов матрицы F могут быть выбраны произвольными, то выберем их так, чтобы

$$F= W^{-1} \tag{2.52}$$

С другой стороны, сопоставление равенств (2.51) и (2.48):

$$A = G'_1W'^{-1}_1, \quad A = FDL$$

показывает, что можно так подобрать произвольные элементы в L, чтобы

$$L = W'^{-1}_1. \tag{2.53}$$

Подставляя в (2.48) вместо F и L их выражения из (2.52) и (2.53), получим:

$$A = W^{-1}D\, W'^{-1}_1 \tag{2.54}$$

Сопоставляя это равенство с равенствами (2.37) и (2.51), мы найдем:

$$G = D\, W'^{-1}_1 \quad G' = W^{-1}D. \tag{2.55}$$

Введем в рассмотрение диагональную матрицу

$$\hat{D} = \left\{ \frac{1}{D_1},\ \frac{D_1}{D_2},\ \ldots,\ \frac{D_{r-1}}{D_r},\ 0, \ldots, 0 \right\}. \tag{2.56}$$

Тогда, поскольку

$$D = D\,\hat{D}\,D,$$

то из (2.54) и (2.55) следует:

$$A = G'_1\,\hat{D}\,G. \tag{2.57}$$

Формула (2.57) показывает, что разложение матрицы A на треугольные множители может быть получено применениям алгоритма Гаусса к матрицам A и A'.

Пусть теперь A — неособенная матрица ($r = n$). Тогда $| D | \neq 0$, $\hat{D} = D^{-1}$. Поэтому из (2.54) следует:

$$A^{-1} = W'_1\,\hat{D}\,W. \tag{2.58}$$

Эта формула дает возможность эффективного вычисления обратной матрицы A^{-1} путем применения алгоритма Гаусса к прямоугольным матрицам

$$(A, E)\ (A', E).$$

В частном случае, когда вместо матрицы A возьмем симметричную матрицу S, матрица G_1 совпадает с G, а матрица W_1 — с матрицей W, и потому формулы (2.57) и (2.58) принимают вид

$$S = G'\,\hat{D}\,G, \tag{2.59}$$

$$S^{-1} = W'\,\hat{D}\,W. \tag{2.60}$$

Уравнение *AX=B*. Многие задачи строительной механики, электротехники, радиоэлектроники и других областей техники связаны с решением системы линейных уравнений, правые части которых принимают различные значения, а матрица системы остается неизменной. Совокупность таких систем $Ax^{(1)} = b^{(1)}, \ldots, Ax^{(m)} = b^{(m)}$ можно рассматривать как одно матричное уравнение $AX=B$, где X и B — матрицы размера $n \times m$, столбцы которых равны соответственно $x^{(j)}$ и $b^{(j)}$ ($j = 1, 2, \ldots, m$).

Как уже указывалось, решение уравнения $AX=B$ можно представить через обратную матрицу $X = A^{-1}B$. Часто, однако, отдают предпочтение процедурам исключения, которые выполняются над расширенной матрицей $[A, B]$. Пусть, например, требуется решить уравнение $Ax=b$ при заданной матрице A и различных векторах в правой части, т.е.

$$A=\begin{bmatrix} 1 & 2 & -3 \\ 3 & -2 & 1 \\ -2 & 1 & 3 \end{bmatrix};\quad b^{(1)}=\begin{bmatrix} 7 \\ 1 \\ 5 \end{bmatrix};\quad b^{(2)}=\begin{bmatrix} 10 \\ 6 \\ -5 \end{bmatrix}.$$

Преобразуем расширенную матрицу [*A, B*] по алгоритму Гаусса—Жордана:

$$\left[\begin{array}{ccc|cc} 1 & 2 & -3 & 7 & 10 \\ 3 & -2 & 1 & 1 & 6 \\ -2 & 1 & 3 & 5 & -5 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 2 & -3 & 7 & 10 \\ 0 & -8 & 10 & -20 & -24 \\ 0 & 5 & -3 & 19 & 15 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|cc} 1 & 0 & -\frac{1}{2} & 2 & 4 \\ 0 & 1 & -\frac{5}{4} & \frac{5}{2} & 3 \\ 0 & 0 & \frac{13}{4} & \frac{13}{2} & 0 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & 5 & 3 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right].$$

Отсюда имеем:

$$X=\begin{bmatrix} 3 & 4 \\ 5 & 3 \\ 2 & 0 \end{bmatrix};\quad x^{(1)}=\begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix};\quad x^{(2)}=\begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}.$$

Неопределенная система. В общем случае, когда число уравнений *m* не равно числу неизвестных *n*, также применима процедура исключения, причем в процессе ее реализации выявляется и характер системы. Пусть, например, дана система уравнений:

$$\left.\begin{array}{r} 2x_1 - x_2 + x_3 + 4x_4 + x_5 = 3 \\ x_1 + x_2 + 2x_3 - x_4 + 5x_5 = 3 \\ 4x_2 - 3x + 6x_4 - 9x_5 = 4 \\ 3x_1 + 2x_2 + 3x_3 + 3x_4 + 6x_5 = 8 \end{array}\right\}.$$

Преобразуем расширенную матрицу системы с помощью алгоритма Гаусса-Жордана:

$$\left[\begin{array}{ccccc|c} 2 & -1 & 1 & 4 & 1 & 3 \\ 1 & 1 & 2 & -1 & 5 & 3 \\ 0 & 4 & -3 & 6 & -9 & 4 \\ 3 & 2 & 3 & 3 & 6 & 8 \end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{ccccc|c} 1 & -\frac{1}{2} & \frac{1}{2} & 2 & \frac{1}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & \frac{3}{2} & -3 & \frac{9}{2} & \frac{3}{2} \\ 0 & 4 & -3 & 6 & -9 & 4 \\ 0 & \frac{7}{2} & \frac{3}{2} & -3 & \frac{9}{2} & \frac{7}{2} \end{array}\right] \rightarrow$$

$$\rightarrow \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & -2 & 3 & 1 \\ 0 & 0 & -7 & 14 & -21 & 0 \\ 0 & 0 & -2 & 4 & -6 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 3 & -1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Последняя нулевая строка соответствует тождеству 0=0, что свидетельствует о зависимости исходных уравнений. Так как независимых уравнений три, то и ранги системы r= 3. Таким образом, получаем эквивалентную систему уравнений:

$$x_1 + 3x_4 - x_5 = 2;\ x_2 = 1;\ x_3 - 2x_4 + 3x_5 == 0,$$

решение которой: $x_1 = 2 - 3x_4 + x_5$; $x_2 = 1$; $x_3 = 2x_4 - 3x_5$, где x_4 и x_5 могут принимать произвольные значения. Рассмотренная система является совместной и неопределенной, так как имеет бесконечную множество решений.

Вообще, если ранг совместной системы меньше ее порядка ($r<n$), то совокупность ее r неизвестных, называемых *основными,* всегда можно выразить через $n - r$ других неизвестных, называемых *свободными.* Очевидно, основными будут те величины, которые исключаются из уравнений, т.е. они соответствуют столбцам, по которым проводится процедура исключения. Свобода в выборе совокупности основных неизвестных ограничивается только возможностью выбора опорного элемента в данном столбце. Так, в рассмотренном примере $x_2 = 1$, и поэтому x_2 не может быть свободной неизвестной. Действительно, пропуская второй столбец и продолжая процесс исключения по третьему столбцу, приходим к следующей ситуации:

$$\left[\begin{array}{ccccc|c} 1 & -\frac{1}{2} & \frac{1}{2} & 2 & \frac{1}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & \frac{3}{2} & -3 & \frac{9}{2} & \frac{3}{2} \\ 0 & 4 & -3 & 6 & -9 & 4 \\ 0 & \frac{7}{2} & \frac{3}{2} & -3 & \frac{9}{2} & \frac{7}{2} \end{array}\right] \quad \left[\begin{array}{ccccc|c} 1 & -1 & 0 & 3 & -1 & 1 \\ 0 & 1 & 1 & -2 & 3 & 1 \\ 0 & 7 & 0 & 0 & 0 & 7 \\ 0 & 2 & 0 & 0 & 0 & 2 \end{array}\right].$$

Продолжать исключение по четвертому столбцу невозможно, так как все элементы в строках, которые остались, равны нулю. Поэтому необходимо возвратиться ко второму столбцу. Отсюда видно, что сама процедура исключения корректирует выбор совокупности основных неизвестных.

При любом числе *m* уравнений ранг системы не может превышать число неизвестных *n* ($r \le n$). Если $m > n$, то не менее $(m - n)$ уравнений совместной системы зависимы и превращаются в тождества в процессе исключения. Совместная система *m* уравнений ранга *r* имеет $m - r$ зависимых уравнений. При $r = n$ она определена, а при $r < n$ неопределенная.

Теорема Кронекера-Капелли. Рассмотрим пример несовместной системы и выясним общий критерий совместности. Пусть, например, в приведенной выше системе правая часть последнего уравнения равна не 8, а 5. Тогда расширенная матрица преобразуется к виду:

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 3 & -1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right].$$

Последняя строка выражает противоречие $0 = -3$ (?), что и свидетельствует о несовместности исходной системы. Такая система не имеет решений. Для того, чтобы система была совместна, необходимо обращение к тождества зависимых уравнений. А это возможно лишь тогда, когда нулевым строкам преобразованной матрицы *A* соответствуют и нулевые элементы преобразованного вектора *b*. Другими словами, ранг матрицы *A* не должен изменяться, если к ней приписать столбец *b*.

Эти соображения объясняют известную теорему Кронекера-Капелли: необходимым и достаточным условием совместности системы линейных уравнений является равенство рангов матрицы системы *A* и ее расширенной матрицы [*A, b*]. В нашем примере ранг *A* равен трем, а ранг [*A, b*] — четырем, поэтому система несовместна.

Однородная система уравнений. Как мы знаем, система, все свободные члены которой равны нулю (b=0), называется *однородной.* Она всегда совместна, так как нулевой столбец b не влияет на ранг расширенной матрицы [A, 0]. Однородная система n уравнений из n неизвестными имеет тривиальный решение x= 0, которое и единственно, если ранг матрицы A равен ее порядку (r — n). При $r < n$ однородная система имеет бесконечное множество решений и сводится к неопределенной системе r уравнений с n неизвестными. Это, в частности, означает, что система n уравнений с n неизвестными имеет нетривиальные решения при условии, что матрица системы особенна, т.е. det$A$= 0.,

Пусть в системе уравнений из предыдущего примера правые части равны нулю. Преобразованная матрица этой системы имеет вид

$$\begin{bmatrix} 1 & 0 & 0 & 3 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

что соответствует решению: $x_1 = -3x_4+x_5$; $x_2 = 0$; $x_3=2x_4-3x_5$

В общем случае решение однородной системы ранга r с n неизвестными

$$x_k = \sum_{s=r+1}^{n} a'_{ks} x_s \ (k = 1, 2, ..., r)$$

где a'_{ks} — элементы матрицы, преобразованной по алгоритму Гаусса-Жордана (предполагается, что основным неизвестным соответствуют первые r столбцов).

Рассмотрим однородную систему n уравнений с n неизвестными Ax=0 или

$$\sum_{j=1}^{n} a_{ij}x_j = 0 \ (i = 1, 2, \ldots, n),$$

матрица A которой просто вырождена, т.е. ее ранг r=n — 1. Разлагая определитель этой матрицы по элементам какой-либо строки, запишем:

$$\sum_{j=1}^{n} a_{ij}\Delta_{ij} = \det A \ (i = 1, 2, \ldots, n).$$

Так как det A=0, то сравнивая записанные соотношения при любом i, находим

$$x_j = k\Delta_{ij} \quad (j=1,\ 2,\ \dots,\ n),$$

где k — произвольное число, не равное нулю.

Таким образом, неизвестные пропорциональны соответствующим алгебраическим дополнениям элементов какой-либо строки матрицы A, т.е. вектор решений можно представить в виде:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = k \begin{bmatrix} \Delta_{i1} \\ \Delta_{i2} \\ \dots \\ \Delta_{in} \end{bmatrix}.$$

Решим систему уравнений

$$\begin{bmatrix} -6 & 8 & -1 \\ -5 & 7 & -1 \\ -4 & 6 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Матрица этой системы просто вырождена, так как ее ранг равен двум, следовательно, достаточно вычислить алгебраические дополнения элементов какой-либо строки, например, первой:

$$\Delta_{11} = \begin{vmatrix} 7 & -1 \\ 6 & -1 \end{vmatrix} = -1; \quad \Delta_{12} = -\begin{vmatrix} -5 & -1 \\ -4 & -1 \end{vmatrix} = 1;$$

$$\Delta_{13} = \begin{vmatrix} -5 & 7 \\ -4 & 6 \end{vmatrix} = -2.$$

Отсюда запишем решение рассмотренной системы уравнений: $x_1 = -k;\ x_2 = k;\ x_3 = -2k$ или

$$x = k \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix},$$

причем это решение удовлетворяет данной системе при любом значении числа k.

2.5. Разбиение матрицы на блоки. Обобщенный алгоритм Гаусса

Часто приходится пользоваться матрицами, разбитыми на прямоугольные части - «клетки» или «блоки». Рассмотрения таких «блочных» матриц мы посвящаем этот раздел.

1. Пусть дана прямоугольная матрица

$$A=||a_{ik}|| \qquad (i= 1,2, ...,m;\ k = 1,2, ...,n). \tag{2.61}$$

С помощью горизонтальных и вертикальных линий рассечем матрицу A на прямоугольные блоки:

$$A = \begin{array}{cc} \begin{matrix} \overbrace{\quad}^{n_1} & \overbrace{\quad}^{n_2} & \dots & \overbrace{\quad}^{n_t} \end{matrix} & \\ \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1t} \\ A_{21} & A_{22} & \dots & A_{2t} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{st} \end{pmatrix} & \begin{matrix} \}\, m_1 \\ \}\, m_2 \\ \vdots \\ \}\, m_s \end{matrix} \end{array}. \tag{2.62}$$

О матрице (2.62) будем говорить, что она разбита на st блоков $A_{\alpha\beta}$ размером $m_\alpha \times n_\beta$ ($\alpha = 1, 2, ..., s;\ \beta = 1, 2, ..., t$) или что она представлена в виде блочной матрицы. Вместо (2.62) будем сокращенно писать:

$$A=(A_{\alpha\beta}) \qquad (\alpha = 1, 2, ..., s;\ \beta = 1, 2, ..., t). \tag{2.63}$$

В случае $s = t$ будем пользоваться и такой записью:

$$A=(A_{\alpha\beta})^s_1. \tag{2.64}$$

Действия над блочными матрицами выполняются по тем же формальным правилам, как и в случае, когда вместо блоков имеем числовые элементы. Пусть, например, дано две прямоугольные матрицы одинаковых размеров и с одинаковым разбиением на блоки:

$$A=(A_{\alpha\beta}),\ B = (B_{\alpha\beta}) \qquad (\alpha = 1, 2, ..., s;\ \beta = 1, 2, ..., t). \tag{2.65}$$

Легко усмотреть, что

$$A+B=(A_{\alpha\beta}+B_{\alpha\beta}) \qquad (\alpha = 1, 2, ..., s;\ \beta = 1, 2, ..., t). \tag{2.66}$$

Подробнее остановимся на умножении блочных матриц. Известно, что при умножении двух прямоугольных матриц A и B длина строк в первом сомножителе A должна совпадать с высотой столбцов во втором сомножителе B. Для возможности «блочного» умножения этих матриц мы дополнительно потребуем, чтобы при разбиении на блоки все горизонтальные размеры в первом сомножителе совпадали с соответствующими вертикальными размерами во второму:

$$A = \begin{array}{c} \begin{array}{ccc} n_1 & n_2 \; \ldots & n_t \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1t} \\ A_{21} & A_{22} & \ldots & A_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ A_{s1} & A_{s2} & \ldots & A_{st} \end{pmatrix} \end{array} \begin{array}{l} \} \, m_1 \\ \} \, m_2 \\ \vdots \\ \} \, m_s \end{array}, \quad B = \begin{array}{c} \begin{array}{ccc} p_1 & p_2 \; \ldots & p_u \end{array} \\ \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1u} \\ B_{21} & B_{22} & \ldots & B_{2u} \\ \cdots & \cdots & \cdots & \cdots \\ B_{t1} & B_{t2} & \ldots & B_{tu} \end{pmatrix} \end{array} \begin{array}{l} \} \, n_1 \\ \} \, n_2 \\ \vdots \\ \} \, n_t \end{array}.$$

(2.67)

Тогда легко проверить, что

$$AB = C = (C_{\alpha\beta}), \quad \text{где} \quad C_{\alpha\beta} = \sum_{\delta=1}^{t} A_{\alpha\delta} B_{\delta\beta} \qquad \begin{pmatrix} \alpha = 1, 2, \ldots, s \\ \beta = 1, 2, \ldots, u \end{pmatrix}.$$

(2.68)

Отдельно отметим тот частный случай, когда одним из сомножителей является *квазидиагональная* матрица. Пусть A — квазидиагональная матрица, т.е. $s = t$ и $A_{\alpha\beta} = 0$ при $\alpha \neq \beta$. В этом случае формула (2.68) нам дает:

$$C_{\alpha\beta} = A_{\alpha\alpha} B_{\alpha\beta} \qquad (\alpha = 1, 2, ..., s; \ \beta = 1, 2, ..., u). \quad (2.69)$$

При умножении блочной матрицы слева на квазидиагональную матрицу все строки блочной матрицы умножаются слева на соответствующие диагональные клетки квазидиагональной матрицы.

Пусть теперь B — квазидиагональная матрица, т.е. $t = u$ и $B_{\alpha\beta}=0$ при $\alpha \neq \beta$. Тогда из (2.68) получаем:

$$C_{\alpha\beta} = A_{\alpha\beta} B_{\beta\beta} \qquad (\alpha = 1, 2, ..., s; \ \beta = 1, 2, ..., u). \quad (2.70)$$

При умножении блочной матрицы справа на квазидиагональную все столбцы блочной матрицы умножаются справа на соответствующие диагональные клетки квазидиагональной матрицы.

Заметим, что умножение квадратных блочных матриц одного и того же порядка всегда осуществимо, когда сомножители разбиты на одинаковые квадратные схемы блоков и в каждом с сомножителей на диагональных местах стоят квадратные матрицы.

Блочная матрица (2.62) называется *верхней (нижней) квазитреугольной,* если $s = t$ и все $A_{\alpha\beta}=0$ при $\alpha>\beta$ (соответственно все $A_{\alpha\beta}=0$ при $\alpha< \beta$). Частным случаем квазитреугольной матрицы является квазидиагональная матрица.

Из формулы (2.68) легко видеть, что

Произведение двух верхних (нижних) квазитреугольных матриц является снова верхней (нижней) квазитреугольной матрицей (при этом предполагается, что блочное умножение можно выполнить); при этом диагональные блоки произведения получаются путем *перемножения соответствующих диагональных блоков сомножителей.*

Действительно, полагая в (2.68) $s = t$ и

$$A_{\alpha\beta}=0, \quad B_{\alpha\beta}=0 \text{ при } \alpha<\beta,$$

найдем:

$$\left.\begin{aligned} &C_{\alpha\beta}=0 \ \text{ при } \ \alpha<\beta \\ &C_{\alpha\alpha}=A_{\alpha\alpha}B_{\alpha\alpha} \end{aligned}\right\} \quad (\alpha, \beta = 1, 2, \ldots, s).$$

Аналогично разбирается случай нижних квазитреугольных матриц.

Отметим правило вычисления определителя квазитреугольной матрицы. Это правило можно получить, исходя из разложения Лапласа.

Если A - квазитреугольная (в частности, квазидиагональная) матрица с квадратными диагональными блоками, то определитель этой матрицы равен произведению определителей диагональных блоков:

$$|A| = |A_{11}|\,|A_{22}| \ldots |A_{ss}|. \qquad (2.71)$$

2. Пусть дана блочная матрица

$$A = \begin{array}{c} \begin{array}{cccc} \overbrace{\phantom{A_{11}}}^{n_1} & \overbrace{\phantom{A_{12}}}^{n_2} & \cdots & \overbrace{\phantom{A_{1t}}}^{n_t} \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1t} \\ A_{21} & A_{22} & \ldots & A_{2t} \\ \ldots & \ldots & \ldots & \ldots \\ A_{s1} & A_{s2} & \ldots & A_{st} \end{pmatrix} \end{array} \begin{array}{l} \\ \}\, m_1 \\ \}\, m_2 \\ \vdots \\ \}\, m_s \end{array}.$$

(2.72)

Прибавим к α-й блочной строке β-ю строку, предварительно помноженную слева на прямоугольную матрицу X размером $m_\alpha \times m_\beta$. Получим блочную матрицу

$$B = \begin{pmatrix} A_{11} & \ldots & A_{1t} \\ \ldots & \ldots & \ldots \\ A_{\alpha 1} + XA_{\beta 1} & \ldots & A_{\alpha t} + XA_{\beta t} \\ \ldots & \ldots & \ldots \\ A_{\beta 1} & \ldots & A_{\beta t} \\ \ldots & \ldots & \ldots \\ A_{s1} & \ldots & A_{st} \end{pmatrix}.$$

(2.73)

Введем вспомогательную квадратную матрицу V, представленную в виде следующей квадратной схемы блоков:

$$V = \begin{array}{c l} \begin{array}{ccccccc} \overbrace{\;}^{m_1} & \dots & \overbrace{\;}^{m_\alpha} & \dots & \overbrace{\;}^{m_\beta} & \dots & \overbrace{\;}^{m_s} \end{array} & \\ \begin{pmatrix} E & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & E & \dots & X & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & E & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & E \end{pmatrix} & \begin{array}{l} \}\, m_1 \\ \vdots \\ \}\, m_\alpha \\ \vdots \\ \}\, m_\beta \\ \vdots \\ \}\, m_s \end{array} \end{array} . \qquad (2.74)$$

В диагональных клетках матрицы V стоят единичные матрицы, порядки которых равны соответственно $m_1, m_2, \ldots, m_s$; все недиагональные блоки матрицы V равны нулю, за исключением блока X, стоящего на пересечении α-й блочной строки с β-м блочным столбцом.

Нетрудно видеть, что

$$VA = B. \qquad (2.75)$$

Отсюда, поскольку V — неособенная матрица, для рангов матриц A и B имеем:

$$r_A = r_B. \qquad (2.76)$$

В частном случае, когда A — квадратная матрица, из (2.75) имеем:

$$|V|\,|A| = |B|. \qquad (2.77)$$

Но определитель квазитреугольной матрицы V равен 1:

$$|V| = 1. \qquad (2.78)$$

Следовательно,

$$|A| = |B|. \qquad (2.79)$$

К таким же выводам можно прийти, если к какому-либо столбцу матрицы (2.71) прибавлять другой столбец, предварительно помноженный справа на прямоугольную матрицу X надлежащих размеров.

Полученные результаты могут быть сформулированы в виде следующей теоремы.

Теорема 3. *Если в блочной матрице A к α-й блочной строке (столбца) прибавить β-ю блочную строку (столбец), предварительно помноженную слева (справа) на прямоугольную матрицу X соответствующих размеров, то при этом преобразовании не* *изменятся ранг матрицы А, а также в случае, когда А - квадратная матрица, и определитель матрицы А.*

3. Рассмотрим теперь тот частный случай, когда в матрице A диагональный блок A_{11} — квадратная и притом неособенная матрица ($|A_{11}| \neq 0$).

К α-й строке матрицы A прибавим первую строку, помноженную слева на $-A_{\alpha 1}A^{-1}_{11}$ ($\alpha = 2, ..., s$). Тогда получим матрицу

$$B_1 = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1t} \\ 0 & A^{(1)}_{22} & \dots & A^{(1)}_{2t} \\ \dots & \dots & \dots & \dots \\ 0 & A^{(1)}_{s2} & \dots & A^{(1)}_{st} \end{pmatrix}, \tag{2.80}$$

где

$$A^{(1)}_{\alpha\beta} = -A_{\alpha 1}A^{-1}_{11}A_{1\beta} + A_{\alpha\beta} \quad (\alpha=2, \dots, s;\ \beta= 2, \dots, t) \tag{2.81}$$

Если $A^{(1)}_{22}$ — квадратная неособенная матрица, то этот процесс можно продолжить. Мы приходим, таким образом, к обобщенному алгоритму Гаусса,

Пусть A — квадратная матрица. Тогда

$$|A| = |B_1| = |A_{11}| \begin{vmatrix} A^{(1)}_{22} & \dots & A^{(1)}_{2t} \\ \dots & \dots & \dots \\ A^{(1)}_{s2} & \dots & A^{(1)}_{st} \end{vmatrix}. \tag{2.82}$$

Формула (2.82) сводит вычисление определителя $|A|$, состоящего из st блоков, к вычислению определителя меньшего порядка, который состоит из $(s - 1)(t - 1)$ блоков. (Если $A^{(1)}_{22}$ — квадратная матрица и $|A^{(1)}_{22}| \neq 0$, то к полученному определителю из $(s - 1)(t - 1)$ блоков мы можем снова применить такое же преобразование и т.д.).

Рассмотрим определитель Δ, разбитый на четыре блока:

$$\Delta = \begin{vmatrix} A & B \\ C & D \end{vmatrix}, \tag{2.83}$$

где A и D — квадратные матрицы.

Пусть $|A| \neq 0$. Тогда вычтем из второй строки первую, предварительно умноженную слева на $-CA^{-1}$. Получим:

$$\Delta = \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix} = |A|\,|D - CA^{-1}B|. \tag{2.84}$$

Точно так же, если | D | ≠ 0, то мы вычтем в Δ из первой строки вторую, предварительно помноженную слева на —BD^{-1}. Получим:

$$\Delta = \begin{vmatrix} A - BD^{-1}C & 0 \\ C & D \end{vmatrix} = |A - BD^{-1}C|\,|D|.$$

(2.85)

В частном случае, когда все четыре матрицы *A, B, C, D* — квадратные (одного и того же порядка *n),* из (2.84) и (2.85) следуют формулы Шура, которые сводят вычисление определителя 2n-го порядка к вычислению определителя n-го порядка:

$$\Delta=|AD- ACCA^{-1}B| \quad (|A| \neq 0), \tag{2.86}$$

$$\Delta = |AD - BD^{-1}CD| \quad (|D| \neq 0). \tag{2.87}$$

Если матрицы A и C перестановочны между собой, то из (2.86) следует:

$$\Delta = |AD - CB| \quad (\text{при условии } ACC = CA). \tag{2.88}$$

Точно так же, если C и D перестановочны между собой, то

$$\Delta = |AD - BC| \quad (\text{при условии } CD = DC). \tag{2.89}$$

Формула (2.88) была получена в предположении $|A| \neq 0$, а формула (2.89) при условии $|D| \neq 0$. Однако, исходя из соображений непрерывности, эти ограничения можно отбросить.

Из формул (2.84)—(2.89) можно получить еще шесть формул, поменяв местами в правых частях A и D и одновременно B и C.

Пример.

$$\Delta = \begin{vmatrix} 1 & 0 & b_1 & b_2 \\ 0 & 1 & b_3 & b_4 \\ c_1 & c_2 & d_1 & d_2 \\ c_3 & c_4 & d_3 & d_4 \end{vmatrix}.$$

По формуле (2.88)

$$\Delta = \begin{vmatrix} d_1 - c_1b_1 - c_2b_3 & d_2 - c_1b_2 - c_2b_4 \\ d_3 - c_3b_1 - c_4b_3 & d_4 - c_3b_2 - c_4b_4 \end{vmatrix}.$$

4. Установим формулу Фробениуса для обращения блочной матрицы. Пусть неособенная квадратная матрица M (| M | ≠ 0) разбита на блоки

$$M = \begin{array}{c} \\ n\{ \\ q\{ \end{array} \begin{array}{c} \begin{matrix} n & q \end{matrix} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{array},$$

(2.90)

и пусть A — также неособенная квадратная матрица (| A | ≠ 0). Требуется определить M^{-1}.

Применим к матрице M обобщенный алгоритм Гаусса. Из второй блочной строки вычтем первую, предварительно помноженную слева на — CA^{-1}. Эта операция равносильна умножению матрицы M слева на матрицу

$$\begin{pmatrix} E & 0 \\ X & E \end{pmatrix},$$

(в первой блочной строке буква E обозначает единичную матрицу n-го порядка, а во второй - единичную матрицу порядка q) где

$$X = -CA^{-1}.$$

Поэтому

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix} M = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}. \tag{2.91}$$

Введем обозначение

$$H=D-CA^{-1}B$$

и заметим, что из равенства (2.91) следует:

$$|M| = |A||H|. \tag{2.92}$$

Поэтому, поскольку $|M|\neq 0$, то и $|H| \neq 0$. (Нам нет необходимости заранее устанавливать неособенность матрицы M, так как это свойство матрицы M следует из того, что $| A |\neq 0$ и $| H | \neq 0$. Если бы оказалось, что $|H|=0$, то тогда и $|M| = 0$, и в этом случае не существует обратной матрицы M^{-1}).

Переходя к обратным матрицам в равенстве (2.91), получим

$$M^{-1} \begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}^{-1} = \begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1}. \tag{2.93}$$

Обратную матрицу для матрицы $\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}$ будем искать в виде

$$\begin{pmatrix} A^{-1} & U \\ 0 & H^{-1} \end{pmatrix}.$$

Тогда из равенства

$$\begin{pmatrix} A & B \\ 0 & H \end{pmatrix} \begin{pmatrix} A^{-1} & U \\ 0 & H^{-1} \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}$$

находим, что $U = -A^{-1}BH^{-1}$. Таким образом,

$$\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BH^{-1} \\ 0 & H^{-1} \end{pmatrix}. \tag{2.94}$$

Но тогда из равенства (2.93) находим

$$M^{-1} = \begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1} \begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix} = \begin{pmatrix} A^{-1} & -A^{-1}BH^{-1} \\ 0 & H^{-1} \end{pmatrix} \begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}. \tag{2.95}$$

Выполняя умножение блочных матриц в правой части равенства (2.95), приходим к формуле Фробениуса

$$M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BH^{-1}CA^{-1} & -A^{-1}BH^{-1} \\ -H^{-1}CA^{-1} & H^{-1} \end{pmatrix}, \tag{2.96}$$

где

$$H = D - CA^{-1}B. \tag{2.97}$$

Формула Фробениуса (2.96) сводит обращение матрицы порядка $n+q$ к обращению двух матриц порядка n и q и к операциям сложения и умножения матриц с размерами их $n\times n$, $q\times q$, $n\times q$ и $q\times n$.

Если предположить, что $|D| \neq 0$ (вместо $|A| \neq 0$) и поменять ролями матрицы A и D, то можно получить другой вид формулы Фробениуса:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} K^{-1} & -K^{-1}BD^{-1} \\ -D^{-1}CK^{-1} & D^{-1} + D^{-1}CK^{-1}BD^{-1} \end{pmatrix}, \tag{2.98}$$

где

$$K = A - BD^{-1}C. \tag{2.99}$$

Пример. Требуется найти элементы обратной матрицы для матрицы

$$M = \begin{Vmatrix} 1 & -1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & 1 & 2 \end{Vmatrix}.$$

Полагаем

$$A = \begin{Vmatrix} 1 & -1 \\ -1 & 2 \end{Vmatrix}, \quad B = \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}, \quad C = \begin{Vmatrix} 0 & 2 \\ 2 & 0 \end{Vmatrix}, \quad D = \begin{Vmatrix} 0 & 1 \\ 1 & 2 \end{Vmatrix}.$$

Находим последовательно

$$A^{-1}=\begin{Vmatrix}2&1\\1&1\end{Vmatrix},\quad CA^{-1}=\begin{Vmatrix}0&2\\2&0\end{Vmatrix}\begin{Vmatrix}2&1\\1&1\end{Vmatrix}=\begin{Vmatrix}2&2\\4&2\end{Vmatrix},$$

$$H=D-CA^{-1}B=D-\begin{Vmatrix}2&2\\4&2\end{Vmatrix}\begin{Vmatrix}0&1\\-1&0\end{Vmatrix}=\begin{Vmatrix}0&1\\1&2\end{Vmatrix}-\begin{Vmatrix}-2&2\\-2&4\end{Vmatrix}=\begin{Vmatrix}2&-1\\3&-2\end{Vmatrix},$$

$$H^{-1}=\begin{Vmatrix}2&-1\\3&-2\end{Vmatrix},\quad A^{-1}B=\begin{Vmatrix}2&1\\1&1\end{Vmatrix}\begin{Vmatrix}0&1\\-1&0\end{Vmatrix}=\begin{Vmatrix}-1&2\\-1&1\end{Vmatrix},$$

$$A^{-1}BH^{-1}=\begin{Vmatrix}-1&2\\-1&1\end{Vmatrix}\begin{Vmatrix}2&-1\\3&-2\end{Vmatrix}=\begin{Vmatrix}4&-3\\1&-1\end{Vmatrix},$$

$$A^{-1}BH^{-1}CA^{-1}=\begin{Vmatrix}4&-3\\1&-1\end{Vmatrix}\begin{Vmatrix}2&2\\4&2\end{Vmatrix}=\begin{Vmatrix}-4&2\\-2&0\end{Vmatrix},$$

$$A^{-1}+A^{-1}BH^{-1}CA^{-1}=\begin{Vmatrix}-2&3\\-1&1\end{Vmatrix},$$

$$H^{-1}CA^{-1}=\begin{Vmatrix}2&-1\\3&-2\end{Vmatrix}\begin{Vmatrix}2&2\\4&2\end{Vmatrix}=\begin{Vmatrix}0&2\\-2&2\end{Vmatrix}.$$

Поэтому по формуле (2.96) находим:

$$M^{-1}=\begin{Vmatrix}-2&3&-4&3\\-1&1&-1&1\\0&-2&2&-1\\2&-2&3&-2\end{Vmatrix}.$$

5. Из теоремы 3 вытекает также

Теорема 4. *Если прямоугольная матрица R представлена в блочном виде*

$$R=\begin{pmatrix}A&B\\C&D\end{pmatrix},\tag{2.100}$$

где A — квадратная неособенная матрица порядка n ($|A|\neq 0$) *, то ранг матрицы R равен n в том и только в том случае, когда*

$$D=CA^{-1}B.\tag{2.101}$$

Доказательство. Вычтем из второй блочной строки матрицы R первую, предварительно помноженную слева на CA^{-1}. Тогда получим матрицу

$$T=\begin{pmatrix}A&B\\0&D-CA^{-1}B\end{pmatrix}.\tag{2.102}$$

Матрицы R и T согласно теореме 3 имеют один и тот же ранг. Ранг же матрицы T совпадает с рангом матрицы A (т.е. с n) тогда и только тогда, когда $D — CA^{-1}B = 0$, т.е. когда имеет место (2.101). Теорема доказана.

Из теоремы 4 вытекает алгоритм построения обратной матрицы A^{-1} и вообще произведения $CA^{-1}B$, где B, C — прямоугольные матрицы размером $n \times p$, $q \times n$.

Приведем с помощью алгоритма Гаусса матрицу

$$\begin{pmatrix} A & B \\ -C & 0 \end{pmatrix} \quad (|A| \neq 0)$$

(2.103)

к виду

$$\begin{pmatrix} G & B_1 \\ 0 & X \end{pmatrix}.$$

(2.104)

(Здесь мы применяем к матрице (2.103) не весь алгоритм Гаусса, а только его первые n этапов, где n — порядок матрицы A. Это можно сделать, если выполняются условия (2.17) при $p = n$. Если же эти условия не выполненные, то, поскольку $|A| \neq 0$, мы можем так перенумеровать первые n строк (или первые n столбцов) матрицы (2.103), чтобы n этапов алгоритма Гаусса оказались выполнимыми. Такой видоизмененный алгоритм Гаусса иногда применяют и при выполнении условий (2.17) для $p= n$).

Докажем, что

$$X = CA^{-1}B. \qquad (2.105)$$

В самом деле, то же преобразование, которое было применено к матрице (2.103), приведет матрицу

$$\begin{pmatrix} A & B \\ -C & -CA^{-1}B \end{pmatrix}$$

(2.106)

к виду

$$\begin{pmatrix} G & B_1 \\ 0 & X - CA^{-1}B \end{pmatrix}.$$

(2.107)

Согласно теореме 4 матрица (2.106) имеет ранг n (n — порядок матрицы A). Но тогда и матрица (2.107) должна иметь ранг n. Отсюда $X — CA^{-1} = 0$, т.е. имеет место (2.105).

В частности, если $B = y$, где y — стовбцевая матрица, и $C = E$, то

$$X = A^{-1}y.$$

Следовательно, применяя алгоритм Гаусса к матрице

$$\begin{pmatrix} A & y \\ -E & 0 \end{pmatrix},$$

мы получаем решение системы уравнений

$$Ax = y.$$

Далее, если в (2.103) положить $B=C=E$, то после применения алгоритма Гаусса к матрице

$$\begin{pmatrix} A & E \\ -E & 0 \end{pmatrix}$$

получим:

$$\begin{pmatrix} G & W \\ 0 & X \end{pmatrix},$$

где

$$X=A^{-1}$$

Проиллюстрируем этот способ нахождения A^{-1} на следующем примере.

Пример. Пусть

$$A = \begin{Vmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 2 \end{Vmatrix}.$$

Требуется вычислить A^{-1}.

Применяем немного видоизмененный метод исключения к матрице

$$\begin{Vmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{Vmatrix}.$$

Ко всем строкам прибавляем вторую строку с некоторым множителем и добиваемся того, чтобы все элементы первого столбца, кроме второго элемента, равнялись нулю. После этого ко всем строкам, кроме второй, прибавляем третью строку с некоторым множителем и достигаем того, чтобы во втором столбце все элементы, кроме второго и третьего, были равны нулю. После этого к последним трем строкам прибавляем первую строку с некоторым множителем и получаем матрицу вида

$$\begin{Vmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ 0 & 0 & 0 & -2 & -1 & 2 \\ 0 & 0 & 0 & 4 & 1 & -3 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{Vmatrix}.$$

Поэтому

$$A^{-1} = \begin{Vmatrix} -2 & -1 & 2 \\ 4 & 1 & -3 \\ 1 & 1 & -1 \end{Vmatrix}.$$

6. Разбиение матрицы на блоки может быть использовано также для нахождения псевдообратной матрицы .

Пусть снова прямоугольная матрица R представлена в виде

$$R = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{2.108}$$

где A — неособенная квадратная матрица ($|A| \neq 0$) и $r_A = r_R$. Тогда согласно теореме 4 справедливо равенство $D = CA^{-1}B$, и потому

$$R = \begin{pmatrix} A \\ C \end{pmatrix} A^{-1}(AB). \tag{2.109}$$

Так как эта формула является результатом двух последовательных скелетных разложений :

$$R = \begin{pmatrix} A \\ C \end{pmatrix} (E \ \ A^{-1}B),\ (E \ \ A^{-1}B) = A^{-1}(AB),$$

то

$$R^{+} = (AB)^{+}(A^{-1})^{+} \begin{pmatrix} A \\ C \end{pmatrix}^{+} = (AB)^{+} A \begin{pmatrix} A \\ C \end{pmatrix}^{+}. \tag{2.110}$$

Применяя формулы (1.45) и (1.46), окончательно найдем:

$$R^{+} = \begin{pmatrix} A^{*} \\ B^{*} \end{pmatrix} (AA^{*} + BB^{*})^{-1} A\, (A^{*}A + C^{*}C)^{-1} (A^{*}C^{*}). \tag{2.111}$$

Формула (2.111) дает явное выражение для псевдообратной матрицы R^{+} через блоки A, B, C.

Пример.

$$R=\left\|\begin{array}{rr|rr} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ \hline 0 & 1 & -1 & 1 \end{array}\right\| = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Здесь $r_R = 2$ и

$$A=\left\|\begin{matrix} 1 & -1 \\ -1 & 2 \end{matrix}\right\|, \quad B=\left\|\begin{matrix} 2 & 0 \\ -3 & 1 \end{matrix}\right\|, \quad C=\|\,0\;\;1\,\|, \quad D=\|\,-1\;\;1\,\|,$$

$$(AA^*+BB^*)^{-1}=\left(3\left\|\begin{matrix} 2 & -3 \\ -3 & 5 \end{matrix}\right\|\right)^{-1}=\frac{1}{3}\left\|\begin{matrix} 5 & 3 \\ 3 & 2 \end{matrix}\right\|,$$

$$(A^*A+C^*C)^{-1}=\left\|\begin{matrix} 2 & -3 \\ -3 & 6 \end{matrix}\right\|^{-1}=\frac{1}{3}\left\|\begin{matrix} 6 & 3 \\ 3 & 2 \end{matrix}\right\|,$$

$$(AA^*+BB^*)^{-1}A(A^*A+C^*C)^{-1}=\frac{1}{9}\left\|\begin{matrix} 15 & 8 \\ 9 & 5 \end{matrix}\right\|.$$

Тогда по формуле (2.111)

$$R^+=\frac{1}{9}\left\|\begin{matrix} 1 & -1 \\ -1 & 2 \\ 2 & -3 \\ 0 & 1 \end{matrix}\right\| \left\|\begin{matrix} 15 & 8 \\ 9 & 5 \end{matrix}\right\| \left\|\begin{matrix} 1 & -1 & 0 \\ -1 & 2 & 1 \end{matrix}\right\| = \left\|\begin{matrix} \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{2}{9} & -\frac{1}{9} & \frac{1}{9} \\ \frac{4}{9} & \frac{1}{9} & \frac{5}{9} \end{matrix}\right\|.$$

Применение блочных матриц. При решении определенной системы уравнений n-го порядка можно воспользоваться разбиением матрицы этой системы на блоки. Представим уравнение $Ax=b$ с неособенной квадратной матрицей A в виде:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} x' \\ x'' \end{bmatrix}=\begin{bmatrix} b' \\ b'' \end{bmatrix},$$

где A_{11} и A_{22}— квадратные матрицы порядков p и q, причем $p+q=n$. Рассматирваемое уравнение равносильно системе двух матричных уравнений:

$$\left.\begin{matrix} A_{11}x'+A_{12}x''=b' \\ A_{21}x'+A_{22}x''=b'' \end{matrix}\right\}.$$

Выразим x'' из второго уравнения

$$x'' = A_{22}^{-1}(b'' - A_{21}x')$$

и подставим его в первое уравнение

$$A_{11}x' + A_{12}A_{22}^{-1}(b'' - A_{21}x') = b'.$$

Отсюда имеем

$$x' = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}(b' - A_{12}A_{22}^{-1}b'').$$

Определив по этой формуле x', можно потом найти и x'' по приведенному выше соотношению.
Проиллюстрируем изложенный метод на примере системы уравнений:

$$\left.\begin{aligned} 2x_1 - x_2 - x_3 - 4x_4 &= 4 \\ 3x_2 - x_3 - 2x_4 &= 2 \\ x_1 + x_2 + 2x_3 + 3x_4 &= 6 \\ 3x_1 - 2x_2 + x_3 + x_4 &= 6 \end{aligned}\right\}.$$

Представим эту систему в виде

$$\left[\begin{array}{cc|cc} 2 & -1 & -1 & -4 \\ 0 & 3 & -1 & -2 \\ \hline 1 & 1 & 2 & 3 \\ 3 & -2 & 1 & 1 \end{array}\right]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \left[\begin{array}{c} 4 \\ 2 \\ \hline 6 \\ 6 \end{array}\right].$$

По формуле для x' имеем

$$x' = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left(\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} -1 & -4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix}\right)^{-1} \times$$

$$\times\left(\begin{bmatrix} 4 \\ 2 \end{bmatrix} - \begin{bmatrix} -1 & -4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 6 \\ 6 \end{bmatrix}\right) = \frac{1}{18}\begin{bmatrix} -3 & 6 \\ -1 & 5 \end{bmatrix}\begin{bmatrix} -8 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Используя полученный результат, определяем x''

$$x'' = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}^{-1}\left(\begin{bmatrix} 6 \\ 6 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

Таким образом, решение системы уравнений: $x_1 = 2$; $x_2 = 1$; $x_3 = 3$; $x_4 = -1$ или в векторной форме

$$x = \begin{bmatrix} 2 \\ 1 \\ 3 \\ -1 \end{bmatrix}.$$

Особенность изложенного метода решения системы уравнений заключается в том, что вместо обращения матрицы n-го порядка необходимо обращать матрицы более низких порядков p и q
$(p+q= — n)$. Этот метод особенно удобен, если матрицу системы уравнений можно представить в таком виде, что одна из матриц A_{12} (или A_{21}) является нулевой (это обычно достигается перестановкой соответствующих строк и столбцов). Так, при $A_{12}=0$

$$x' = A_{11}^{-1}b'; \quad x'' = A_{22}^{-1}(b'' - A_{21}x'),$$

а при $A_{21} = 0$

$$x' = A_{11}^{-1}(b' - A_{12}A_{22}^{-1}b''); \quad x'' = A_{22}^{-1}b''.$$

Если одновременно $A_{12}=0$ и $A_{21}=0$, то исходная система распадается на две независимые системы уравнений $A_{11}x' = b'$ и $A_{22}x'' = b''$, решения которых: $x' = A^{-1}{}_{11}b'$ и $x'' = A^{-1}{}_{22}b''$.

Исключение переменных. Иногда требуется представить определенную систему уравнений n-го порядка относительно переменных $x_1, x_2, ..., x_n$, исключив из нее совокупность переменных $x_{s+1}, ..., x$. Обозначив $x' = (x_1, x_2, ... x_s)$ и $x'' = (x_{s+1} ..., x_n)$, на основе результатов предыдущего пункта запишем:

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})\,x' = b' - A_{12}A_{22}^{-1}b''.$$

Таким образом, приходим к системе уравнений s-го порядка в матричной форме $A'x'=c$, где матрица системы A' и вектор свободных членов c выражаются соотношениями:

$$A' = A_{11} - A_{12}A_{22}^{-1}A_{21}; \quad c = b' - A_{12}A_{22}^{-1}b''.$$

В частности, если $b''=0$, получаем сокращенную систему в виде $A'x' = b'$. В этом случае можно представить элементы a'_{ij} $(i, j =1, 2, ..., s)$ матрицы A' через миноры матрицы A в виде:

$$a'_{ij} = \frac{\Delta_s^{ij}}{\Delta_s} \quad (i, \ j = 1, \ 2, \ \ldots, \ s).$$

Здесь $\Delta_s = \Delta_{11,\ 22,\ldots,ss}$ - s-кратное алгебраическое дополнение матрицы A, получаемое вычеркиваниям из ее определителя первых s строк и столбцов. Его можно рассматривать также как минор матрицы A, образованный последними $(n — s)$ строками и столбцами с номерами
$s+1$, $s+2$, ..., n (рис. 2.6, *а)*. Величина Δ_s^{ij} — это минор матрицы A $(n —s+1)$-го порядка, образованный i-й строкой и j-м столбцом, а также последними $(n — s)$ строками и столбцами (рис. 2.6, б).

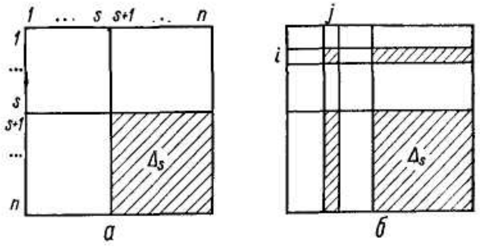


Рис. 2.6. Схема образования миноров матрицы системы:
а — минор Δ_s; *б* — минор Δ_s^{ij}.

Рассмотрим, например, систему уравнений

$$\left[\begin{array}{cc|ccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \\ \hline x_3 \\ x_4 \\ x_5 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \\ \hline 0 \\ 0 \\ 0 \end{array}\right].$$

Исключая переменные x_3, x_4, x_5, приходим к системе

$$\begin{bmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

где

$$a'_{11} = \frac{1}{\Delta_s} \begin{vmatrix} a_{11} & a_{13} & a_{14} & a_{15} \\ a_{31} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{53} & a_{54} & a_{55} \end{vmatrix}; \quad a'_{12} = \frac{1}{\Delta_s} \begin{vmatrix} a_{12} & a_{13} & a_{14} & a_{15} \\ a_{32} & a_{33} & a_{34} & a_{35} \\ a_{42} & a_{43} & a_{44} & a_{45} \\ a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix};$$

$$a'_{21} = \frac{1}{\Delta_s} \begin{vmatrix} a_{21} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{53} & a_{54} & a_{55} \end{vmatrix}; \quad a'_{22} = \frac{1}{\Delta_s} \begin{vmatrix} a_{22} & a_{23} & a_{24} & a_{25} \\ a_{32} & a_{33} & a_{34} & a_{35} \\ a_{42} & a_{43} & a_{44} & a_{45} \\ a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix},$$

а также

$$\Delta_s = \begin{vmatrix} a_{33} & a_{34} & a_{35} \\ a_{43} & a_{44} & a_{45} \\ a_{53} & a_{54} & a_{55} \end{vmatrix}.$$

Микромодуль 5
Примеры решения типовых задач

1. Решить систему уравнений матричным методом и по формулам Крамера.

$$\begin{cases} 2x_1 + x_2 - 3x_3 = 3, \\ x_1 + x_2 + 2x_3 = 2, \\ -x_1 + x_2 + x_3 = 0. \end{cases}$$

Решение. Матричный метод

$$A = \begin{pmatrix} 2 & 1 & -3 \\ 1 & 1 & 2 \\ -1 & 1 & 1 \end{pmatrix}$$ - основная матрица системы.

Используя свойства определителей, вычисляем определитель основной матрицы:

$$\Delta(A) = \begin{vmatrix} 2 & 1 & -3 \\ 1 & 1 & 2 \\ -1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 3 & -1 \\ 1 & 2 & 3 \\ -1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 3 & -1 \\ 0 & 2 & 3 \\ -1 & 0 & 0 \end{vmatrix} = 0 - 9 + 0 - (2 + 0 + 0) = -11 \neq 0$$

Поскольку $\Delta(A) \neq 0$, то обратная матрица существует. Находим ее.

$$A_{11} = \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1, \quad A_{12} = \begin{vmatrix} 1 & -3 \\ 1 & 1 \end{vmatrix} = -4, \quad A_{13} = \begin{vmatrix} 1 & -3 \\ 1 & 2 \end{vmatrix} = 5,$$

$$A_{12} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = -3, \quad A_{22} = \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1, \quad A_{32} = \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -7,$$

$$A_{13} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2, \quad A_{23} = \begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} = -3, \quad A_{33} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1,$$

Обратная матрица имеет вид $A^{-1} = \dfrac{-1}{11} \begin{pmatrix} -1 & -4 & 5 \\ -3 & -1 & -7 \\ 2 & -3 & 1 \end{pmatrix}$

тогда по формуле $X = A^{-1}B$ получаем

$$X=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\frac{-1}{11}\begin{pmatrix}-1-4 & 5\\-3-1-7\\2-3 & 1\end{pmatrix}\begin{pmatrix}3\\2\\0\end{pmatrix}=\frac{-1}{11}\begin{pmatrix}-1\cdot3-4\cdot2 & 5\cdot0\\-3\cdot3-1\cdot2-7\cdot0\\2\cdot3-3\cdot2 & 1\cdot0\end{pmatrix}=\frac{-1}{11}\begin{pmatrix}-11\\-11\\0\end{pmatrix}=\begin{pmatrix}1\\1\\0\end{pmatrix}$$

Следовательно, $x_1=x_2=1$, $x_3=0$ - решение данной системы.

Метод *Крамера.* Используем теперь формулы Крамера:

$$\Delta=-11,\quad \Delta_1=\begin{vmatrix}3\,1-3\\2\,1\quad 2\\0\,1\quad 1\end{vmatrix}=-11,\qquad \Delta_2=\begin{vmatrix}2\,3-3\\1\,2\quad 2\\-1\,0\quad 1\end{vmatrix}=-11,\quad \Delta_3=$$

$$\begin{vmatrix}2 & 1 & 3\\1 & 1 & 2\\-1 & 1 & 0\end{vmatrix}=0,$$

$x_1=\Delta_1/\Delta=1$, $x_2=\Delta_2/\Delta=1$, $x_3=\Delta_3/\Delta=0$

2. Решить систему уравнений, используя метод Гаусса.

$$\begin{cases}x_1+2x_2+3x_3=12\\2x_1+x_2+2x_3=9\\-3x_1+x_2+4x_3=-4\end{cases},$$

Решение. Записываем расширенную матрицу данной системы

$$B=\left(\begin{array}{ccc|c}1 & 2 & 3 & 12.\\2 & 1 & 2 & 9\ .\\-3 & 1 & 4 & 10.\end{array}\right)$$

и выполняя преобразование над строками, сведем ее к трапецеподобному виду:

$$\left(\begin{array}{ccc|c}1 & 2 & 3 & 12.\\2 & 1 & 2 & 9\ .\\-3 & 1 & 4 & 10.\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 2 & 3 & 12\ .\\0 & -3 & -4 & -15.\\0 & 7 & 13 & 46\ .\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 2 & 3 & 12\ .\\0 & -3 & -42 & -15.\\0 & 1 & 5 & 16\ .\end{array}\right)\sim$$

$$\sim\left(\begin{array}{ccc|c}1 & 2 & 3 & 12\ .\\0 & 1 & 5 & 16\ .\\0 & -3 & -\ 4 & -15.\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 2 & 3 & 12.\\0 & 1 & 5 & 16.\\0 & 0 & 11 & 33.\end{array}\right)$$

Последняя строка соответствует уравнению $11x_3=33$, откуда $x_3=3$.

Далее записываем уравнение:

$x_2+5x_3=16$,

тогда

x_2=16-5x_3=16-15=1, x_2=1.

x_1+2x_2+3x_3=12,

тогда

x_1=12-2x_2-3x_3=12-2-9=1

Следовательно, решение системы таково: $x_1 = 1$, $x_2 = 1$, $x_3 = 3$.

3. Исследуйте систему уравнений $\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 1 \\ 3x_1 + 2x_2 + x_3 - x_4 = 1 \\ 2x_1 + 3x_2 + x_3 + x_4 = 1 \\ 2x_1 + 2x_2 + 2x_3 - x_4 = 1 \end{cases}$

на совместимость и в случае совместимости найдите ее решение:

Решение. Выписываем расширенную матрицу данной системы

$$B = \left(\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 3 & 2 & 1 & -1 & 1 \\ 2 & 3 & 1 & 1 & 1 \\ 2 & 2 & 2 & -1 & 1. \end{array}\right)$$

Находим ранг этой матрицы (и одновременно основной матрицы), выполняя элементарные преобразования строк:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 3 & 2 & 1 & -1 & 1 \\ 2 & 3 & 1 & 1 & 1 \\ 2 & 2 & 2 & -1 & 1. \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 0 & -4 & -8 & 2 & -2 \\ 0 & -1 & -5 & 3 & -1 \\ 0 & -1 & 1 & -2 & 0. \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 0 & -2 & -4 & 1 & -1 \\ 0 & -1 & -5 & 3 & -1 \\ 0 & -1 & 1 & -2 & 0. \end{array}\right) \sim$$

$$\sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 0 & 0 & 6 & -5 & 1 \\ 0 & -1 & -5 & 3 & -1 \\ 0 & 0 & 6 & -5 & 1. \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 0 & -1 & -5 & 3 & -1 \\ 0 & 0 & 6 & -5 & 1 \\ 0 & 0 & 0 & 0 & 0. \end{array}\right)$$

Из вида последней матрицы делаем вывод, что ранг основной матрицы равен 3. Ранг расширенной матрицы также равен трем, поскольку ранг расширенной матрицы не меньший за ранг основной матрицы, а последняя строка, которая содержит лишь нулевые элементы, не увеличивает ранга (все миноры четвертого порядка равны нулю).

Вывод. Данная система имеет множество решений. Эти решения находим так. По виду последней матрицы, записываем систему

$$\begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 1 \\ \quad -2x_2 - 5x_3 + 3x_4 = -1 \\ \qquad\qquad 6x_3 - 5x_4 = 1 \end{cases}$$

которая эквивалентна исходной системе. Двигаясь от последнего уравнения к первому, последовательно находим:

$$x_3=\frac{1+5x_4}{6};\quad x_2=-5x_3+3x_4+1=-5\cdot\frac{1+5x_4}{6}+3x_4+1=\frac{1-7x_4}{6};$$

$$x_1=1-x_2-3x_3+x_4=1-2\cdot\frac{1-7x_4}{6}-3\frac{1+5x_4}{6}+x_4=\frac{1+5x_4}{6};$$

Ответ можно записать так:

$$x_1=\frac{1+5t}{6};\quad x_2=\frac{1-7t}{6};\quad x_3=\frac{1+5t}{6};\quad x_4=t;\quad где\quad t\in R$$

4. Найдите все решения системы

$$\begin{cases}3x_1+2x_2-x_3=0\\2x_1-x_2+3x_3=0\\x_1+3x_2-4x_3=0\end{cases}$$

Решение. Выписываем расширенную матрицу данной системы

$$B=\left(\begin{array}{ccc|c}3&2&-1&0.\\2&-1&3&0.\\1&3&-4&0.\end{array}\right)$$

Находим ранг этой матрицы (и одновременно основной матрицы), выполняя элементарные преобразования строк:

$$\left(\begin{array}{ccc|c}3&2&-1&0.\\2&-1&3&0.\\1&3&-4&0.\end{array}\right)\sim\left(\begin{array}{ccc|c}1&3&-4&0.\\2&-1&3&0.\\3&2&-1&0.\end{array}\right)\sim\left(\begin{array}{ccc|c}1&3&-4&0.\\0&-7&11&0.\\0&-7&11&0.\end{array}\right)\sim\left(\begin{array}{ccc|c}1&3&-4&0.\\0&-7&11&0.\\&&&.\end{array}\right).$$

Поскольку угловой минор

$$M^{(2)}=\begin{vmatrix}1&3\\0&-7\end{vmatrix}=7\neq0$$

,

то ранг равен 2 (2<3). Следовательно, система имеет множество решений. По последней матрице записываем равносильную систему

$$\begin{cases}x_1+3x_2-4x_3=0\\-7x_2+11x_3=0\end{cases}$$

Значения $x_3=7t$, $x_2=11t$, где $t\in R$ удовлетворяют второе уравнение. Теперь из первого уравнения получаем $x_1=4x_3-3x_2=28t-33t=-5t$,

Ответ $x_1=-5t$, $x_2=11t$, $x_3=7t$, где $t\in R$

5. Найдите собственные числа и собственные векторы матрицы

$$A = \begin{pmatrix} 3 - 1 & 1 \\ -1 & 5 - 1 \\ 1 - 1 & 3 \end{pmatrix}$$

Решение. Решим характеристическое уравнение матрицы

$$\begin{vmatrix} 3-\lambda & -1 & 1 \\ -1 & 5-\lambda & -1 \\ 1 & -1 & 3-\lambda \end{vmatrix} = 0$$

$$(3-\lambda)(5-\lambda)(3-\lambda)+1+1-5+3-\lambda-3+\lambda,$$

$$(\lambda-3)(\lambda-2)(\lambda-6)=0, \quad \lambda_1=2, \quad \lambda_2=3, \quad \lambda_3=0$$

Следовательно, матрица A имеет три собственных числа. Найдем теперь собственные векторы, подставляя поочередно значения $\lambda_1, \lambda_2, \lambda_3$ в систему

$$\begin{cases} (3-\lambda)x_1 - x_2 + x_3 = 0, \\ -x_1 + (5-\lambda)x_2 - x_3 = 0, \\ x_1 - x_2 + (3-\lambda)x_3 = 0 \end{cases}$$

Собственному числу λ_2=2 соответствует собственный вектор X_1, координаты которого удовлетворяют систему

$$\begin{cases} (3-2)x_1 - x_2 + x_3 = 0, \\ -x_1 + (5-2)x_2 - x_3 = 0, \\ x_1 - x_2 + (3-2)x_3 = 0 \end{cases} \text{ или } \begin{cases} x_1 - x_2 + x_3 = 0, \\ -x_1 + 3x_2 - x_3 = 0, \\ x_1 - x_2 + x_3 = 0 \end{cases}$$

Отсюда
$$\begin{cases} 2x_2 = 0 \\ x_1 = -x_3 \end{cases}$$

и собственный вектор
$$X_1 = C_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

Аналогично находят собственные векторы

$$X_2 = C_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ и } X_3 = C_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

где C_1, C_2, C_3 - отличные от нуля действительные числа.

Микромодуль 5
Индивидуальные тестовые задачи

1.Решить систему:

а) матричным методом;

б) по формуле Крамара

2.1.1. $\begin{cases} x_1 + x_2 - x_3 = 0 \\ 2x_1 + x_2 + 2x_3 = 10 \\ x_1 - 3x_2 + x_3 = -2 \end{cases}$ 2.1.2. $\begin{cases} 2x_1 + x_2 + x_3 = 2 \\ x_1 + 2x_2 + x_3 = 3 \\ x_1 + x_2 + 2x_3 = -1 \end{cases}$

2.1.3. $\begin{cases} x_1 + 2x_2 + 3x_3 = -1 \\ 2x_1 + x_2 - 4x_3 = -3 \\ 3x_1 + 2x_2 + x_3 = -1 \end{cases}$ 2.1.4. $\begin{cases} 3x_1 + 4_2 - x_3 = -2 \\ 5x_1 + 3x_2 - 4x_3 = -2 \\ 4x_1 + 2x_2 + 3x_3 = 5 \end{cases}$

2.1.5. $\begin{cases} 2x_1 - x_2 + 3x_3 = -4 \\ x_1 + 2x_2 - 4x_3 = 19 \\ -3x_1 + 4x_2 + 2x_3 = 3 \end{cases}$ 2.1.6. $\begin{cases} 3x_1 - 2x_2 + 4x_3 = -17 \\ 4x_1 + 3x_2 - 2x_3 = 18 \\ 3x_1 + x_2 + 3x_3 = -7 \end{cases}$

2.1.7. $\begin{cases} x_1 + x_2 - x_3 = 4 \\ 2x_1 + x_2 + 2x_3 = 2 \\ 3x_1 - x_2 + x_3 = 0 \end{cases}$ 2.1.8. $\begin{cases} x_1 + 2x_2 - 2x_3 = -3 \\ 2x_1 + x_2 + 3x_3 = 8 \\ 3x_1 - 4x_2 + x_3 = 5 \end{cases}$

2.1.9. $\begin{cases} x_1 + 3x_2 - x_3 = 0 \\ 2x_1 - 4x_2 + 4x_3 = 6 \\ 3x_1 + 2x_2 + x_3 = 4 \end{cases}$ 2.1.10. $\begin{cases} 2x_1 + 3x_2 - 3x_3 = 2 \\ x_1 - 4x_2 + 5x_3 = 2 \\ 3x_1 + 4x_2 - 3x_3 = 4 \end{cases}$

2.1.11. $\begin{cases} 3x_1 + 4x_2 + 3x_3 = 3 \\ 4x_1 + 5x_2 - 3x_3 = 4 \\ 2x_1 + 3x_2 - 4x_3 = 2 \end{cases}$ 2.1.12. $\begin{cases} 2x_1 - 2x_2 + x_3 = 1 \\ x_1 \quad + \quad 2x_3 = 7 \\ x_1 \quad + 3x_2 = 7 \end{cases}$

2.1.13. $\begin{cases} 3x_1 + x_2 + 2x_3 = 2 \\ x_1 + 3x_2 + 2x_3 = 2 \\ 2x_1 + x_2 - 2x_3 = -1 \end{cases}$ 2.1.14. $\begin{cases} 3x_1 - 2x_2 + x_3 = 0 \\ 7x_1 + x_2 + 4x_3 = 0 \\ \qquad x_2 + 5x_3 = 4 \end{cases}$

2.1.15. $\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 \quad + 2x_3 = 4 \\ 3x_1 + x_2 = 5 \end{cases}$ 2.1.16. $\begin{cases} 2x_1 + x_2 + x_3 = 2 \\ 3x_1 + 3x_2 + 2x_3 = 5 \\ x_1 - x_2 + x_3 = -3 \end{cases}$

2.1.17. $\begin{cases} x_1 + 2x_2 + x_3 = 2 \\ 2x_1 \quad + \quad 4x_3 = -13 \\ 3x_1 - 3x_2 - 4x_3 = 0 \end{cases}$ 2.1.18. $\begin{cases} 3x_1 + x_2 + x_3 = 4 \\ x_1 + 3x_2 + x_3 = 6 \\ x_1 + x_2 + 3x_3 = 0 \end{cases}$

2.1.19. $\begin{cases} x_1 + 3x_2 \quad = 6 \\ x_1 - x_2 - x_3 = 0 \\ 2x_1 + 2x_2 + 3x_3 = -2 \end{cases}$ 2.1.20. $\begin{cases} 3x_1 + 3x_2 + 4x_3 = 4 \\ 5x_1 - 7x_2 + 8x_3 = 20 \\ 4x_1 + 5x_2 - 7x_3 = -8 \end{cases}$

2. Решить однородную систему линейных алгебраических уравнений

2.2.1. $\begin{cases} x_2 - x_3 + 2x_4 = 0 \\ -2x_1 + x_2 + 3x_3 - x_4 = 0 \\ 3x_1 - 2x_2 + x_3 + x_4 = 0 \\ x_1 + 3x_3 + 2x_4 = 0 \end{cases}$

2.2.2. $\begin{cases} x_1 + x_2 - 2x_3 + x_4 = 0 \\ -x_1 + 2x_2 + x_3 - x_4 = 0 \\ 3x_1 + x_3 + x_4 = 0 \\ 3x_1 + 3x_2 + x_3 + x_4 = 0 \end{cases}$

2.2.3. $\begin{cases} -2x_1 + 2x_2 - x_3 - x_4 = 0 \\ 3x_1 + x_2 - x_4 = 0 \\ -x_1 + 2x_2 + 2x_3 + 2x_4 = 0 \\ 4x_2 + 3x_3 + 2x_4 = 0 \end{cases}$

2.2.4. $\begin{cases} x_1 - x_2 + 2x_3 + x_4 = 0 \\ 5x_1 - 2x_2 + 5x_3 + 4x_4 = 0 \\ -2x_1 + x_2 - x_3 - 2x_4 = 0 \\ -2x_1 + x_2 - x_3 - 2x_4 = 0 \end{cases}$

2.2.5. $\begin{cases} 3x_1 + x_2 + 3x_3 + 2x_4 = 0 \\ -x_1 + 2x_2 + 3x_3 + x_4 = 0 \\ 2x_1 + 4x_2 + 5x_3 + 5x_4 = 0 \\ x_2 - x_3 + 2x_4 = 0 \end{cases}$

2.2.6. $\begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ -x_1 + x_2 - x_3 - x_4 = 0 \\ x_1 - x_2 + 2x_3 + 3x_4 = 0 \\ x_1 + x_2 + 2x_3 + x_4 = 0 \end{cases}$

2.2.7. $\begin{cases} 2x_1 + x_2 - 2x_3 + x_4 = 0 \\ x_1 - x_2 + x_4 = 0 \\ -x_1 + 2x_2 + x_3 - x_4 = 0 \\ 2x_1 + 2x_2 - x_3 + 2x_4 = 0 \end{cases}$

2.2.8. $\begin{cases} 3x_1 - x_2 + 2x_3 + 3x_4 = 0 \\ -2x_1 + x_2 - x_3 - 2x_4 = 0 \\ x_1 + 2x_2 + x_3 - x_4 = 0 \\ 2x_1 + 2x_2 + x_3 = 0 \end{cases}$

2.2.9. $\begin{cases} -x_1 + x_3 - x_4 = 0 \\ -2x_1 + x_2 - 2x_3 + x_4 = 0 \\ 2x_1 + 3x_2 - x_3 + x_4 = 0 \\ -x_1 + 4x_2 - 2x_3 + x_4 = 0 \end{cases}$

2.2.10. $\begin{cases} x_1 - 2x_2 + 2x_3 - x_4 = 0 \\ 3x_1 + x_2 - x_3 + x_4 = 0 \\ 2x_1 + 3x_3 + x_4 = 0 \\ 6x_1 - x_2 + 4x_3 + x_4 = 0 \end{cases}$

2.2.11. $\begin{cases} 4x_1 + x_2 + 3x_3 - 3x_4 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ -3x_1 - x_2 - 2x_3 + x_4 = 0 \\ 2x_1 - 2x_2 + x_3 - 2x_4 = 0 \end{cases}$

2.2.12. $\begin{cases} x_1 - x_2 + 2x_3 - x_4 = 0 \\ 3x_1 + 2x_2 - x_3 + x_4 = 0 \\ -2x_1 + x_2 + x_3 + 2x_4 = 0 \\ x_1 + x_2 + x_3 + x_4 = 0 \end{cases}$

2.2.13. $\begin{cases} -2x_1 + 2x_2 + x_3 - x_4 = 0 \\ 3x_1 - x_3 + x_4 = 0 \\ x_1 - x_2 + 3x_3 + 2x_4 = 0 \\ 2x_1 + x_2 + 3x_3 + 2x_4 = 0 \end{cases}$

2.2.14. $\begin{cases} -3x_1 + x_2 + 2x_3 - x_4 = 0 \\ 2x_1 + 3x_2 - x_3 + x_4 = 0 \\ x_1 - 2x_2 + 3x_3 + 2x_4 = 0 \\ x_2 + 2x_3 + x_4 = 0 \end{cases}$

2.2.15. $\begin{cases} x_1 + 3_2 - x_3 + 2x_4 = 0 \\ 2x_1 + x_2 + 3x_3 + x_4 = 0 \\ 3x_1 - 2_2 - x_3 - 2x_4 = 0 \\ 2x_1 + 2x_2 + x_3 + x_4 = 0 \end{cases}$

2.2.16. $\begin{cases} x_1 + 2x_3 - x_4 = 0 \\ -2x_1 + x_2 + x_4 = 0 \\ 2x_1 + x_2 + x_3 + 2x_4 = 0 \\ x_1 + 2x_2 + 2x_3 + 2x_4 = 0 \end{cases}$

2.2.17. $\begin{cases} 3x_1+2x_2-x_3+2x_4=0 \\ -x_1+x_3+3x_4=0 \\ 2x_1+x_2+2x_3+x_4=0 \\ -x_2+2x_3-x_4=0 \end{cases}$ 2.2.18. $\begin{cases} -2x_1+3x_2+2x_3+x_4=0 \\ x_1-2x_2+x_3-x_4=0 \\ 3x_1+x_2-x_3=0 \\ x_1+x_2+x_3=0 \end{cases}$

2.2.19. $\begin{cases} x_1+x_2+2x_3+x_4=0 \\ 2x_2+3x_3-x_4=0 \\ -2x_1-3x_2-x_3+x_4=0 \\ -x_1-x_2+4x_3+x_4=0 \end{cases}$ 2.2.20. $\begin{cases} 2x_1+3x_2+2x_3-x_4=0 \\ -x_1+x_2+2x_4=0 \\ 3x_1-2x_2-x_3=0 \\ 4x_1+2x_2+x_3+x_4=0 \end{cases}$

2.2.21. $\begin{cases} -3x_1+2_2-2x_3+2x_4=0 \\ 2x_1-x_2+x_3-x_4=0 \\ -x_1-x_3+x_4=0 \\ -2x_1+x_2-2x_3+2x_4=0 \end{cases}$ 2.2.22. $\begin{cases} 4x_1-2x_2+3x_3-x_4=0 \\ -2x_1-x_2-2x_3+x_4=0 \\ x_2+x_3-4x_4=0 \\ 2x_1-2x_2+2x_3+x_4=0 \end{cases}$

2.2.23. $\begin{cases} 2x_2-x_3+2x_4=0 \\ x_1-x_2+2x_3-x_4=0 \\ 3x_1+x_3+x_4=0 \\ 4x_1+x_2+2x_3+2x_4=0 \end{cases}$ 2.2.24. $\begin{cases} 3x_1+2x_3+2x_4=0 \\ -2x_1+x_2-x_3=0 \\ 3x_1-2x_2+x_3+x_4=0 \\ 4x_1-x_2+2x_3+3x_4=0 \end{cases}$

2.2.25. $\begin{cases} x_1+x_2+x_3+x_4=0 \\ 2x_1-x_2+2x_3-x_4=0 \\ -3x_1+x_2-2x_3+2x_4=0 \\ x_1+2x_2+2x_3+3x_4=0 \end{cases}$ 2.2.26. $\begin{cases} 3x_1+x_2-x_3+3x_4=0 \\ -x_1+x_2+2x_3+3x_4=0 \\ -x_1-x_2+x_3=0 \\ x_1+x_2+2x_3+x_4=0 \end{cases}$

2.2.27. $\begin{cases} 3x_1+3x_2-4x_4=0 \\ 2x_1+x_2+x_3+x_4=0 \\ x_1-2x_2-3x_4=0 \\ 2x_1+2x_2+x_3-2x_4=0 \end{cases}$ 2.2.28. $\begin{cases} x_1+x_2+x_3+x_4=0 \\ -2x_1-x_2+2x_3+x_4=0 \\ 2x_1-x_3+x_4=0 \\ x_1+x_3+3x_4=0 \end{cases}$

2.2.29. $\begin{cases} -4x_1+2x_2-x_4=0 \\ 3x_1-x_2+x_3+2x_4=0 \\ 2x_1+3x_2-2x_3+x_4=0 \\ x_1+3x_2-x_3+2x_4=0 \end{cases}$ 2.2.30. $\begin{cases} x_1-x_2+2x_3+x_4=0 \\ 3x_2+x_3+2x_4=0 \\ -2x_1+x_2-x_3-2x_4=0 \\ -x_1+x_3-x_4=0 \end{cases}$

3. Дана определенная система уравнений:

$$\left.\begin{array}{r} x_1+2x_2+3x_3+4x_4=30 \\ -x_1+2x_2+3x_3+4x_4=10 \\ x_2-x_3+x_4=3 \\ x_1+x_2+x_3+x_4=10 \end{array}\right\}.$$

Найти ее решение с помощью:

а) правила Крамера;

б) обращение матрицы системы;

в) алгоритма Гаусса;

г) алгоритма Гаусса -Жордаиа.

4. Найти *LU*-разложение для матрицы системы уравнений из задачи 3, воспользовавшись:

а) компактной схемой; б) методом исключения.

5. Существует ли решение приведенных ниже систем уравнений и является ли оно единственным?

$$\text{а)}\ \left.\begin{aligned} 3x_1+2x_2+x_3&=7\\ 2x_1+x_2-2x_3&=8\\ x_1+\frac{3}{4}x_2+x_3&=5\end{aligned}\right\};\qquad \text{б)}\ \left.\begin{aligned} 3x_1+2x_2+4x_3&=7\\ 2x_1+x_2+x_3&=4\\ x_1+3x_2+5x_3&=2\end{aligned}\right\};$$

$$\text{в)}\ \left.\begin{aligned} 2x_1+7x_2&=4\\ 3x_1+3x_2+6x_3&=3\\ 2x_1+2x_2+4x_3&=2\end{aligned}\right\};\qquad \text{г)}\ \left.\begin{aligned} 2x_1+8x_2+7x_3&=0\\ x_1+2x_2+4x_3&=0\\ 2x_1+4x_2+6x_3&=0\end{aligned}\right\}.$$

6. Решите с помощью алгоритма Гаусса те системы уравнений из задачи 5, решения которых существуют.

7. При каких значениях α система

$$\left.\begin{aligned} (5-\alpha)x_1-3x_2+2x_3&=0\\ 6x_1-(4+\alpha)x_2+4x_3&=0\\ 4x_1-4x_2+(5-\alpha)x_3&=0\end{aligned}\right\}$$

имеет ненулевые решения? Решите эту систему при одном из найденных значений α.

8. Дана система уравнений

$$\begin{bmatrix} 2 & 2 & 5 & 3\\ 6 & 1 & 5 & 4\\ 4 & -1 & 0 & 1\\ 2 & 0 & 1 & 1\end{bmatrix}\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}=\begin{bmatrix} 5\\ 5\\ 0\\ 1\end{bmatrix}.$$

Проверьте ее на совместимость и найдите решение, если оно существует.

9. Решите уравнение *Ax=b* при заданной матрице *A* и различных значениях вектора $b = b^{(i)}$ $(i = 1, 2, 3)$:

$$A=\begin{bmatrix} 2 & 3 & 1\\ 4 & -1 & 1\\ 5 & 1 & 1\end{bmatrix};\quad b^{(1)}=\begin{bmatrix} 1\\ 5\\ 1\end{bmatrix};\quad b^{(2)}=\begin{bmatrix} -3\\ -1\\ -1\end{bmatrix};\quad b^{(3)}=\begin{bmatrix} 2\\ 5\\ 3\end{bmatrix}.$$

Запишите условие этой задачи в виде единого матричного уравнения $AX = B$ и представьте его решение как матрицу *X*.

10. Решите матричное уравнение $XA = B$, где

$$A=\begin{bmatrix} 1 & -1 & -2 \\ 2 & -1 & -1 \\ -1 & 3 & 2 \end{bmatrix}; \quad B=\begin{bmatrix} -1 & 1 & -1 \\ 3 & -1 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

11. Покажите, что матричное уравнение $AX = B$ не имеет решения для матрицы X, если ранг B больше ранга A.

10. Уравнение n-узловой электрической схемы с одним входом и одним выходом, имеющим общий узел O (рис. 2.7), можно представить в виде:

$$\begin{bmatrix} y_{11} & y_{12} & y_{13} & \dots & y_{1n} \\ y_{21} & y_{22} & y_{23} & \dots & y_{2n} \\ y_{31} & y_{32} & y_{33} & \dots & y_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ y_{n1} & y_{n2} & y_{n3} & \dots & y_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_n \end{bmatrix} = \begin{bmatrix} i_1 \\ -i_2 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

или в матричной записи $Yu = i$, где Y — матрица проводимости схемы; $u = (u_1,\ u_2.....u_n)$ — вектор узловых напряжений, отсчитываемых от базисного узла O; i_1 и u_2 — соответственно входной и выходный токи.

Рис. 2.7. Электрическая схема с одним входом и одним выходом, которые имеют общий узел O.

а) Воспользовавшись правилом Крамера, покажите, что входное и исходное напряжение можно представить в виде:

$$\left.\begin{aligned} u_1 &= \frac{\Delta_{11}}{\Delta} i_1 - \frac{\Delta_{21}}{\Delta} i_2 \\ u_2 &= \frac{\Delta_{12}}{\Delta} i_1 - \frac{\Delta_{22}}{\Delta} i_2 \end{aligned}\right\},$$

где Δ — определитель матрицы Y; Δ_{ij} $(u, j = 1,2)$ — алгебраические дополнения соответствующих элементов этой матрицы.

б) Исключив переменные u_3, u_4..... u_n, приведите исходную систему уравнений к виду:

$$\left.\begin{aligned} i_1 &= \frac{1}{\Delta_{11,22}} (\Delta_{22} u_1 - \Delta_{21} u_2) \\ i_2 &= \frac{1}{\Delta_{11,22}} (\Delta_{12} u_1 - \Delta_{11} u_2) \end{aligned}\right\}.$$

в) Запишите общие выражения для передачи напряжения при холостом ходе K_U и передачи тока при коротком замыкании K_I, т.е.

$$K_U = \left(\frac{u_2}{u_1}\right)_{i_2=0}, \qquad K_I = \left(\frac{i_2}{i_1}\right)_{u_2=0}.$$

13. Система уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{aligned}\right\}$$

представляется на плоскости двумя прямыми (рис. 2.8).

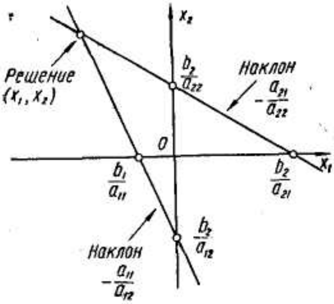


Рис. 2.8. Геометрическая интерпретация системы двух уравнений на плоскости.

Если

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{12}a_{22} - a_{12}a_{21} \neq 0,$$

то данная система определена и ее решение:

$$x_1 = \frac{1}{\Delta}(a_{22}b_1 - a_{12}b_2);$$
$$x_2 = \frac{1}{\Delta}(a_{11}b_2 - a_{21}b_1)$$

изображается точкой (x_1, x_2) пересечение этих прямых. Рассмотрите геометрическое представление данной системы в случаях, когда она:

а) неоднородная ($b_1 \neq 0$, $b_2 \neq 0$) и совместная при $\Delta = 0$;

б) неоднородная и несовместная;

в) однородная ($b_1 = b_2 = 0$) при $\Delta \neq 0$;

г) однородная при $\Delta = 0$.

14. Покажите, что в частном случае исключения переменных при $b' = 0$ и $s = 2$ система уравнений n-го порядка приводится к виду:

$$\frac{1}{\Delta_{11,22}}\begin{bmatrix} \Delta_{22} & -\Delta_{21} \\ -\Delta_{12} & \Delta_{11} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Модуль 3
Матричные многочлены и функции от матриц

Микромодуль 6
Матричные многочлены

С каждой квадратной матрицей связаны два многочлена: характеристический и минимальный. Эти многочлены играют большую роль в различных вопросах теории матриц. Так, например, понятие о функции от матрицы, которое мы введем в следующем микромодуле, будет целиком основываться на понятии о минимальном многочлене матрицы. В этом модуле рассматриваются свойства характеристического и минимального многочлена. Этому исследованию предпосылаются основные сведения о многочленах с матричными коэффициентами и о действиях над ними.

3.1. Сложение и умножение матричных многочленов

Рассмотрим квадратную *многочленную* матрицу $A(\lambda)$, т.е. квадратную матрицу, элементами которой являются многочлены относительно λ (с коэффициентами из данного числового поля *K):*

$$A(\lambda) = \| a_{ik}(\lambda) \|_1^n = \| a_{ik}^{(0)}\lambda^m + a_{ik}^{(1)}\lambda^{m-1} + \ldots + a_{ik}^{(m)} \|_1^n. \tag{3.1}$$

Матрицу $A(\lambda)$ можно представить в виде многочлена с матричными коэффициентами, расположенного по степеням λ:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m, \tag{3.2}$$

где

$$A_j = \| a_{ik}^{(j)} \|_1^n \quad (j = 0, 1, \ldots, m). \tag{3.3}$$

Число m называется *степенью* многочлена, если $A_0 \neq 0$. Число n называется *порядком* многочлена. Многочлен (3.1) будем называть *регулярным,* если $| A_0 | \neq 0$.

Многочлен с матричными коэффициентами мы будем иногда называть *матричным многочленом.* В отличие от матричного многочлена обычный многочлен со скалярными коэффициентами будем называть *скалярным многочленом.*

Рассмотрим основные действия над матричными многочленами. Пусть даны два матричных многочлены одного и того же порядка $A(\lambda)$ и B (λ). Обозначим через m наибольшую из степеней этих многочленов. Эти многочлены можно записать в виде

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m,$$
$$B(\lambda) = B_0\lambda^m + B_1\lambda^{m-1} + \ldots + B_m.$$

Тогда

$$A(\lambda) \pm B(\lambda) = (A_0 \pm B_0)\lambda^m + (A_1 \pm B_1)\lambda^{m-1} + \ldots + A_m \pm B_m,$$

т.е. *сумма (разность) двух матричных многочленов одного и того же порядка может быть представлена в виде многочлена, степень которого не превосходит наибольшего из степеней данных многочленов.*

Пусть даны два матричных многочлена A (λ) и B (λ) степеней m и p одного и того же порядка n:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m \qquad (A_0 \neq 0),$$
$$B(\lambda) = B_0\lambda^p + B_1\lambda^{p-1} + \ldots + B_p \qquad (B_0 \neq 0).$$

Тогда

$$A(\lambda)\,B(\lambda) = A_0B_0\lambda^{m+p} + (A_0B_1 + A_1B_0)\lambda^{m+p-1} + \ldots + A_mB_p. \tag{3.4}$$

Если бы мы перемножили $B(\lambda)$ на $A(\lambda)$ (т.е. изменили бы порядок сомножителей), то мы получили бы, вообще говоря, другой многочлен.

Умножение матричных многочленов обладает еще одним специфическим свойством. В отличие от произведения скалярных многочленов произведение матричных многочленов (3.4) может иметь степень, меньшую $m+p$, т.е. меньшую суммы степеней сомножителей. Действительно, в (3.4) произведение матриц A_0B_0 может равняться нулю при $A_0 \neq 0$ и $B_0 \neq 0$. Однако, если хотя бы одна из матриц A_0 и B_0 неособенная, то изс $A_0 \neq 0$ и $B_0 \neq 0$ следует: $A_0B_0 \neq 0$. Таким образом, *произведение двух матричных многочленов равно многочлену, степень которого меньше или равна сумме степеней сомножителей. Если хотя бы один из двух сомножителей - регулярный многочлен, то в этом случае степень произведения всегда равна сумме степеней сомножителей.*

Матричный многочлен n-го порядка $A(\lambda)$ можно записать двояко:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m \tag{3.5}$$

и

$$A(\lambda) = \lambda^m A_0 + \lambda^{m-1}A_1 + \ldots + A_m. \tag{3.6}$$

Обе записи при скалярном λ дают один и тот же результат. Однако если мы пожелаем вместо скалярного аргумента λ подставить квадратную матрицу n-го порядка Λ, то результаты подстановок в (3.5) и (3.6) будут, вообще говоря, различны, так как степени матрицы Λ могут не быть перестановочными с матричными коэффициентами $A_0, A_1, \ldots, Am$.

Положим

$$A(\Lambda) = A_0\Lambda^m + A_1\Lambda^{m-1} + \ldots + A_m \tag{3.7}$$

и

$$\hat{A}(\Lambda) = \Lambda^m A_0 + \Lambda^{m-1}A_1 + \ldots + A_m \tag{3.8}$$

и будем называть $A(\Lambda)$ правым, а $\hat{A}(\Lambda)$ левым значением матричного многочлена $A(\lambda)$ при подстановке вместо λ матрицы Λ (в «правом» значении $A(\lambda)$ степени матрицы Λ стоят справа от коэффициентов, а в «левом» — слева).

Рассмотрим снова два матричных многочлена

$$A(\lambda) = \sum_{i=0}^{m} A_{m-i}\lambda^i, \quad B(\lambda) = \sum_{k=0}^{p} B_{p-k}\lambda^k$$

и их произведение

$$P(\lambda) = \sum_{i=0}^{m}\sum_{k=0}^{p} A_{m-i}\lambda^i B_{p-k}\lambda^k = \sum_{i=0}^{m}\sum_{k=0}^{p} A_{m-i}B_{p-k}\lambda^{i+k} = \sum_{j=0}^{m+p}\Big(\sum_{i+k=j} A_{m-i}B_{p-k}\Big)\lambda^j \tag{3.9}$$

и

$$P(\lambda) = \sum_{i=0}^{m}\sum_{k=0}^{p} \lambda^i A_{m-i}\lambda^k B_{p-k} = \sum_{i=0}^{m}\sum_{k=0}^{p} \lambda^{i+k}A_{m-i}B_{p-k} = \sum_{j=0}^{m+p}\lambda^j \sum_{i+k=j} A_{m-i}B_{p-k}. \tag{3.10}$$

Преобразования в тождестве (3.9) сохраняют свою силу при замене λ матрицей n-го порядка Λ, если только матрица Λ перестановочна со всеми матричными коэффициентами B_{p-k} (в этом случае и любая степень матрицы Λ перестановочний со всеми коэффициентами B_{p-k}).

Аналогично в тождестве (3.10) можно заменить скаляр λ матрицей Λ, если матрица Λ перестановочна со всеми коэффициентами A_{m-i}. В первом случае получаем:

$$P(\Lambda) = A(\Lambda)\, B(\Lambda), \tag{3.11}$$

во втором

$$\hat{P}(\Lambda) = \hat{A}(\Lambda)\, \hat{B}(\Lambda). \tag{3.12}$$

Таким образом, *правое (левое) значение произведения двух матричных многочленов равно произведению правых (левых) значений сомножителей, если матрица-аргумент Λ перестановочна со всеми коэффициентами правого (левого) сомножителя.*

Если $S(\lambda)$ — сумма двух матричных многочленов n-го порядка $A(\lambda)$ и $B(\lambda)$, то при замене скаляра λ любой матрицей n-го порядка Λ всегда справедливы тождества

$$S(\Lambda) = A(\Lambda) + B(\Lambda), \quad \hat{S}(\Lambda) = \hat{A}(\Lambda) + \hat{B}(\Lambda). \tag{3.13}$$

3.2. Правое и левое деление матричных многочленов.

Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ одного и того же порядка n, причем $B(\lambda)$ — регулярный многочлен:

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m \quad (A_0 \neq 0),$$
$$B(\lambda) = B_0\lambda^p + B_1\lambda^{p-1} + \ldots + B_p \quad (|B_0| \neq 0).$$

Мы будем говорить, что матричные многочлены $Q(\lambda)$ и $R(\lambda)$ являются соответственно *правым частным* и *правым остатком* при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda) = Q(\lambda)\, B(\lambda) + R(\lambda) \tag{3.14}$$

и степень $R(\lambda)$ меньше степени $B(\lambda)$.

Аналогично будем называть многочлены $\hat{Q}(\lambda)$ и $\hat{R}(\lambda)$ соответственно *левым частным* и *левым остатком* при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda) = B(\lambda)\,\hat{Q}(\lambda) + \hat{R}(\lambda) \tag{3.15}$$

и степень $\hat{R}$ (λ) меньше степени $\hat{B}$ (λ).

Обращаем внимание читателя, что при «правом» делении (т.е. при нахождении правого частного и правого остатка) в (3.5) на «делитель» B(λ) частное Q(λ) умножается справа, а при «левом» делении в (3.7) на делитель B (λ) частное $\hat{Q}$ (λ) умножается слева. В общем случае многочлены Q (λ) и R (λ) не совпадают с $\hat{Q}$ (λ) и $\hat{R}$ *(λ).*

Покажем, что *как правое, так и левое деление матричных многочленов одного и того же порядка всегда осуществимо и однозначное, если делитель — регулярный многочлен.*

Рассмотрим правое деление A (λ) на B (λ). Если $m < p$, можно положить Q (λ) = 0, R (λ) = A (λ). В случае $m \geq p$ для нахождения частного Q (λ) и остатка R(λ) применим обычную схему деления многочлена на многочлен. «Разделим» старший член делимого A_0 λ^m на старший член делителя $B_0\lambda^{p.}$ Получим старший член искомого частного A_0B^{-1}λт-^{p}. Умножим этот член справа на делитель B (λ) и полученное произведение вычтем из A (λ). Найдем «первый остаток» $A^{(1)}$ (λ):

$$A(\lambda) = A_0B_0^{-1}\lambda^{m-p}\,B(\lambda) + A^{(1)}(\lambda). \tag{3.16}$$

Степень $m^{(1)}$ многочлена $A^{(1)}$(λ) меньше m:

$$A^{(1)}(\lambda) = A_0^{(1)}\lambda^{m^{(1)}} + \dots \quad (A_0^{(1)} \neq 0,\ m^{(1)} < m). \tag{3.17}$$

Если $m^{(1)} \geq p$, то, повторяя этот процесс, получаем:

$$\left.\begin{aligned} A^{(1)}(\lambda) &= A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p}B(\lambda) + A^{(2)}(\lambda), \\ A^{(2)}(\lambda) &= A_0^{(2)}\lambda^{m^{(2)}} + \dots \qquad (m^{(2)} < m^{(1)}) \end{aligned}\right\} \tag{3.18}$$

Так как степени многочленов A (λ), $A^{(1)}$(λ), $A^{(2)}$(λ), ... убывают, то на некотором этапе мы придем к остатку R(λ), степень которого меньше p. Тогда из (3.16)—(3.18) будет следовать:

$$A(\lambda) = Q(\lambda)\,B(\lambda) + R(\lambda),$$

где

$$Q(\lambda) = A_0B_0^{-1}\lambda^{m-p} + A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p} + \dots$$

Докажем теперь однозначность правого деления. Пусть одновременно

$$A(\lambda) = Q(\lambda)\,B(\lambda) + R(\lambda)$$

(3.19)

и

$$A(\lambda) = Q^*(\lambda) B(\lambda) + R^*(\lambda),$$

(3.20)

где степени многочленов $R(\lambda)$ и $R^*(\lambda)$ меньше степени $B(\lambda)$, т.е. меньше p. Вычитая почленно (3.15) из (3.16), получим:

$$[Q(\lambda) - Q^*(\lambda)] B(\lambda) = R^*(\lambda) - R(\lambda).$$

(3.21)

Если бы $Q(\lambda) — Q^*(\lambda) \not\equiv 0$, то, поскольку $|B_0| \neq 0$, степень левой части равенства (3.21) равнялась бы сумме степеней $B(\lambda)$ и $Q(\lambda) — Q^*(\lambda)$ и потому была бы $\geq p$. Это невозможно, так как степень многочлена, стоящего в правой части равенства (3.21), меньше p. Таким образом,
$Q(\lambda) — Q^*(\lambda) \equiv 0$, а тогда из (3.21) $R^*(\lambda)-R(\lambda) \equiv 0$, т.е.

$$Q(\lambda) = Q^*(\lambda), \quad R(\lambda) = R^*(\lambda).$$

Аналогично устанавливаются существование и единственность левого частного и левого остатка.

Заметим, что возможность и однозначность левого деления $A(\lambda)$ на $B(\lambda)$ следует из возможности и однозначности правого деления транспонированных матриц $A'(\lambda)$ и $B'(\lambda)$. (Из регулярности $B(\lambda)$ следует регулярность $B'(\lambda)$.) Действительно, из

$$A'(\lambda) = Q_1(\lambda) B'(\lambda) + R_1(\lambda)$$

следует:

$$A(\lambda) = B(\lambda) Q_1'(\lambda) + R_1'(\lambda).$$

(3.22)

В силу этих же соображений левое деление $A(\lambda)$ на $B(\lambda)$ однозначно, так как из неоднозначности левого деления $A(\lambda)$ на $B(\lambda)$ следовала бы неоднозначность правого деления $A'(\lambda)$ на $B'(\lambda)$.

Сопоставление (3.15) и (3.22) дает:

$$\hat{Q}(\lambda) = Q_1'(\lambda), \quad \hat{R}(\lambda) = R_1'(\lambda).$$

Пример.

$$A(\lambda)=\begin{Vmatrix} \lambda^3+\lambda & 2\lambda^3+\lambda^2 \\ -\lambda^3-2\lambda^2+1 & 3\lambda^3+\lambda \end{Vmatrix}=$$

$$=\overbrace{\begin{Vmatrix} 1 & 2 \\ -1 & 3 \end{Vmatrix}}^{A_0}\lambda^3+\begin{Vmatrix} 0 & 1 \\ -2 & 0 \end{Vmatrix}\lambda^2+\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}\lambda+\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix},$$

$$B(\lambda)=\begin{Vmatrix} 2\lambda^2+3 & -\lambda^2+1 \\ -\lambda^2-1 & \lambda^2+2 \end{Vmatrix}=\overbrace{\begin{Vmatrix} 2 & -1 \\ -1 & 1 \end{Vmatrix}}^{B_0}\lambda^2+\begin{Vmatrix} 3 & 1 \\ -1 & 2 \end{Vmatrix},$$

$$|B_0|=1,\quad B_0^{-1}=\begin{Vmatrix} 1 & 1 \\ 1 & 2 \end{Vmatrix},\quad A_0B_0^{-1}=\begin{Vmatrix} 3 & 5 \\ 2 & 5 \end{Vmatrix},\quad A_0B_0^{-1}B(\lambda)=\begin{Vmatrix} \lambda^2+4 & 2\lambda^2+13 \\ -\lambda^2+1 & 3\lambda^2+12 \end{Vmatrix},$$

$$A^{(1)}(\lambda)=\begin{Vmatrix} \lambda^3+\lambda & 2\lambda^3+\lambda^2 \\ -\lambda^3-2\lambda^2+1 & 3\lambda^3+\lambda \end{Vmatrix}-\begin{Vmatrix} \lambda^3+4\lambda & 2\lambda^3+13\lambda \\ -\lambda^3+\lambda & 3\lambda^3+12\lambda \end{Vmatrix}=$$

$$=\begin{Vmatrix} -3\lambda & \lambda^2-13\lambda \\ -2\lambda^2-\lambda+1 & -11\lambda \end{Vmatrix},$$

$$A^{(1)}(\lambda)=\begin{Vmatrix} 0 & 1 \\ -2 & 0 \end{Vmatrix}\lambda^2+\begin{Vmatrix} -3 & -13 \\ -1 & -11 \end{Vmatrix}\lambda+\begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix},$$

$$A_0^{(1)}B_0^{-1}=\begin{Vmatrix} 0 & 1 \\ -2 & 0 \end{Vmatrix}\cdot\begin{Vmatrix} 1 & 1 \\ 1 & 2 \end{Vmatrix}=\begin{Vmatrix} 1 & 2 \\ -2 & -2 \end{Vmatrix},$$

$$A_0^{(1)}B_0^{-1}B(\lambda)=\begin{Vmatrix} 1 & 2 \\ -2 & -2 \end{Vmatrix}\cdot\begin{Vmatrix} 2\lambda^2+3 & -\lambda^2+1 \\ -\lambda^2-1 & \lambda^2+2 \end{Vmatrix}=\begin{Vmatrix} 1 & \lambda^2+5 \\ -2\lambda^2-4 & -6 \end{Vmatrix},$$

$$R(\lambda)=A^{(1)}(\lambda)-A_0^{(1)}B_0^{-1}B(\lambda)=$$

$$=\begin{Vmatrix} -3\lambda & \lambda^2-13\lambda \\ -2\lambda^2-\lambda+1 & -11\lambda \end{Vmatrix}-\begin{Vmatrix} 1 & \lambda^2+5 \\ -2\lambda^2-4 & -6 \end{Vmatrix}=\begin{Vmatrix} -3\lambda-1 & -13\lambda-5 \\ -\lambda+5 & -11\lambda+6 \end{Vmatrix},$$

$$Q(\lambda)=A_0B_0^{-1}\lambda+A_0^{(1)}B_0^{-1}=\begin{Vmatrix} 3 & 5 \\ 2 & 5 \end{Vmatrix}\lambda+\begin{Vmatrix} 1 & 2 \\ -2 & -2 \end{Vmatrix}=\begin{Vmatrix} 3\lambda+1 & 5\lambda+2 \\ 2\lambda-2 & 5\lambda-2 \end{Vmatrix}.$$

Предлагаем читателю как упражнение непосредственно проверить, что

$$A(\lambda)=Q(\lambda)B(\lambda)+R(\lambda).$$

Рассмотрим произвольный матричный многочлен n-го порядка

$$F(\lambda)=F_0\lambda^m+F_1\lambda^{m-1}+\ldots+F_m\qquad(F_0\neq 0). \tag{3.23}$$

Разделим его на бином $\lambda E - A$ справа и слева:

$$F(\lambda)=Q(\lambda)(\lambda E-A)+R,\qquad F(\lambda)=(\lambda E-A)\hat{Q}(\lambda)+\hat{R}. \tag{3.24}$$

В этом случае правый остаток R и левый остаток $\hat{R}$ не будут зависеть от λ. Для определения правого значения $F(A)$ и левого $\hat{F}(A)$ можно соответственно в тождествах (3.24) заменить скаляр λ на матрицу A, поскольку матрица A перестановочна с матричными коэффициентами бинома $\lambda E - A$ (см. п. 3.1):

$$F(A) = Q(A)(A - A) + R = R, \qquad \hat{F}(A) = (A - A)\hat{Q}(A) + \hat{R} = \hat{R}. \tag{3.25}$$

Теорема 1 (обобщенная теорема Безу). *При правом (левом) делении матричного многочлена $F(\lambda)$ на бином $\lambda E - A$ остаток от деления равен $F(A)$ (соответственно $\hat{F}(A)$).*

Из теоремы следует, что *многочлен $F(\lambda)$ делится без остатка справа (слева) на бином $\lambda E - A$ тогда и только тогда, когда $F(A) = 0$ (соответственно $\hat{F}(A) = 0$).*

Пример. Пусть $A = \| a_{ik} \|^n_1$ и $f(\lambda)$ — многочлен относительно λ. Тогда

$$F(\lambda) = f(\lambda)E - f(A)$$

делится (слева и справа) без остатка на $\lambda E - A$. Это следует непосредственно из обобщенной теоремы Безу, поскольку в данном случае $F(A) = \hat{F} A) = 0$.

3.3. Характеристический многочлен матрицы. Присоединенная матрица

1. Рассмотрим матрицу $A = \| a_{ik} \|^n_1$. *Характеристической матрицей* для матрицы A называется матрица $\lambda E - A$. Определитель характеристической матрицы

$$\Delta(\lambda) = | \lambda E - A | = | \lambda\delta_{ik} - a_{ik} |^n_1$$

представляет собой скалярный многочлен относительно λ и называется *характеристическим многочленом* матрицы A (этот многочлен отличается множителем $(-1)^n$ от многочлена $\Delta(\lambda)$).

Матрицу $B(\lambda) = \| b_{ik}(\lambda) \|^n_1$, где $b_{ik}(\lambda)$ — алгебраическое дополнение элемента $\lambda\delta_{ik} - a_{ki}$ в определителе $\Delta(\lambda)$, мы будем называть *присоединенной матрицей* для матрицы A.

Так, например, для матрицы

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

будем иметь:

$$\lambda E - A = \begin{Vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{Vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \lambda^3 - (a_{11} + a_{22} + a_{33})\lambda^2 + \ldots,$$

$$B(\lambda) = \begin{Vmatrix} \lambda^2 - (a_{22} + a_{33})\lambda + a_{22}a_{33} - a_{23}a_{32} & * & * \\ a_{21}\lambda + a_{23}a_{31} - a_{21}a_{33} & * & * \\ a_{31}\lambda + a_{21}a_{32} - a_{22}a_{31} & * & * \end{Vmatrix}.$$

Из приведенных определений следуют тождества относительно λ:

$$(\lambda E - A)\, B(\lambda) = \Delta(\lambda)E, \qquad (3.26)$$

$$B(\lambda)\,(\lambda E - A) = \Delta(\lambda)E. \qquad (3.27)$$

Правые части этих равенств мы можем рассматривать как многочлены с матричными коэффициентами (каждый из этих коэффициентов равен произведению скаляра на единичную матрицу E). Многочленную матрицу $B(\lambda)$ можно также представить в виде многочлена, расположенного по степеням λ. Равенства (3.26) и (3.27) показывают, что $\Delta(\lambda)E$ делится слева и справа на $\lambda E — A$ без остатка. Согласно обобщенной теоремы Безу это возможно лишь тогда, когда остаток $(A)E=\Delta\ (A)$ равен нулю.

Теорема 2 (Гамильтона — Кэли). *Всякая квадратная матрица удовлетворяет своему характеристическому уравнению,* т.е.

$$\Delta(A) = 0. \qquad (3.28)$$

Пример.

$$A = \begin{Vmatrix} 2 & 1 \\ -1 & 3 \end{Vmatrix},$$

$$\Delta(\lambda) = \begin{Vmatrix} \lambda-2 & -1 \\ 1 & \lambda-3 \end{Vmatrix} = \lambda^2 - 5\lambda + 7,$$

$$\Delta(A) = A^2 - 5A + 7E = \begin{Vmatrix} 3 & 5 \\ -5 & 8 \end{Vmatrix} - 5\begin{Vmatrix} 2 & 1 \\ -1 & 3 \end{Vmatrix} + 7\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix} = 0.$$

2. Обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_n$ все характеристические числа матрицы A, т.е. все корни характеристического многочлена $\Delta(\lambda)$ (каждое из чисел λ_i повторяется в этом ряде столько раз, какова его кратность как корня многочлена $\Delta(\lambda)$). Тогда

$$\Delta(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_г). \qquad (3.29)$$

Пусть дано произвольный скалярный многочлен $g(\mu)$. Найдем характеристические числа матрицы $g(A)$. Для этого разложим $g(\mu)$ на линейные множители

$$g(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2) \ldots (\mu - \mu_l). \qquad (3.30)$$

Подставим в обе части этого тождества вместо μ матрицу A:

$$g(A) = a_0 (A - \mu_1 E)(A - \mu_2 E) \ldots (A - \mu_l E). \tag{3.31}$$

Переходя к определителям в обеих частях равенства (3.31) и используя равенства (3.29) и (3.30), получим:

$$|g(A)| = a_0^n |A - \mu_1 E| \, |A - \mu_2 E| \ldots |A - \mu_l E| =$$
$$= (-1)^{nl} a_0^n \Delta(\mu_1) \Delta(\mu_2) \ldots \Delta(\mu_l) =$$
$$= (-1)^{nl} a^n \prod_{i=1} \prod_{k=1}^{n} (\mu_i - \lambda_k) = g(\lambda_1) g(\lambda_2) \ldots g(\lambda_n).$$

Заменив в равенстве

$$|g(A)| = g(\lambda_1) g(\lambda_2) \ldots g(\lambda_n) \tag{3.32}$$

многочлен $g(\mu)$ на $\lambda - g(\mu)$, где λ - некоторый параметр, найдем:

$$|\lambda E - g(A)| = [\lambda - g(\lambda_1)] [\lambda - g(\lambda_2)] \ldots [\lambda - g(\lambda_n)]. \tag{3.33}$$

Из этого равенства вытекает следующая теорема.

Теорема 3. *Если* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *— все характеристические числа* (с *учетом кратностей) матрицы* A, *а* $g(\mu)$ *— некоторый скалярный многочлен, то* $g(\lambda_1)$, $g(\lambda_2)$, ..., $g(\lambda_n)$ *— все характеристические числа матрицы* $g(A)$.

В частности, если матрица A имеет характеристические числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, то матрица A^k имеет характеристические числа $\lambda^k_1, \lambda^k_2, \ldots, \lambda^k_n$ ($k = 0, 1, 2, 3, \ldots$).

3. Укажем эффективную формулу, которая выражает присоединенную матрицу $B(\lambda)$ через характеристический многочлен $\Delta(\lambda)$.

Пусть

$$\Delta(\lambda) = \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \ldots - p_n. \tag{3.34}$$

Разность $\Delta(\lambda) - \Delta(\mu)$ делится без остатка на $\lambda - \mu$. Поэтому

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^{n-1} + (\mu - p_1) \lambda^{n-2} + (\mu^2 - p_1 \mu - p_2) \lambda^{n-3} + \ldots \tag{3.35}$$

есть многочлен относительно λ и μ.

Тождество

$$\Delta(\lambda) - \Delta(\mu) = \delta(\lambda, \mu)(\lambda - \mu) \qquad (3.36)$$

не нарушится, если в него вместо λ и μ подставить перестановочные между собой матрицы λE и A. Тогда, поскольку согласно теореме Гамильтона — Кэли $\Delta(A) = 0$,

$$\Delta(\lambda) E = \delta(\lambda E, A)(\lambda E - A). \qquad (3.37)$$

Сопоставляя между собой равенства (3.27) и (3.37), из однозначности частного получаем искомую формулу

$$B(\lambda) = \delta(\lambda E, A). \qquad (3.38)$$

Отсюда в силу (3.35)

$$B(\lambda) = \lambda^{n-1} + B_1\lambda^{n-2} + B_2\lambda^{n-3} + \ldots + B_{n-1}, \qquad (3.39)$$

где

$$B_1 = A - p_1E, \quad B_2 = A^2 - p_1A - p_2E, \ldots$$

и вообще

$$B_k = A^k - p_1A^{k-1} - p_2A^{k-2} - \ldots - p_kE \quad (k = 1, 2, \ldots, n-1). \qquad (3.40)$$

Матрицы $B_1, B_2, \ldots, B_{n-1}$ можно вычислять последовательно, исходя из рекуррентных соотношений

$$B_k = AB_{k-1} - p_kE \quad (k = 1, 2.....n - 1;\ B_0 = E). \qquad (3.41)$$

При этом

$$AB_{n-1} - p_nE = 0. \qquad (3.42)$$

Из (3.41) следуют равенства (3.40). Если в (3.42) подставить выражение для B_{n-1} из (3.40), то получим: $\Delta(A)=0$. Этот вывод теоремы Гамильтона-Кэли не опирается явным образом на обобщенную теорему Безу, но неявно содержит в себе эту теорему.

Соотношение (3.41) и (3.42) непосредственно получаются из тождества (3.26), если в обеих частях этого тождества приравнять между собой коэффициенты при одинаковых степенях λ.

Если A - неособенная матрица, то

$$p_n = (-1)^{n-1} \mid A \mid \neq 0$$

и из (35) следует:

$$A^{-1} = \frac{1}{p_n} B_{n-1}. \qquad (3.43)$$

Пусть λ_0 — характеристическое число матрицы A, т.е. $\Delta(\lambda_0)=0$. Подставив в (3.26) вместо λ значение λ_0, найдем:

$$(\lambda_0 E - A)B(\lambda_0) = 0. \qquad (3.44)$$

Предположим, что матрица $B(\lambda_0)\neq 0$, и обозначим через b любой ненулевой столбец этой матрицы. Тогда из (3.44) $(\lambda_0 E - A)b = 0$ или

$$Ab = \lambda_0 b. \tag{3.45}$$

Следовательно, любой ненулевой столбец матрицы $B(\lambda_0)$ определяет собственный вектор, соответствующий характеристическому числу λ_0. (Если характеристическому числу λ_0 соответствует d_0 линейно независимых собственных векторов ($n - d_0$ — ранг матрицы $\lambda_0 E - A$), то ранг матрицы $B(\lambda_0)$ не превосходит d_0. В частности, если числу λ_0 соответствует только одно собственное направление, то в матрице $B(\lambda_0)$ элементы любых двух столбцов пропорциональны).

Таким образом, *если коэффициенты характеристического многочлена известны, то присоединенная матрица может быть найдена по формуле* (3.38). *Если данная матрица A неособенная, то по формуле* (3.43) *находится обратная матрица* A^{-1}. *Если* λ_0— *характеристическое число матрицы A, то ненулевые столбцы матрицы B* (λ_0) *являются собственными векторами матрицы A для* $\lambda = \lambda_0$.

Пример.

$$A = \begin{Vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{Vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda-2 & 1 & -1 \\ 0 & \lambda-1 & -1 \\ 1 & -1 & \lambda-1 \end{vmatrix} = \lambda^3 - 4\lambda^2 + 5\lambda - 2,$$

$$\delta(\lambda,\mu) = \frac{\Delta(\lambda)-\Delta(\mu)}{\lambda-\mu} = \lambda^2 + \lambda(\mu-4) + \mu^2 - 4\mu + 5,$$

$$B(\lambda) = \delta(\lambda E, A) = \lambda^2 E + \lambda\underbrace{(A-4E)}_{B_1} + \underbrace{A^2 - 4A + 5E}_{B_2}.$$

Но

$$B_1 = A - 4E = \begin{Vmatrix} -2 & -1 & 1 \\ 0 & -3 & 1 \\ -1 & 1 & -3 \end{Vmatrix},$$

$$B_2 = AB_1 + 5E = \begin{Vmatrix} 0 & 2 & -2 \\ -1 & +3 & -2 \\ 1 & -1 & 2 \end{Vmatrix},$$

$$B(\lambda) = \begin{Vmatrix} \lambda^2 - 2\lambda & -\lambda + 2 & \lambda - 2 \\ -1 & \lambda^2 - 3\lambda + 3 & \lambda - 2 \\ -\lambda + 1 & \lambda - 1 & \lambda^2 - 3\lambda + 2 \end{Vmatrix},$$

$$|A| = 2, \quad A^{-1} = \frac{1}{2} B_2 = \begin{Vmatrix} 0 & 1 & -1 \\ -\frac{1}{2} & \frac{3}{2} & -1 \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{Vmatrix}.$$

Далее

$$\Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2).$$

Первый столбец матрицы $B(1)$ дает собственный вектор (1, 1, 0) для характеристического числа $\lambda = 1$.

Первый столбец матрицы $B(2)$ дает собственный вектор (0, 1, 1), соответстующийхарактеристическому числу $\lambda = 2$.

3.4. Метод Д. К. Фаддеева одновременного вычисления коэффициентов характеристического многочлена и присоединенной матрицы

Д. К. Фаддеев предложил метод одновременного определения скалярных коэффициентов $p_1, p_2, \ldots, p_n$ характеристического многочлена

$$\Delta(\lambda) = \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \ldots - p_n \qquad (3.46)$$

матричных коэффициентов $B_1, B_2, \ldots, B_{n-1}$ присоединенной матрицы $B(\lambda)$.

Для изложения метода Д. К. Фаддеева введем понятие о следе матрицы. Под *следом* матрицы $A = \|a_{ik}\|^n_1$ (обозначение: Sp A) понимают сумму диагональных элементов этой матрицы:

$$\operatorname{Sp} A = \sum_{i=1}^{n} a_{ii}.$$

(3.47)

Нетрудно видеть, что

$$\operatorname{Sp} A = p_1 = \sum_{i=1}^{n} \lambda_i,$$

(3.48)

если $\lambda_1, \lambda_2, ..., \lambda_n$ — характеристические числа матрицы A, т.е.

$$\Delta(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n). \qquad (3.49)$$

Так как, согласно теореме 3, степень матрицы A^k имеет своими характеристическими числами степени $\lambda^k_1, \lambda^k_2, ..., \lambda^k_n$ $(k = 0, 1, 2, 3, ...)$, то

$$\operatorname{Sp} A^k = s_k = \sum_{i=1}^{n} \lambda_i^k \quad (k = 0, 1, 2, \dots).$$

(3.50)

Суммы s_k $(k = 1, 2, ..., n)$ степеней корней многочлена (3.46) связаны с коэффициентами этого уравнения формулами Ньютона

$$kp_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1 \quad (k = 1, 2, \dots, n). \qquad (3.51)$$

Если вычислить следы $s_1, s_2, ..., s_n$ матриц $A, A^2, ..., A^n$, то потом можно из уравнений (3.51) последовательно определить коэффициенты $p_1, p_2, \dots, p_n$. В этом состоит метод Леверрье определения коэффициентов характеристического многочлена по следам степеней матрицы.

Д. К. Фаддеев предложил вместо следов степеней $A, A^2, ..., A^n$ вычислить последовательно следы некоторых других матриц $A_1, A_2, \dots, A_n$ и с их помощью определить $p_1, p_2, \dots, p_n$ и $B_1, B_2, ..., B_n$ следующими формулами:

$$\left.\begin{array}{lll}
A_1 = A & p_1 = \operatorname{Sp} A_1, & B_1 = A_1 - p_1 E, \\
A_2 = AB_1, & p_2 = \frac{1}{2}\operatorname{Sp} A_2, & B_2 = A_2 - p_2 E, \\
\dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\
A_{n-1} = AB_{n-2}, & p_{n-1} = \frac{1}{n-1}\operatorname{Sp} A_{n-1}, & B_{n-1} = A_{n-1} - p_{n-1} E, \\
A_n = AB_{n-1}, & p_n = \frac{1}{n}\operatorname{Sp} A_n, & B_n = A_n - p_n E = 0.
\end{array}\right\}$$

(3.52)

Последнее равенство $B_n = A_n - p_n E = 0$ может быть использовано для контроля вычислений.

Для того чтобы убедиться, что числа $p_1, p_2, \dots, p_n$ и матрицы $B_1, B_2, \dots, B_{n-1}$, последовательно определяемые по формулам (3.52), являются коэффициентами $\Delta(\lambda)$ и $B(\lambda)$, заметим, что из (3.52) следуют следующие формулы для A_k и B_k ($k = 1, 2, \dots, n$):

$$A_k = A^k - p_1 A^{k-1} - \dots - p_{k-1} A, \quad B_k = A^k - p_1 A^{k-1} - \dots - p_{k-1} A - p_k E. \tag{3.53}$$

Приравняем между собой следы левой и правой частей первой из этих формул; получим:

$$k p_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1.$$

Но эти формулы совпадают с формулами Ньютона (3.51), по которым последовательно определяются коэффициенты характеристического многочлена $\Delta(\lambda)$. Следовательно, числа $p_1, p_2, \dots, p_n$, определяемые по формулам (3.52), и являются коэффициентами $\Delta(\lambda)$. Но тогда вторые формулы (3.53) совпадают с формулами (3.40), по которым определяются матричные коэффициенты $B_1, B_2, \dots, B_{n-1}$ присоединенной матрицы $B(\lambda)$. Следовательно, формулы (3.52) определяют и коэффициенты $B_1, B_2, \dots, B_{n-1}$ матричного многочлена $B(\lambda)$.

Пример. Для контроля за вычислениями мы под каждой из матриц A_1, A_2, A_3 подписываем ее суммарную строку. Произведения суммарной строки первого сомножителя на столбцы второго должны дать элементы суммарной строки произведения.

$$\begin{array}{c} A=\left\|\begin{matrix} 2 & -1 & 1 & 2\\ 0 & 1 & 1 & 0\\ -1 & 1 & 1 & 1\\ 1 & 1 & 1 & 0 \end{matrix}\right\| \\ \qquad\;\; 2 \quad\; 2 \quad\; 4 \quad\; 3 \end{array}, \quad p_1=\mathrm{Sp}\,A=4, \quad B_1=A-4E=\left\|\begin{matrix} -2 & -1 & 1 & 2\\ 0 & -3 & 1 & 0\\ -1 & 1 & -3 & 1\\ 1 & 1 & 1 & -4 \end{matrix}\right\|,$$

$$\begin{array}{c} A_2=AB_1=\left\|\begin{matrix} -3 & 4 & 0 & -3\\ -1 & -2 & -2 & 1\\ 2 & 0 & -2 & -5\\ -3 & -3 & -1 & 3 \end{matrix}\right\| \\ \qquad\qquad -5 \; -1 \; -5 \; -4 \end{array}, \quad p_2=\frac{1}{2}\mathrm{Sp}\,A_2=-2, \quad B_2=A_2+2E=\left\|\begin{matrix} -1 & 4 & 0 & -3\\ -1 & 0 & -2 & 1\\ 2 & 0 & 0 & -5\\ -3 & -3 & -1 & 5 \end{matrix}\right\|;$$

$$\begin{array}{c} A_3=AB_2=\left\|\begin{matrix} -5 & 2 & 0 & -2\\ 1 & 0 & -2 & -4\\ -1 & -7 & -3 & 4\\ 0 & 4 & -2 & -7 \end{matrix}\right\| \\ \qquad\qquad -5 \; -1 \; -7 \; -9 \end{array}, \quad p_3=\frac{1}{3}\mathrm{Sp}\,A_3=-5, \quad B_3=A_3+5E=\left\|\begin{matrix} 0 & 2 & 0 & -2\\ 1 & 5 & -2 & -4\\ -1 & -7 & 2 & 4\\ 0 & 4 & -2 & -2 \end{matrix}\right\|;$$

$$A_4=AB_3=\left\|\begin{matrix} -2 & 0 & 0 & 0\\ 0 & -2 & 0 & 0\\ 0 & 0 & -2 & 0\\ 0 & 0 & 0 & -2 \end{matrix}\right\|, \quad p_4=-2,$$

$$\Delta(\lambda)=\lambda_4-4\lambda^3+2\lambda^2+5\lambda+2,$$

$$|A|=2, \quad A^{-1}=\frac{1}{p_4}B_3=\left\|\begin{matrix} 0 & -1 & 0 & 1\\ -\frac{1}{2} & -\frac{5}{2} & 1 & 2\\ \frac{1}{2} & \frac{7}{2} & -1 & -2\\ 0 & -2 & 1 & 1 \end{matrix}\right\|.$$

Замечание. Если мы хотим определить p_1, p_2, p_3, p_4 и только первые столбцы в B_1, B_2, B_3, то достаточно вычислить в A_2 элементы 1-го столбца и только диагональные элементы остальных столбцов, в A_3 — только элементы 1-го столбца, в A_4 — только два первых элемента 1-го столбца.

3.5. Минимальный многочлен матрицы

Определение 1. Скалярный многочлен $f(\lambda)$ называется *аннулирующим многочленом* квадратной матрицы A, если

$$f(A) = 0.$$

Аннулирующий многочлен $\psi(\lambda)$ наименьшей степени со старшим коэффициентом, равным единицы, называется *минимальным многочленом* матрицы A.

Согласно теореме Гамильтона-Кэли характеристический многочлен $\Delta(\lambda)$ матрицы A является аннулирующиим для этой матрицы. Однако, как будет показано ниже, в общем случае он не является минимальным.

Разделим произвольный аннулирующиий многочлен $f(\lambda)$ на минимальный:

$$f(\lambda)=\psi(\lambda)q(\lambda)+r(\lambda),$$

где степень $r(\lambda)$ меньше степени $\psi(\lambda)$. Отсюда имеем:

$$f(A)=\psi(A)q(A)+r(\lambda).$$

Поскольку $f(A)=0$ и $\psi(A)=0$, то, значит, и $r(A)=0$. Но степень $r(\lambda)$ меньше степени минимального многочлена $\psi(\lambda)$. Поэтому $r(\lambda)\equiv 0$. (В противном случае существовал бы аннулирующиий многочлен степень которого была бы меньше степени минимального многочлена). Таким образом, *произвольный аннулирующий многочлен матрицы всегда делится дез остатку на ее минимальный многочлен.*

Пусть два многочлена $\psi(\lambda)$ и $\psi(\lambda)$ являются минимальными для одной и той же матрицы. Тогда каждый из них делится на другой многочлен без остатка, т.е. эти многочлены отличаются постоянным множителем. Этот постоянный множитель равен единице, поскольку равны единице старшие коэффициенты в $\psi(\lambda)$ и $\psi(\lambda)$. Мы доказали *единственность минимального многочлена* для данной матрицы A.

Выведем формулу, которая связывает минимальный многочлен с характеристическим. Обозначим через $D_{n-1}(\lambda)$ наибольший общий делитель всех миноров $(n - 1)$-го порядка характеристической матрицы $\lambda E - A$, т.е. наибольший общий делитель всех элементов присоединенной матрицы $B(\lambda)=\|b_{ik}(\lambda)\|^n_1$ (см. предыдущий раздел); при этом старший коэффициент в $D_{n-1}(\lambda)$ берем равным единице. Тогда

$$B(\lambda)=D_{n-1}(\lambda)C(\lambda), \tag{3.54}$$

где $C(\lambda)$ — некоторая многочленная матрица, *«приведенная» присоединенная матрица* для $\lambda E - A$. Из (3.26) и (3.54) находим:

$$\Delta(\lambda)E=(\lambda E-A)C(\lambda)D_{n-1}(\lambda). \tag{3.55}$$

Отсюда следует, что $\Delta(\lambda)$ делится без остатка на $D_{n-1}(\lambda)$ (в этом можно и непосредственно убедиться, раскладывая характеристический определитель $\Delta(\lambda)$ по элементам какой-либо строки):

$$\frac{\Delta(\lambda)}{D_{n-1}(\lambda)}=\psi(\lambda), \tag{3.56}$$

где $\psi(\lambda)$ — некоторый многочлен. Обе части тождества (3.55) можно сократить на $D_{n-1}(\lambda)$:

$$\psi(\lambda)E=(\lambda E-A)C(\lambda). \tag{3.57}$$

(в данном случае наряду с (3.57) имеет место и тождество [см. (3.27)] $\psi(\lambda)E=C(\lambda)(\lambda E - A)$, т.е. $C(\lambda)$ есть одновременно и левым и правым частным от деления $\psi(\lambda)E$, на $\lambda E - A$).

Поскольку $\psi(\lambda)E$ делится без остатка слева на $\lambda E - A$, то в силу обобщенной теоремы Безу

$$\psi(A) = 0.$$

Таким образом, многочлен $\psi(\lambda)$, определенный формулой (3.56), является аннулирующиим многочленом для матрицы A. Докажем, что он является минимальным многочленом.

Обозначим минимальный многочлен через $\psi^*(\lambda)$. Тогда $\psi(\lambda)$ делится без остатка на $\psi^*(\lambda)$:

$$\psi(\lambda)=\psi^*(\lambda)\chi(\lambda). \quad (3.58)$$

Поскольку $\psi^*(A)=0$, то в силу обобщенной теоремы Безу матричный многочлен $\psi^*(\lambda)E$ делится слева без остатка на $\lambda E - A$:

$$\psi^*(\lambda)E=(\lambda E-A)C^*(\lambda). \quad (3.59)$$

Из (3.58) и (3.59) следует:

$$\psi(\lambda)E = (\lambda E-A)C^*(\lambda)\chi(\lambda). \quad (3.60)$$

Тождества (3.57) и (3.60) показывают, что как $C(\lambda)$, так и $C^*(\lambda)\chi(\lambda)$ являются левыми частными при делении $\psi(\lambda)E$ на $\lambda E - A$. В силу однозначности деления

$$C(\lambda) = C^*(\lambda)\chi(\lambda).$$

Отсюда следует, что $\chi(\lambda)$ является общим делителем всех элементов многочленной матрицы $C(\lambda)$. Но, с другой стороны, наибольший общий делитель всех элементов приведенной присоединенной матрицы $C(\lambda)$ равен единице, поскольку эта матрица была получена из $B(\lambda)$ путем деления на $D_{n-1}(\lambda)$. Поэтому $\chi(\lambda) = \text{const}$. Так как старшие коэффициенты в $\psi(\lambda)$ и $\psi^*(\lambda)$ равны единице, то в (3.58) $\chi(\lambda) = 1$, т.е. $\psi(\lambda) = \psi^*(\lambda)$, что и требовалось доказать.

Мы установили следующую формулу для минимального многочлена:

$$\psi(\lambda) = \frac{\Delta(\lambda)}{D_{n-1}(\lambda)}. \quad (3.61)$$

Для приведенной присоединенной матрицы $C(\lambda)$ имеем формулу, аналогичную формуле (3.38):

$$C(\lambda) = \Psi(\lambda E, A), \quad (3.62)$$

где многочлен $\Psi(\lambda, \mu)$ определяется равенством

$$\Psi(\lambda, \mu) = \frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu} \quad (3.63)$$

(Формула (3.62) выводится совсем так же, как и формула (3.38). В обе части тождества $\psi(\lambda) - \psi(\mu) = (\lambda - \mu)\Psi(\lambda, \mu)$ вместо λ и μ подставляются матрицы λE и A и полученное матричное равенство сопоставляется с (3.57)).

Кроме того,

$$(\lambda E-A)C(\lambda)= \psi(\lambda)E. \tag{3.64}$$

Переходя к определителям в обеих частях равенства (3.64), получаем:

$$\Delta(\lambda)\,|\,C(\lambda)\,| = [\psi(\lambda)]^n. \tag{3.65}$$

Таким образом, $\Delta(\lambda)$ делится без остатка на $\psi(\lambda)$, а некоторая степень $\psi(\lambda)$ делится без остатка на $\Delta(\lambda)$, т.е. совокупность всех раздичных между собой корней у многочлена $\Delta(\lambda)$ и $\psi(\lambda)$ одна та же. Иначе говоря, *корнями* $\psi(\lambda)$ *служат все различные между собой характеристические числа матрицы A.*

Если

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\dots(\lambda - \lambda_s)^{n_s}$$

$$(\lambda_i \neq \lambda_j \ \text{при} \ i \neq j; \ \ n_i > 0, \ \ i, \ \ j = 1, 2, \dots, s), \tag{3.66}$$

то

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\dots(\lambda - \lambda_s)^{m_s}, \tag{3.67}$$

где

$$0<m_k\leq n_k \quad (k= 1,2, ..., s). \tag{3.68}$$

Отметим еще одно свойство матрицы $C(\lambda)$. Пусть λ_0 — какое-нибудь характеристическое число матрицы $A = \| a_{ik} \|^n_1$. Тогда $\psi(\lambda_0)=0$, и потому согласно (3.64)

$$(\lambda_0 E-A)C(\lambda_0) = 0. \tag{3.69}$$

Заметим, что всегда $C(\lambda_0)\neq 0$. Действительно, в противном случае все элементы приведенной присоединенной матрицы $C(\lambda)$ делились бы без остатка на $\lambda - \lambda_0$, что невозможно.

Обозначим через c любой ненулевой столбец матрицы $C(\lambda_0)$. Тогда из (3.69)

$$(\lambda_0 E-A)c = 0,$$

т.е.

$$Ac = \lambda_0 c. \tag{3.70}$$

Иначе говоря, любой ненулевой столбец матрицы $C(\lambda_0)$ (а такой столбец всегда имеется) определяет собственный вектор для $\lambda = \lambda_0$.

Пример.

$$A = \begin{Vmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{Vmatrix},$$

$$\Delta(\lambda) = \begin{vmatrix} \lambda-3 & 3 & -2 \\ 1 & \lambda-5 & 2 \\ 1 & -3 & \lambda \end{vmatrix} = \lambda^3 - 8\lambda^2 + 20\lambda - 16 = (\lambda-2)^2(\lambda-4),$$

$$\delta(\lambda, \mu) = \frac{\Delta(\mu) - \Delta(\lambda)}{\mu - \lambda} = \mu^2 + \mu(\lambda - 8) + \lambda^2 - 8\lambda + 20,$$

$$B(\lambda) = A^2 + (\lambda - 8)A + (\lambda^2 - 8\lambda + 20)E =$$

$$= \begin{Vmatrix} 10 & -18 & 12 \\ -6 & 22 & -12 \\ -6 & 18 & -8 \end{Vmatrix} + (\lambda - 8)\begin{Vmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{Vmatrix} + (\lambda^2 - 8\lambda + 20)\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix} =$$

$$= \begin{Vmatrix} \lambda^2 - 5\lambda + 6 & -3\lambda + 6 & 2\lambda - 4 \\ -\lambda + 2 & \lambda^2 - 3\lambda + 2 & -2\lambda + 4 \\ -\lambda + 2 & 3\lambda - 6 & \lambda^2 - 8\lambda + 12 \end{Vmatrix}.$$

Все элементы матрицы $B(\lambda)$ делятся на $D_2(\lambda) = \lambda - 2$. Сокращая этот множитель, получим:

$$C(\lambda) = \begin{Vmatrix} \lambda-3 & -3 & 2 \\ -1 & \lambda-1 & -2 \\ -1 & 3 & \lambda-6 \end{Vmatrix}$$

и

$$\psi(\lambda) = \frac{\Delta(\lambda)}{\lambda - 2} = (\lambda - 2)(\lambda - 4).$$

Подставим в $C(\lambda)$ вместо λ значение $\lambda_0 = 2$:

$$C(2) = \begin{Vmatrix} -1 & -3 & 2 \\ -1 & 1 & -2 \\ -1 & 3 & -4 \end{Vmatrix}.$$

Первый столбец дает нам собственный вектор (1, 1, 1) для $\lambda_0 = 2$. Второй столбец дает нам собственный вектор (—3, 1, 3) для того же характеристического числа $\lambda_0 = 2$. Третий столбец есть линейная комбинация первых двух.

Точно так же, полагая $\lambda = 4$, из первого столбца матрицы $C(4)$ найдем собственный вектор (1, —1, —1), отвечающий характеристическому числу $\lambda_0 = 4$.

Обращаем внимание читателя на то, что $\psi(\lambda)$ и $C(\lambda)$ можно было бы определить иначе.

Находим сначала $D_2(\lambda)$. $D_2(\lambda)$ может иметь своими корнями только числа 2 и 4. При $\lambda = 4$ в $\Delta(\lambda)$ минор второго порядка

$$\begin{vmatrix} 1 & \lambda-5 \\ 1 & -3 \end{vmatrix} = -\lambda + 2$$

не обращается в нуль. Поэтому $D_2(4) \neq 0$. При $\lambda = 2$ столбцы матрицы A становятся пропорциональными. Поэтому все миноры второго порядка в $\Delta(\lambda)$ при $\lambda = 2$ равны нулю: $D_2(2) = 0$. Так как вычисленный минор имеет первую степень, то $D_2(\lambda)$ не может делиться на $(\lambda - 2)^2$. Следовательно,

$$D_2(\lambda) = \lambda - 2.$$

Отсюда

$$\psi(\lambda) = \frac{\Delta(\lambda)}{\lambda - 2} = (\lambda - 2)(\lambda - 4) = \lambda^2 - 6\lambda + 8,$$

$$\psi(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu + \lambda - 6,$$

$$C(\lambda) = \psi(\lambda E, A) = A + (\lambda - 6)E = \left\| \begin{matrix} \lambda - 3 & -3 & 2 \\ -1 & \lambda - 1 & -2 \\ -1 & 3 & \lambda - 6 \end{matrix} \right\|.$$

Микромодуль 6
Индивидуальные тестовые задачи

1. Сравните понятие скалярного многочлена, многочлена от матрицы, матричного многочлена и многочленной матрицы. Приведите примеры.
2. Дана матрица:

$$A = \begin{bmatrix} -7 & 3 & 6 & 6 \\ -4 & 2 & 4 & 4 \\ -3 & 5 & 0 & -1 \\ -2 & -2 & 4 & 5 \end{bmatrix}.$$

а) Найдите присоединенную матрицу $G(\lambda)=\mathrm{Adj}(\lambda E - A)$ и характеристический многочлен $\Delta(\lambda)=\det(\lambda E - A)$ через алгебраические дополнения характеристической матрицы $F(\lambda)= \lambda E - A$.

б) Определите $G(\lambda)$ и $\Delta(\lambda)$ с помощью алгоритма Фаддеева и сравните с полученным ранее результатом.

в) Проверьте соотношение $F(\lambda)\, G(\lambda) = \Delta(\lambda)E$.

г) Запишите $G(\lambda)$ в виде матричного многочлена.

3. Проверьте справедливость теоремы Гамильтона- Кэли на примере матрицы

$$A = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}.$$

4. Покажите, что с помощью теоремы Гамильтона-Кэли можно представить многочлен $N(A)$ любой степени $m > n$ от квадратной матрицы A n-го порядка через многочлен от A, степень которого не превышает $n - 1$. Определите $N(A) = A^4+A^3+A^2+A+E$ для матрицы A из предыдущей задачи

а) прямым вычислением;

б) с помощью понижения порядка многочлена.

5. Используя теорему Гамильтона-Кэли, получите формулу обращения неособенной квадратной матрицы A n-го порядка

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 E),$$

где a_i ($i = 0, 1, ..., n-1$) — коэффициенты характеристического многочлена $\Delta(\lambda)=\det(\lambda E - A)$. Найдите с помощью этой формулы обратную для матрицы

$$A = \begin{bmatrix} 3 & -2 & 0 & -1 \\ 0 & 2 & 2 & 1 \\ 1 & -2 & -3 & -2 \\ 0 & 1 & 2 & 1 \end{bmatrix}$$

и проверьте результат каким-либо другим способом.

Микромодуль 7
Дифференциальные уравнения

3.6. Однородная система дифференциальных уравнений

1. Представление в матричной форме. Теория матриц оказалась эффективным средством исследования и решения дифференциальных уравнений. Среди них наиболее простыми являются линейные *дифференциальные уравнения с постоянными коэффициентами,* к которым приводятся много задач физики и техники. Здесь рассматриваются только такие уравнения и для краткости будем называть их просто дифференциальными уравнениями. Дифференциальное уравнение первого порядка, относительно неизвестной функции $x= x(t)$ имеет вид

$$\frac{dx(t)}{dt} = ax(t) + f(t),$$

где a — постоянный коэффициент; $f(t)$ — непрерывная функция времени, определенная на некотором интервале $t_1 < t < t_2$.

Решением уравнения является функция $x(t)$, подстановка которой в это уравнение обращает его в тождество. При $f(t)=0$ уравнение называется *однородным* и его общее решение выражается как $x = ce^{at}$, где c — произвольная стала. Общее решение исходного *неоднородного* уравнения $(f(t) \neq 0)$ выражается формулой

$$x = \left(c + \int_{t_0}^{t} e^{-a\tau} v(\tau)\, d\tau\right) e^{at} = ce^{at} + \int_{t_0}^{t} e^{a(t-\tau)} f(\tau)\, d\tau.$$

Это решение представляет собой сумму общего решения однородного и частного решения неоднородного дифференциальных уравнений. Оно удовлетворяет начальному условию $x(0) = x_0$ при $c = x_0$, т.е.

$$x = x_0 e^{at} + \int_{0}^{t} e^{a(t-\tau)} f(\tau)\, d\tau.$$

Переходя к системам дифференциальных уравнений, рассмотрим их представление в *нормальной форме:*

$$\left.\begin{aligned}\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + f_1(t)\\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + f_2(t)\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + f_n(t)\end{aligned}\right\}$$

к которой, как известно, можно привести любую систему линейных дифференциальных уравнений. В матричной записи эта система представляется одним уравнениям

$$\frac{dx}{dt} = Ax + f,$$

где x — вектор (столбец) неизвестных функций $x_1 = x_1(t)$; $x_2 = x_2(t)$; ...; $x_n = x_n(t)$; f — вектор (столбец) задаючих функций $f_1 = f_1(t)$; $f_2 = f_2(t); \ldots; f_n = f_n(t)$ и A - квадратная матрица постоянных *коэффициентов* a_{ij} $(i, j = 1, 2, \ldots, n)$:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}; \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}; \quad f = \begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_n \end{bmatrix}.$$

Задачу об отыскании решения системы дифференциальных уравнений, которые удовлетворяют заданным начальным значением скаляра $t=0$ и вектора $x_0 = x(0)$, называют *задачей Коши*. По аналогии с дифференциальным уравнением первого порядка можно записать искомое решение для вектора неизвестных функций в виде:

$$x = e^{At}x_0 + \int_0^t e^{A(t-\tau)} f(\tau)\,d\tau$$

Необходимо установить допустимость такого представления решения, а также выяснить смысл и способы определения входящей в него матрицы e^{At}.

2. Решение нормальной однородной системы. В матричной форме нормальная однородная система дифференциальных уравнений ($f = 0$) имеет вид:

$$\frac{dx}{dt} = Ax$$

Будем искать ее решение в виде $x = he^{\lambda t}$, где h — вектор (столбец) произвольных постоянных. Подставляя x в исходное уравнение, получаем $h\lambda e^{\lambda t} = Ahe^{\lambda t}$ или после сокращения на скаляр $e^{\lambda t}$ и перенесения Ah в левую часть равенства:

$$(\lambda E - A)h = 0.$$

Заметим, что сокращать на вектор h нельзя, так как операция деления на вектор в общем случае не имеет смысла. Вынося за скобки вектор h, необходимо умножить предварительно $h\lambda = \lambda h$ на единичную матрицу E.

Уравнение $(\lambda E — A)h=0$ имеет нетривиальные решения при условии, которое определитель матрицы $(\lambda E — A)$ обращается в нуль, т.е. $|\lambda E — A| = 0$ или

$$\Delta(\lambda) = \begin{vmatrix} \lambda - a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \lambda - a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix} = 0.$$

Так как порядок матрицы A равен n, то $\Delta(\lambda)$ является многочленом n-й степени относительно λ, т.е. $\Delta(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$. Корни уравнения $\Delta(\lambda) = 0$ (нули многочлена $\Delta(\lambda)$), число которых в соответствии с основной теоремой алгебры равно n, дадут значения $\lambda_1, \lambda_2, \ldots, \lambda_n$, при которых исходная система имеет нетривиальные решения.

Рассмотрим наиболее простой случай, когда всех корни уравнения $\Delta(\lambda) = 0$ простые (попарно различные). Тогда при $\lambda = \lambda_i$ имеем однородное уравнение $(\lambda_i — A)\, h^{(i)} = 0$, из которого можно определить вектор $h^{(i)}$. Таким образом, решение нормальной системы дифференциальных уравнений, соответствующее корню λ_i, будет $x^{(i)} = h^{(i)}e^{\lambda_i t}$. Всего получим n таких решений, соответствующих n корням $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Для любой квадратной матрицы A по установившейся терминологии $(\lambda E — A)$ называется *характеристической матрицей,* а $\Delta(\lambda) = |\lambda E — A| = 0$ — *характеристическим уравнением.* Корни $\lambda_1, \lambda_2, \ldots, \lambda_n$ уравнение $\Delta(\lambda)=0$ называются *собственными значениями (характеристическими числами),* а векторы $h^{(1)}, h^{(2)}, \ldots, h^{(n)}$ — *собственными векторами* матрицы A.

3. Пространство решений однородной системы. Можно показать, что множество всех решений однородной системы дифференциальных уравнений образует n-мерное линейное пространство.

Действительно, если $x^{(i)}$ и $x^{(j)}$ — какие-либо решения системы, то их сумма $x^{(i)}+x^{(j)}$ также будет решением, что вытекает из следующего:

$$\frac{d}{dt}(x^{(i)}+x^{(j)})=\frac{dx^{(i)}}{dt}+\frac{dx^{(j)}}{dt}=Ax^{(i)}+Ax^{(j)}=A(x^{(i)}+x^{(j)}).$$

Если $x^{(i)}$ — решение системы, то его произведение на число α, т.е. $\alpha x^{(i)}$, также будет решением, что вытекает из соотношений:

$$\frac{d}{dt}(\alpha x^{(i)})=\alpha\frac{dx^{(i)}}{dt}=\alpha Ax^{(i)}=A(\alpha x^{(i)}).$$

Так как для суммы матриц и произведения матрицы на число выполняются все аксиомы линейного просторную, то с учетом полученных результатов следует, что множество всевозможных решений системы дифференциальных уравнений образует линейное пространство. Если все собственные значения матрицы A системы уравнений различны, то в качестве базиса этого пространства можно принять n решений $x^{(i)}=h^{(i)}e^{\lambda_i t}$ $(i = 1, 2, ..., n)$. Тогда общее решение имеет такой вид:

$$x = c_1x^{(1)} + c_2x^{(2)} +\ldots+c_nx^{(n)} = \sum_{i=1}^{n} c_i\, h^{(i)}e^{\lambda_i t}\ .$$

4. Матричная запись решения однородной системы. Представим полученное выражение в матричной форме. Рассматривая векторы $x^{(i)}$ как столбцы матрицы X, а c_i — как элементы столбца произвольных постоянных c, запишем:

$$x=[x^{(1)},\ x^{(2)},\ \ldots,\ x^{(n)}]\begin{bmatrix} c_1 \\ c_2 \\ \ldots \\ c_n \end{bmatrix}=Xc.$$

В свою очередь, матрица X выражается следующим образом:

$$X=[h^{(1)}e^{\lambda_1 t},\ h^{(2)}e^{\lambda_2 t},\ \ldots,\ h^{(n)}e^{\lambda_n t}]=$$

$$=[h^{(1)},\ h^{(2)},\ \ldots,\ h^{(n)}]\cdot\begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_n t} \end{bmatrix}=H\varphi(t).$$

Здесь через H обозначена матрица n-го порядка, состоящая из столбцов $h^{(i)}$, а элементами диагональной матрицы $\varphi(t)$ являются экспонентные функции $e^{\lambda_i t}$ $(i=1, 2, ..., n)$. Итак, решение нормальной однородной системы линейных дифференциальных уравнений представляется в виде:

$$x = H\,\varphi(t)\,c.$$

При $t = 0$ матрица $\varphi(t)$ равна единичной матрице, ведь начальное условие $x_0 = Hc$, откуда $c=H^{-1}x_0$. Подставляя это значение c в общее решение, получаем

$$x = H\varphi(t)\,H^{-1}x_0 = \Phi(t)\,x_0.$$

Матрица n-го порядка

$$\Phi(t)=H\varphi(t)\,H^{-1}$$

называется *фундаментальной матрицей.* Ее вычисление сводится к определению собственных значений и собственных векторов матрицы A системы дифференциальных уравнений.

5. Определение фундаментальной матрицы. В вычислительной математике задача отыскания собственных значений и векторов ставится применительно к любой квадратной матрицк без непосредственной связи с дифференциальными уравнениями и называется *полной проблемой собственных значений.* Для ее решения разработано большое число различных алгоритмов, которые изложены в специальной литературе. Естественно, любой из них можно использовать для определения фундаментальной матрицы $\Phi(t)$. Ограничимся лишь иллюстрацией использования основных соотношений при вычислении $\Phi(t)$. Матрица H, называемая *модальной матрицей,* может быть получена как совокупность n столбцов $h^{(i)}$, которые являются решениями однородных уравнений $(\lambda_i E — A)\,h^{(i)} = 0$ (i=1, 2, ... n). Ранг матрицы $(\lambda_i E — A)$ при различных собственных значениях равен n—1. Действительно, так как определитель этой матрицы равен нулю, то она особенная и ее ранг не может быть большим, чем n—1. В то же время он не может быть и меньшим, чем n—1, так как при этом все миноры (n—1)-го порядка равнялись бы нулю, что означало бы кратность собственных значений. Таким образом, ранг $(\lambda_i E — A)$ точно равен (n—1), поэтому при вычислении столбцов модальной матрицы можно воспользоваться методом, изложенным в предыдущем микромодуле.

Рассмотрим в качестве примера однородную систему дифференциальных уравнений:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= 4x_1 - 8x_2 + x_3 \\ \frac{dx_2}{dt} &= 5x_1 - 9x_2 + x_3 \\ \frac{dx_3}{dt} &= 4x_1 - 6x_2 - x_3 \end{aligned}\right\}.$$

Для этой системы

$$A=\begin{bmatrix}4 & -8 & 1\\ 5 & -9 & 1\\ 4 & -6 & -1\end{bmatrix};\quad \lambda E-A=\begin{bmatrix}\lambda-4 & 8 & -1\\ -5 & \lambda+9 & -1\\ -4 & 6 & \lambda+1\end{bmatrix}.$$

Поскольку для вычисления h^i необходимы алгебраические дополнения какой-либо строки матрицы $(\lambda E - A)$, то определитель этой матрицы удобно получать разложением по элементам той же строки.

Алгебраические дополнения элементов первой строки:

$$\Delta_{11}(\lambda)=\begin{bmatrix}\lambda+9 & -1\\ 6 & \lambda+1\end{bmatrix}=\lambda^2+10\lambda+15;$$

$$\Delta_{12}(\lambda)=-\begin{vmatrix}-5 & -1\\ -4 & \lambda+1\end{vmatrix}=5\lambda+9;\quad \Delta_{13}(\lambda)=\begin{vmatrix}-5 & \lambda+9\\ -4 & 6\end{vmatrix}=$$
$$=4\lambda+6.$$

Характеристический многочлен и собственные значения:

$$\Delta(\lambda)=(\lambda-4)(\lambda^2+10\lambda+15)+8(5\lambda+9)-1(4\lambda+6)=$$
$$=\lambda^3+6\lambda^2+11\lambda+6=(\lambda+1)(\lambda+2)(\lambda+3);$$
$$\lambda_1=-1;\quad \lambda_2=-2;\quad \lambda_3=-3.$$

Собственные векторы:

$$h^{(i)}=k_i\begin{bmatrix}\Delta_{11}(\lambda_i)\\ \Delta_{12}(\lambda_i)\\ \Delta_{13}(\lambda_i)\end{bmatrix};\quad h^{(1)}=k_1\begin{bmatrix}6\\ 4\\ 2\end{bmatrix};\quad h^{(2)}=k_2\begin{bmatrix}-1\\ -1\\ -2\end{bmatrix};$$

$$h^{(3)}=k_3\begin{bmatrix}-6\\ -6\\ -6\end{bmatrix}.$$

Принимая $k_1=-(1/2)$; $k_2=-1$ и $k_3=-(1/6)$ (эти значения произвольны и выбираются по соображением удобства), получаем модальную матрицу, а также обратную к ней:

$$H=\begin{bmatrix}3 & 1 & 1\\ 2 & 1 & 1\\ 1 & 2 & 1\end{bmatrix};\quad H^{-1}=\begin{bmatrix}1 & -1 & 0\\ 1 & -2 & 1\\ -3 & 5 & -1\end{bmatrix}.$$

Фундаментальная матрица

$$\Phi(t) = H\varphi(t)H^{-1} =$$

$$= \begin{bmatrix} 3 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 & 0 \\ 0 & e^{-2t} & 0 \\ 0 & 0 & e^{-3t} \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 1 & -2 & 1 \\ -3 & 5 & -1 \end{bmatrix},$$

что после перемножения матриц приводит к следующему результату:

$$\Phi(t) = \begin{bmatrix} 3e^{-t} + e^{-2t} - 3e^{-3t} & -3e^{-t} - 2e^{-2t} + 5e^{-3t} & e^{-2t} - e^{-3t} \\ 2e^{-t} + e^{-2t} - 3e^{-3t} & -2e^{-t} - 2e^{-2t} + 5e^{-3t} & e^{-2t} - e^{-3t} \\ e^{-t} + 2e^{-2t} - 3e^{-3t} & -e^{-t} - 4e^{-2t} + 5e^{-3t} & 2e^{-2t} - e^{-3t} \end{bmatrix}.$$

Таким образом, согласно соотношению $x=\Phi(t)x_0$ общее решение рассматриваемой однородной системы дифференциальных уравнений:

$$x_1 = (3e^{-t} + e^{-2t} - 3e^{-3t})\,x_{10} + (-3e^{-t} - 2e^{-2t} + 5e^{-3t})\,x_{20} + \\ + (e^{-2t} - e^{-3t})\,x_{30};$$

$$x_2 = (2e^{-t} + e^{-2t} - 3e^{-3t})\,x_{10} + (-2e^{-t} - 2e^{-2t} + 5e^{-3t})\,x_{20} + \\ + (e^{-2t} - e^{-3t})\,x_{30};$$

$$x_3 = (e^{-t} + 2e^{-2t} - 3e^{-3t})\,x_{10} + (-e^{-t} - 4e^{-2t} + 5e^{-3t})\,x_{20} + \\ + (2e^{-2t} - e^{-3t})\,x_{30},$$

где x_{10}, x_{20}, x_{30} — элементы вектора x_0, равные начальным значениям соответствующих переменных при $t=0$.

6. Экспонентная функция от матрицы. Воспользовавшись разложением в степенной ряд экспонентной функции от скалярной переменной

$$e^{\lambda t} = 1 + \frac{\lambda t}{1!} + \frac{(\lambda t)^2}{2!} + \cdots = \sum_{s=0}^{\infty} \frac{(\lambda t)^s}{s!},$$

представим функциональную матрицу $\varphi(t)$ в виде:

$$\varphi(t) = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{\infty} \frac{(\lambda_1 t)^s}{s!} & & \\ & \ddots & \\ & & \sum_{s=0}^{\infty} \frac{(\lambda_n t)^s}{s!} \end{bmatrix} =$$

$$= \sum_{s=0}^{\infty} \frac{t^s}{s!} \begin{bmatrix} \lambda_1^s & & \\ & \ddots & \\ & & \lambda_n^s \end{bmatrix}.$$

Диагональную матрицу под знаком суммы можно рассматривать как результат возведения в *s*-ю степень диагональной матрицы Λ, элементами которой являются собственные числа матрицы *A,* т.е.

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}; \quad \Lambda^s = \begin{bmatrix} \lambda_1^s & & \\ & \ddots & \\ & & \lambda_n^s \end{bmatrix}.$$

Таким образом, по аналогии со скалярным случаем можно записать:

$$\varphi(t) = \sum_{s=0}^{\infty} \frac{\Lambda^s t^s}{s!} = e^{\Lambda t} = \exp(\Lambda t),$$

т.е. матрица $\varphi(t)$ представляет собой экспонентную функцию от матрицы Λt.

Выясним характер фундаментальной матрицы $\Phi(t)$. Подставляя решение в однородное дифференциальное уравнение $dx/dt = Ax$, получаем тождество:

$$\frac{d}{dt}[\Phi(t)x_0] = A\Phi(t)x_0; \quad \frac{d\Phi(t)}{dt}x_0 = A\Phi(t)x_0.$$

Так как в этих тождествах x_0 — вектор начальных значений не зависящий от времени, то $\Phi(t) = A\Phi(t)$, т.е. $\Phi(t)$ — это такая матрица, производная которой по времени равна произведению матрицы A на саму матрицу. Аналогичными свойствами обладает единственная скалярная функция $\exp(at)$, поэтому интуиция подсказывает соотношение:

$$\Phi(t) = e^{At}; \quad \frac{d\Phi(t)}{dt} = \frac{d}{dt}e^{At} = Ae^{At} = A\Phi(t).$$

Можно привести и более строгие соображения в пользу этих соотношений. Подставив $\Phi(t)=H\varphi(t)H^{-1}$ в тождество $\Phi(t)=A\Phi(t)$, имеем:

$$\frac{\partial}{dt}(H\varphi(t)H^{-1}) = A(H\varphi(t)H^{-1}); \quad H\frac{\partial\varphi(t)}{dt}H^{-1} = AH\varphi(t)H^{-1}.$$

Умножая слева на H^{-1} и справа на H, получаем:

$$\frac{\partial\varphi(t)}{dt} = H^{-1}AH\varphi(t); \quad \Lambda\varphi(t) = (H^{-1}AH)\varphi(t).$$

Здесь дифференцирование диагональной матрицы выполнено по формуле

$$\dot{\varphi}(t) = \frac{d}{dt}\exp(\Lambda t) = \Lambda \exp(\Lambda t),$$

справедливость которой вытекает из следующего:

$$\frac{d\varphi(t)}{dt} = \frac{d}{dt}\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} = \begin{bmatrix} \frac{d}{dt}e^{\lambda_1 t} & & \\ & \ddots & \\ & & \frac{d}{dt}e^{\lambda_n t} \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 e^{\lambda_1 t} & & \\ & \ddots & \\ & & \lambda_n e^{\lambda_n t} \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} = \Lambda\varphi(t) = \Lambda e^{\Lambda t}.$$

В результате получаем важные соотношения, которые устанавливают связь между матрицами Λ и A с помощью модальной матрицы H:

$$\Lambda = H^{-1}AH; \quad A = H\Lambda H^{-1}.$$

Как видно, любая квадратная матрица A, все собственные значения которой различны, преобразуется в диагональную матрицу, элементами которой являются эти собственные значения. Это частный случай *преобразования подобия,* которое играет большую роль в теории матриц и ее приложениях. Воспользуемся полученными соотношениями для представления $\Phi(t)$ в экспонентной форме. Так как $\Phi(t) = H\varphi(t)H^{-1}$, то

$$\Phi(t) = H\sum_{s=0}^{\infty}\frac{\Lambda^s t^s}{s!}H^{-1} = \sum_{s=0}^{\infty}\frac{t^s}{s!}(H\Lambda^s H^{-1}).$$

Но $\Lambda^s = \Lambda\Lambda^{s-1} = (H^{-1}AH)\Lambda\Lambda^{s-2} = (H^{-1}AH)(H^{-1}AH)\ldots$
$(H^{-1}AH) = H^{-1}A(HH^{-1})A(HH^{-1})A\ldots AH = H^{-1}A^sH.$

Следовательно,

$$\Phi(t) = \sum_{s=0}^{\infty}\frac{A^s t^s}{s!} = e^{At}.$$

Вообще, экспоненциальную функцию от любой квадратной матрицы *X* можно представить в виде сходящегося ряда:

$$e^X = 1 + X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots = \sum_{s=0}^{\infty} \frac{X^s}{s!}.$$

Обратной к матрице e^X является функциональная матрица

$$(e^X)^{-1} = e^{-X} = 1 - X + \frac{X^2}{2!} - \frac{X^3}{3!} + \cdots = \sum_{s=0}^{\infty} (-1)^s \frac{X^s}{s!}.$$

Дифференцирование экспоненциальной функции от матрицы выполняется по обычному правилу

$$\frac{d}{dt} e^{At} = Ae^{At}.$$

Следует иметь в виду, что вообще $e^X e^Y \neq e^Y e^X$, причем соотношение $e^X e^Y = e^{X+Y}$ имеет смысл только в случаях, когда *X* и *Y* — перестановочные матрицы. Через экспоненциальную функцию выражаются также другие функции от матриц:

$$\sin X = \frac{1}{2i}(e^{iX} - e^{-iX}); \quad \cos X = \frac{1}{2}(e^{iX} + e^{-iX});$$

$$\operatorname{sh} X = \frac{1}{2}(e^X - e^{-X}); \quad \operatorname{ch} X = \frac{1}{2}(e^X + e^{-X});$$

$$e^{iX} = \cos X + i \sin X; \quad e^{-iX} = \cos X - i \sin X.$$

Разложение в ряд также используется для вычисления Ф(t)= exp(At), так как доказано, что этот ряд сходится равномерно и абсолютно.

7. Преобразование подобия и замена переменных. Преобразование подобия $\Lambda = H^{-1}AH$ приводит матрицу *A* с различными собственными значениями к диагональной форме Λ (к диагональной форме приводятся и некоторые другие типы матриц, например, симметричные). Так как экспоненциальная функция от диагональной матрицы определяется без труда, то при вычислении экспоненциальной функции от матрицы удобно пользоваться соотношением:

$$\exp(At) = H \exp(\Lambda t) H^{-1}.$$

В то же время *H* можно рассматривать как матрицу преобразования переменных дифференциального уравнения, т.е. $x=Hy$, где y — новый вектор неизвестных функций. Подставив x в исходное уравнение, получим:

$$H\frac{dy}{dt} = AHy; \quad \frac{\partial y}{dt} = (H^{-1}AH)\,y; \quad \frac{dy}{dt} = \Lambda y,$$

т.е. замена переменных приводит к дифференциальному уравнению с диагональной матрицей Λ. Фундаментальная матрица при этом будет также диагональна и решение получаем в виде:

$$y = e^{\Lambda t}\, y_0;\; y_0 = H^{-1} x_0.$$

Преобразование $x = Hy$ можно рассматривать как переход к новому базису n-мерного пространства с *нормальными координатами.* Значение этого преобразования сотоит в том, что исходная система уравнений «развязывается» относительно новых переменных. Действительно, представив преобразованное уравнение $dy/dt = \Lambda y$ в развернутом виде

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} \lambda_1 y_1 & & & \\ & \lambda_2 y_2 & & \\ & & \ddots & \\ & & & \lambda_n y_n \end{bmatrix},$$

найдем:

$$\frac{dy_1}{dt} = \lambda_1 y_1;\quad \frac{dy_2}{dt} = \lambda_2 y_2;\quad \dots;\quad \frac{dy_n}{dt} = \lambda_n y_n.$$

Соответственно и в решении относительно преобразованных координат собственные значения не связаны между собой:

$$y_1 = y_{10} e^{\lambda_1 t};\; y_2 = y_{20} e^{\lambda_2 t};\dots;\; y_n = y_{n0} e^{\lambda_n t}.$$

Такая форма представляет существенные удобства при анализе линейных дифференциальных систем. Так, для примера из (п. 5) имеем:

$$y_1 = y_{10} e^{-t};\; y_2 = y_{20} e^{-2t};\; y_3 = y_{30} e^{-3t},$$

где начальные значения

$$y_0 = \begin{bmatrix} y_{10} \\ y_{20} \\ y_{30} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -2 & 1 \\ -3 & 5 & -1 \end{bmatrix}\begin{bmatrix} x_{10} \\ x_{20} \\ x_{30} \end{bmatrix} = \begin{bmatrix} x_{10} - x_{20} \\ x_{10} - 2x_{20} + x_{30} \\ -3x_{10} + 5x_{20} - x_{30} \end{bmatrix}.$$

Переходим к исходным функциям:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3y_1 + y_2 + y_3 \\ 2y_1 + y_2 + y_3 \\ y_1 + 2y_2 + y_3 \end{bmatrix}.$$

$$x_1 = (x_{10} - x_{20})\, 3e^{-t} + (x_{10} - 2x_{20} + x_{30})\, e^{-2t} + (-3x_{10} + 5x_{20} - x_{30})\, e^{-3t},$$

$$x_2 = (x_{10} - x_{20})\, 2e^{-t} + (x_{10} - 2x_{20} + x_{30})\, e^{-2t} + (-3x_{10} + 5x_{20} - x_{30})\, e^{-3t};$$

$$x_3 = (x_{10} - x_{20})\, e^{-t} + (x_{10} - 2x_{20} + x_{30})\, 2e^{-2t} + (-3x_{10} + 5x_{20} - x_{30})\, e^{-3t}.$$

Эти выражения можно получить из решения в (п. 5) группированием слагаемых относительно экспонент.

3.7. Неоднородная система дифференциальных уравнений

8. Неоднородная система уравнений. Будем искать решение неоднородной системы дифференциальных уравнений в матричной форме $\dot{x}(t) = Ax(t) + f(t)$ в виде $x(t) = \exp(At)\xi(t)$, где $\xi(t)$ — векторная функция времени, подлежащая определению. Подставляя выражение для $x(t)$ и ее производной в исходное уравнение, имеем:

$$Ae^{At}\xi(t) + e^{At}\frac{d\xi(t)}{dt} = Ae^{At}\xi(t) + f(t)$$

или после очевидных упрощений

$$Ae^{At}\xi(t) + e^{At}\frac{d\xi(t)}{dt} = Ae^{At}\xi(t) + f(t)$$

При начальных условиях $x(0)=x_0$ начальное значение искомой функции $\xi(0) = [\exp(-At)\, x(t)]_{t=0} = x_0$. Интегрированием получаем

$$\xi(t) = x_0 + \int_0^t e^{-A\tau} f(\tau)\, d\tau.$$

Используя это выражение, находим решение неоднородного уравнения, которое удовлетворяет начальному условию $x(0) = x_0$,

$$x(t) = e^{At}\left(x_0 + \int_0^t e^{-A\tau} f(\tau)\, d\tau\right) = e^{At}x_0 + \int_0^t e^{A(t-\tau)} f(\tau)\, d\tau,$$

которое совпадает с выражением, приведеннім в (п. 1), и называется *формулой Коши*. Его можно рассматривать как сумму решения соответствующего однородного уравнения (при $f(t) = 0$) и решения неоднородного уравнения при нулевых начальных условиях ($x_0 = 0$).

Пусть дана неоднородная система дифференциальных уравнений в нормальной форме:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= -2x_2 + 3t \\ \frac{dx_2}{dt} &= 2x_1 + 4 \end{aligned}\right\}.$$

Найдем фундаментальную матрицу системы:

$$A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}; \quad |\lambda E - A| = \begin{bmatrix} \lambda & 2 \\ -2 & \lambda \end{bmatrix} \Delta(\lambda) = \lambda^2 + 4;$$

$$\lambda_1 = 2i; \quad \lambda_2 = -2i; \quad \Delta_{11}(\lambda) = \lambda; \quad \Delta_{12}(\lambda) = 2;$$

$$h^{(1)} = k_1 \begin{bmatrix} \Delta_{11}(\lambda_1) \\ \Delta_{12}(\lambda_2) \end{bmatrix} = k_1 \begin{bmatrix} 2i \\ 2 \end{bmatrix}; \quad h^{(2)} = k_2 \begin{bmatrix} -2i \\ 2 \end{bmatrix}.$$

Полагая $k_1 = k_2 = 1/2$, имеем:

$$H = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}; \quad H^{-1} = \frac{1}{2i} \begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix};$$

$$\Phi(t) = e^{At} = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{2it} & 0 \\ 0 & e^{-2it} \end{bmatrix} \frac{1}{2i} \begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix} =$$

$$= \frac{1}{2i} \begin{bmatrix} i(e^{2it} + e^{-2it}) & -(e^{2it} - e^{-2it}) \\ e^{2it} - e^{-2it} & i(e^{2it} + e^{-2it}) \end{bmatrix} = \begin{bmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{bmatrix}.$$

Решение задачи Коши для однородной системы

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{bmatrix} \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} = \begin{bmatrix} x_{10}\cos 2t - x_{20}\sin 2t \\ x_{10}\sin 2t + x_{20}\cos 2t \end{bmatrix}.$$

Найдем интеграл в выражении для частного решения неоднородной системы при $x_0 = 0$:

$$\int_0^t e^{-A\tau} f(\tau)\, d\tau = \int_0^t \begin{bmatrix} \cos 2\tau & \sin 2\tau \\ -\sin 2\tau & \cos 2\tau \end{bmatrix} \begin{bmatrix} 3\tau \\ 4 \end{bmatrix} d\tau =$$

$$= \int_0^t \begin{bmatrix} 3\tau \cos 2\tau + 4 \sin 2\tau \\ -3\tau \sin 2\tau + 4 \cos 2\tau \end{bmatrix} dt = \begin{bmatrix} -\frac{5}{4}\cos 2t + \frac{3}{2} t \sin 2t + \frac{5}{4} \\ \frac{5}{4} \sin 2t + \frac{3}{2} t \cos 2t \end{bmatrix}.$$

Частное решение неоднородной системы:

$$e^{At} \int_0^t e^{-A\tau} f(\tau)\, d\tau = \begin{bmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{bmatrix} \times$$

$$\times \begin{bmatrix} -\frac{5}{4}\cos 2t + \frac{3}{2} t \sin 2t + \frac{5}{4} \\ \frac{5}{4} \sin 2t + \frac{3}{2} t \cos 2t \end{bmatrix} = \begin{bmatrix} \frac{5}{4}\cos 2t - \frac{5}{4} \\ \frac{5}{4} \sin 2t + \frac{3}{2} t \end{bmatrix}.$$

Таким образом, решение неоднородной системы, которая удовлетворяет начальным условиям $x(0)= x_0$, запишется следующим образом:

$$x = \begin{bmatrix} x_{10} \cos 2t - x_{20} \sin 2t \\ x_{10} \sin 2t + x_{20} \cos 2t \end{bmatrix} + \begin{bmatrix} \frac{5}{4}\cos 2t - \frac{5}{4} \\ \frac{5}{4} \sin 2t + \frac{3}{2} t \end{bmatrix} =$$

$$= \begin{bmatrix} \left(x_{10} + \frac{5}{4}\right) \cos 2t - x_{20} \sin 2t - \frac{5}{4} \\ \left(x_{10} + \frac{5}{4}\right) \sin 2t + x_{20} \cos 2t + \frac{3}{2} t \end{bmatrix}.$$

9. Модальная матрица. Значение этой матрицы в теории дифференциальных уравнений и других приложениях настолько велико, что она заслуживает более детального рассмотрения. Обозначим характеристическую матрицу $[\lambda E - A]$ через $F(\lambda)$ и присоединенную к ней Adj $[\lambda E - A]$ - через $G(\lambda)$. Согласно (п. 3.1) $F(\lambda)G(\lambda) = G(\lambda) F(\lambda)= \Delta(\lambda) E$, где $\Delta(\lambda) = \det [\lambda E - A]$. Для значений $\lambda = \lambda_i$, $(i=1, 2, \ldots, n)$ $\Delta(\lambda_i) = 0$ и, следовательно, $F(\lambda_i)G(\lambda_i)=0$. Это матричное уравнение распадается на n уравнений относительно столбцов $g_i^{(j)}$ и матрицы $G(\lambda_i)$:

$$F(\lambda_i)\, g^{(1)}{}_i=0,\ F(\lambda_i)\, g^{(2)}{}_i=0, \ldots, F(\lambda_i) g^{(n)}{}_i=0$$

решения которых с точностью до постоянных совпадают с решением однородного уравнения, рассмотренного в (п. 2). Отсюда следует вывод, что в матрице $G(\lambda_i)$ ранга единицы все столбцы пропорциональны, поэтому любой ненулевой из них (или произведение его на произвольное число) можно принять в качестве столбеца $h^{(i)}$ модальной матрицы H. Самая матрица $G(\lambda_i)$ при этом имеет вид

$$G(\lambda_i) = [h^{(i)}\nu_{i1},\ h^{(i)}\nu_{i2},\ \dots,\ h^{(i)}\nu_{in}] = h^{(i)}[\nu_{i1},\ \nu_{i2},\ \dots,\ \nu_{in}] = h^{(i)}\nu_{(i)},$$

где числа $\nu_{i_1}, \nu_{i_2}, \dots, \nu_{i_n}$ — коэффициенты пропорциональности между столбцами матрицы $G(\lambda_i)$ и собственными векторами $h^{(i)}$; $\nu_{(i)}$ — строка, элементами которой являются эти числа.

Таким образом, для определения модальной матрицы H можно воспользоваться присоединенной матрицей $G(\lambda)$, причем векторы $h^{(i)}$ получаются подстановкой $\lambda=\lambda_i$ в $G(\lambda)$ и выбором из $G(\lambda_i)$ одного из столбцов или пропорционального ему столбца. Этот путь представляется избыточным, так как для решения этой задачи достаточно иметь один столбец матрицы $G(\lambda)$, который совпадает с соответствующей строкой матрицы алгебраических дополнений (п. 5). Однако, как будет показано дальше, $G(\lambda)$ понадобится и для получения обратной модальной матрицы H^{-1}. Для вычисления $G(\lambda_i)$ существуют различные методы. Один из них основан на соотношении

$$G(\lambda_j) = (-1)^{n-1} \prod_{i \neq j} F(\lambda_i).$$

Так, для примера из (п. 5) имеем (λ_1 = —1, λ_2 = —2, λ_3 = —3):

$$F(\lambda)=\begin{bmatrix}\lambda-4 & 8 & -1\\ -5 & \lambda+9 & -1\\ -4 & 6 & \lambda+1\end{bmatrix};$$

$$G(\lambda)=\begin{bmatrix}\lambda^2+10\lambda+15 & -8\lambda-14 & \lambda+1\\ 5\lambda+9 & \lambda^2-3\lambda-8 & \lambda+1\\ 4\lambda+6 & -6\lambda-8 & \lambda^2+5\lambda+4\end{bmatrix};$$

$$G(-1)=F(-2)F(-3)=\begin{bmatrix}-6 & 8 & -1\\ -5 & 7 & -1\\ -4 & 6 & -1\end{bmatrix}\begin{bmatrix}-7 & 8 & -1\\ -5 & 6 & -1\\ -4 & 6 & -2\end{bmatrix}=$$

$$=\begin{bmatrix}6 & -6 & 0\\ 4 & -4 & 0\\ 2 & -2 & 0\end{bmatrix};$$

$$G(-2)=F(-1)F(-3)=\begin{bmatrix}-1 & 2 & -1\\ -1 & 2 & -1\\ -2 & 4 & -2\end{bmatrix};$$

$$G(-3)=F(-1)F(-2)=\begin{bmatrix}-6 & 10 & -2\\ -6 & 10 & -2\\ -6 & 10 & -2\end{bmatrix}.$$

Приняв модальные столбцы пропорциональными столбцам каждой из полученных матриц, получим

$$H=\begin{bmatrix}3 & 1 & 1\\ 2 & 1 & 1\\ 1 & 2 & 1\end{bmatrix}.$$

При этом

$$G(-1)=h^{(1)}\nu_{(1)}=\begin{bmatrix}3\\ 2\\ 1\end{bmatrix}[2,\ -2,\ 0];$$

$$G(-2)=h^{(2)}\nu_{(2)}=\begin{bmatrix}1\\ 1\\ 2\end{bmatrix}[-1,\ 2\ -1];$$

$$G(-3)=h^{(3)}\nu_{(3)}=\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}[-6,\ 10,\ -2].$$

Строки $v_{(1)}$, $v_{(2)}$, $v_{(3)}$ образуют матрицу констант

$$N = \begin{bmatrix} 2 & -2 & 0 \\ -1 & 2 & -1 \\ 6 & 10 & -2 \end{bmatrix}.$$

Интересно отметить, что $NH=D$, где D — диагональная матрица с элементами $\Delta'(\lambda_i)$, которые равны значениям производной $\Delta(\lambda)$,) по λ при $\lambda = \lambda_i$ $(i = 1, 2, \ldots, n)$. Действительно, в нашем примере

$$\Delta'(\lambda) = 3\lambda^2 + 12\lambda + 11 \text{ и } \Delta'(-1) = 2;\ \Delta'(-2) = -1;\ \Delta'(-3) = 2.$$

В то же время

$$NH = \begin{bmatrix} 2 & -2 & 0 \\ -1 & 2 & -1 \\ -6 & 10 & -2 \end{bmatrix} \begin{bmatrix} -3 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} \Delta'(\lambda_1) & & \\ & \Delta'(\lambda_2) & \\ & & \Delta'(\lambda_3) \end{bmatrix}.$$

Можно показать, что это соотношение всегда имеет место, если все собственные значения различны. Обращая обе его части $H^{-1}N^{-1}=D^{-1}$ и умножая справа на N, находим

$$H^{-1} = D^{-1}N = \begin{bmatrix} \frac{1}{\Delta'(\lambda_1)} & & \\ & \ddots & \\ & & \frac{1}{\Delta'(\lambda_n)} \end{bmatrix} \begin{bmatrix} v_{(1)} \\ \cdots \\ v_{(n)} \end{bmatrix} = \begin{bmatrix} \frac{1}{\Delta'(\lambda_1)} v_{(1)} \\ \cdots \\ \frac{1}{\Delta'(\lambda_n)} v_{(n)} \end{bmatrix}.$$

Итак, обращение модальной матрицы при известной матрице N сводится в основном к вычислению значений производной определителя для различных собственных значений. В рассматриваемом примере:

$$H^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & -2 & 0 \\ -1 & 2 & -1 \\ -6 & 10 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -2 & 1 \\ -3 & 5 & -1 \end{bmatrix},$$

что совпадает с полученным ранее результатом.

10. Формула Коши. Используем полученное в (п. 7) соотношение для фундаментальной матрицы $\Phi(t) = He^{\Lambda t}H^{-1}$. Обозначив через $h^{(i)}$ столбцы матрицы H и через $\tilde{h}_{(i)}$ — строки матрицы H^{-1}, запишем:

$$\Phi(t) = [h^{(1)}e^{\lambda_1 t}, \ \ldots, \ h^{(n)}e^{\lambda_n t}] \begin{bmatrix} \tilde{h}_{(1)} \\ \cdots \\ \tilde{h}_{(n)} \end{bmatrix} = \sum_{i=1}^{n} h^{(i)} e^{\lambda_i t} \tilde{h}_{(i)} = \sum_{i=1}^{n} h^{(i)} \tilde{h}_{(i)} e^{\lambda_i t}.$$

Как следует из $H^{-1} = D^{-1}N$, i-я строка матрицы H^{-1} имеет вид:

$$\tilde{h}_{(i)} = \frac{\nu_{(i)}}{\Delta'(\lambda_i)}.$$

Тогда

$$h^{(i)}\tilde{h}_{(i)} = h^{(i)}\nu_{(i)} \frac{1}{\Delta'(\lambda_i)} = \frac{G(\lambda_i)}{\Delta'(\lambda_i)}.$$

Это важное соотношение позволяет представить фундаментальную матрицу формулой:

$$\Phi(t) = e^{At} = \sum_{i=1}^{n} \frac{G(\lambda_i)}{\Delta'(\lambda_i)} e^{\lambda_i t}.$$

Решение задачи Коши для однородной системы линейных дифференциальных уравнений получаем в виде:

$$x(t) = \Phi(t)\, x_0 = \sum_{i=1}^{n} \frac{G(\lambda_i)}{\Delta'(\lambda_i)} e^{\lambda_i t} x_0,$$

где x_0 — вектор начальных значений неизвестных при $t_0 = 0$.

Приведем к соответствующему виду частное решение при нулевых начальных значениях. Интеграл в формуле Коши (п. 8) преобразуем следующим образом:

$$\int_0^t e^{A(t-\tau)} f(\tau)\, d\tau = \int_0^t He^{\Delta(t-\tau)} H^{-1} f(\tau)\, d\tau =$$

$$= \int_0^t [h^{(1)}e^{\lambda_1(t-\tau)}, \ \ldots, \ h^{(n)}e^{\lambda_n(t-\tau)}] \begin{bmatrix} \tilde{h}_{(1)} \\ \cdots \\ \tilde{h}_{(n)} \end{bmatrix} f(\tau)\, d\tau =$$

$$= \int_0^t \sum_{i=1}^{n} h^{(i)}\tilde{h}_{(i)} e^{\lambda_i (t-\tau)} f(\tau)\, d\tau = \sum_{i=1}^{n} \frac{G(\lambda_i)}{\Delta'(\lambda_i)} e^{\lambda_i t} \int_0^t e^{-\lambda_i \tau} f(\tau)\, d\tau.$$

Записав сумму полученных решений, получим выражение для формулы Коши

$$x(t) = \sum_{i=1}^{n} \frac{G(\lambda_i)}{\Delta'(\lambda_i)} e^{\lambda_i t} \left(x_0 + \int_0^t e^{-\lambda_i \tau} f(\tau)\, d\tau \right),$$

или

$$x(t) = \sum_{i=1}^{n} \frac{G(\lambda_i)}{\Delta'(\lambda_i)} \left(x_0 e^{\lambda_i t} + \int_0^t e^{\lambda_i (t-\tau)} f(\tau)\, d\tau \right).$$

Решим, например, уравнения электрической схемы (рис. 2.9, *а*) при L = 1 Г, C = 0,5 Ф, R = 3 Ом и входном воздействии в виде единичной ступенчатой функции $e(t)$= 1 при $t \geq 0$ и $e(t)$ = 0 при $t < 0$ (рис. 2.9, *б):*

$$Ri(t) + L\frac{di(t)}{dt} + u(t) = e(t);$$

$$i(t) = C\frac{du(t)}{dt}.$$

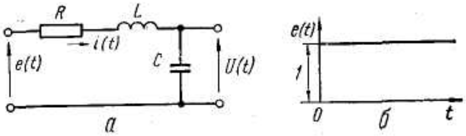


Рис. 2.9. Электрическая схема *(а)* и график входного воздействия (б).

Преобразовав к нормальной форме и подставив значение параметров, получим

$$\frac{d}{dt}\begin{bmatrix} i(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} -3 & -1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} i(t) \\ u(t) \end{bmatrix} + \begin{bmatrix} e(t) \\ 0 \end{bmatrix}.$$

Для рассматриваемой системы

$$A = \begin{bmatrix} -3 & -1 \\ 2 & 0 \end{bmatrix}; \quad F(\lambda) = \begin{bmatrix} \lambda + 3 & 1 \\ -2 & \lambda \end{bmatrix}; \quad f(t) = \begin{bmatrix} e(t) \\ 0 \end{bmatrix}:$$

Определим присоединенные матрицы и производные характеристического определителя для собственных значений:

$$\Delta(\lambda) = \lambda^2 + 3\lambda + 2; \quad \lambda_1 = -1; \quad \lambda_2 = -2;$$

$$\Delta'(\lambda) = 2\lambda + 3; \quad \Delta'(\lambda_1) = 1; \quad \Delta'(\lambda_2) = -1;$$

$$G(\lambda_1) = -F(\lambda_2) = \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix}; \quad G(\lambda_1) = -F(\lambda_2) = \begin{bmatrix} -2 & -1 \\ 2 & 1 \end{bmatrix}.$$

По формуле Коши имеем

$$\begin{bmatrix} i(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix} e^{-t} \left(\begin{bmatrix} i(0) \\ u(0) \end{bmatrix} + \int_0^t e^{\tau} \begin{bmatrix} 1 \\ 0 \end{bmatrix} d\tau \right) +$$

$$+ \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} e^{-2t} \left(\begin{bmatrix} i(0) \\ u(0) \end{bmatrix} + \int_0^t e^{2\tau} \begin{bmatrix} 1 \\ 0 \end{bmatrix} d\tau \right),$$

что приводит к решению ($t \geq 0$):

$$\begin{bmatrix} i(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} -e^{-t} + 2e^{-2t} & -e^{-t} + e^{-2t} \\ 2e^{-t} - 2e^{-2t} & 2e^{-t} - e^{-2t} \end{bmatrix} \begin{bmatrix} i(0) \\ u(0) \end{bmatrix} + \begin{bmatrix} e^{-t} - e^{-2t} \\ 1 - 2e^{-t} + e^{-2t} \end{bmatrix}.$$

11. Уравнение *n*-го порядка. Рассмотрим дифференциальное уравнение n-го порядка

$$\frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \cdots + a_{n-1} \frac{dx}{dt} + a_n x = \xi(t).$$

Задача Коши для такого уравнения состоит в отыскании решения $x(t)$, удовлетворяющего начальным условиям при $t = 0$,

$$x = x_0, \quad \frac{dx}{dt} = x_0', \quad \ldots, \quad \frac{d^{n-1} x}{dt^{n-1}} = x_0^{(n-1)}.$$

Уравнение n-го порядка приводится к нормальной форме серией подстановок:

$$x = x_1; \quad \frac{dx}{dt} = x_2; \quad \ldots; \quad \frac{d^{n-1} x}{dt^{n-1}} = x^n.$$

Тогда получаем эквивалентную систему линейных дифференциальных уравнений первого порядка, разрешенных относительно производных от переменных $x_1, x_2, \ldots, x_n$:

$$\left.\begin{aligned} &\frac{dx_1}{dt} = x_2; \ \frac{dx_2}{dt} = x_3; \quad \ldots, \ \frac{dx_{n-1}}{dt} = x_n \\ &\frac{dx_n}{dt} = -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n + \xi(t) \end{aligned}\right\}.$$

В матричной записи эта система имеет вид:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_2 & -a_1 \end{bmatrix} \times$$

$$\times \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \\ \xi(t) \end{bmatrix}.$$

Характеристическая матрица

$$F(\lambda) = (\lambda E - A) = \begin{bmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & \lambda + a_1 \end{bmatrix}.$$

Определитель этой матрицы раскроем по элементам последней строки. После удаления этой строки и j-го столбца получаем определитель, элементы которого выше главной диагонали равны нулю, а по главной диагонали располагаются $j — 1$ элементов λ и $n — j$ элементов, равных - 1. Следовательно, алгебраические дополнения элементов последней строки

$$\Delta_{nj}(\lambda) = (-1)^{n+j}(-1)^{n-j}\lambda^{j-1} = \lambda^{j-1} \quad (j = 1,\ 2,\ \ldots,\ n),$$

и характеристическое уравнение получаем в виде:

$$\Delta(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0.$$

Заметим, что это уравнение можно получить непосредственно из однородного дифференциального уравнения n-го порядка заменой операторов дифференцирования $\dfrac{d^i}{dt^i}$ на λ^i $(i = 1, 2, ..., n)$. Воспользовавшись выражениями для $\Delta_{nj}(\lambda)$ запишем модальную матрицу, которая в случае различных собственных чисел $\lambda_1, \lambda_2, ..., \lambda_n$ принимает стандартную форму:

$$H = \begin{bmatrix} \Delta_{n1}(\lambda_1) & \Delta_{n1}(\lambda_2) & \dots & \Delta_{n1}(\lambda_n) \\ \Delta_{n2}(\lambda_1) & \Delta_{n2}(\lambda_2) & \dots & \Delta_{n2}(\lambda_n) \\ \dots & \dots & \dots & \dots \\ \Delta_{nn}(\lambda_1) & \Delta_{nn}(\lambda_2) & \dots & \Delta_{nn}(\lambda_n) \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \dots & \dots & \dots & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{bmatrix}.$$

Определитель этой матрицы известен под названием *определитель Вандермонда*.

Так как решением однородного дифференциального уравнения n-й степени является только первый элемент $x_1 = x$ общего решения H exp (Λt) $H^{-1}x_0 = H$ exp $(\Lambda t)c$, то при перемножении матриц достаточно использовать только первую строку H, все элементы которой равны единицы, т.е.

$$x = \|1,\ 1,\ \dots,\ 1\| \begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \dots & \\ & & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix},$$

где $c = H^{-1}x_0$ или

$$\begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \dots & \dots & \dots & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ x_0' \\ \dots \\ x_0^{(n-1)} \end{bmatrix}.$$

В результате получим

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} + \dots + c_n e^{\lambda_n t} = \sum_{i=1}^{n} c_i e^{\lambda_i t}.$$

Следует заметить, что обращение модальной матрицы H всегда возможно, ибо определитель Вандермонда не равен нулю.

12. Формула Коши для неоднородного уравнения n-го порядка. Выражение для первого элемента $x_1 = x$ получим, заменив в формуле Коши (п. 10) матрицу $G(\lambda_i)$ ее первой строкой $g_{(1)} = [g_{11}, \dots, g_{1n}]$, т.е.

$$x(t) = \sum_{i=1}^{n} \frac{g_{(1)}(\lambda_i)}{\Delta'(\lambda_i)} \left(x_0 e^{\lambda_i t} + \int_0^t e^{\lambda_i (t-\tau)} f(\tau)\, d\tau \right).$$

Найдем общее выражение для элементов первой строки $g_{(1)}$ присоединенной матрицы $G(\lambda)$ = Adj[λE — A]. На основании соотношения $G(\lambda)F(\lambda) = \Delta(\lambda) E$ можно записать:

$$[g_{11}, g_{12}, \ldots, g_{1n}] \begin{bmatrix} \lambda & -1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_n & a_{n-1} & \cdots & \lambda + a_1 \end{bmatrix} = [\Delta(\lambda), 0, \ldots, 0].$$

Отсюда получаем систему уравнений относительно подлежащих определению элементов g_{1j} (j = 1, 2, ..., n) первой строки матрицы $G(\lambda)$:

$$g_{11}\lambda + g_{1n}a_n = \Delta(\lambda);$$
$$-g_{1j} + g_{1,\,j+1}\lambda + g_{1n}a_{n-j} = 0 \quad (j = 1, 2, \ldots, n-2);$$
$$-g_{1,\,n-1} + g_{1n}(\lambda + a_1) = 0.$$

Так как $g_{1n} = \Delta_{n1}(\lambda) = 1$, то приходим к рекуррентной формуле

$$g_{1j} = g_{1,\,j+1}\lambda + a_{n-j}$$
$$(j = n-1, \ldots, 1).$$

Элемент g_{1j} представим в общем виде

$$g_{1j} = \lambda^{n-j} + a_1\lambda^{n-j-1} + \cdots +$$
$$+ a_{n-j-1}\lambda + a_{n-j}.$$

Это выражение соответствует и первому уравнению. Действительно,

$$g_{11} = \frac{1}{\lambda}(\Delta(\lambda) - a_n) = \frac{1}{\lambda}(\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n - a_n) =$$
$$= \lambda^{n-1} + a_1\lambda^{n-2} + \cdots + a_{n-2}\lambda + a_{n-1},$$

что совпадает с приведенным выше при j = 1. Процесс образования элементов g_{1j} можно представить как последовательный сдвиг коэффициентов характеристического уравнения $\Delta(\lambda)$ (рис. 2.10).

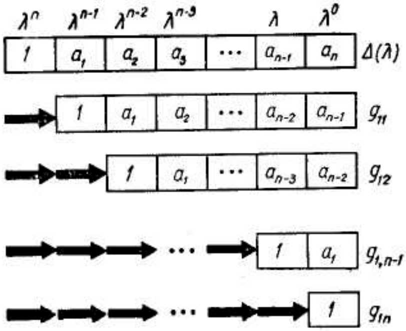


Рис. 2.10. Процесс образования элементов g_{1j} (j = 1, 2, ..., n)

Полученные результаты позволяют записать формулу Коши для линейного дифференциального уравнения n-го порядка следующим образом:

$$x(t)=\sum_{i=1}^{n}\frac{1}{\Delta'(\lambda_i)}\left(g_{(1)}(\lambda_i)\,x_0e^{\lambda_i t}+\ g_{(1)}(\lambda_i)\int_0^t e^{\lambda_i(t-\tau)}f(\tau)\,d\tau\right).$$

Заменив произведение строки $g_{(1)}$ на столбец x_0 начальных значений суммой и учтя, что все элементы столбца $f(\tau)$, кроме последнего, равны нулю и g_{1n} =1, получим окончательно

$$x(t)=\sum_{i=1}^{n}\frac{1}{\Delta'(\lambda_i)}\left(\sum_{j=1}^{n}g_{1j}(\lambda_i)\,x_0^{(j-1)}e^{\lambda_i t}+\int_0^t e^{\lambda_i(t-\tau)}\xi(\tau)\,d\tau\right).$$

Найдем, например, решение данного ниже уравнения при начальных условиях x_0 = 2; x_0' = 5:

$$\frac{d^2x}{dt^2}+5\frac{dx}{dt}+6x=36t.$$

Определим величины, которые входят в общую формулу:

$$\Delta(\lambda)=\lambda^2+5\lambda+6=(\lambda+2)(\lambda+3);\quad \lambda_1=-2;\quad \lambda_2=-3;$$
$$\Delta'(\lambda)=2\lambda+5;\quad \Delta'(\lambda_1)=1;\quad \Delta'(\lambda_2)=-1;$$
$$g_{11}=\lambda+5;\quad g_{12}=1;\quad g_{11}(\lambda_1)=3;\quad g_{11}(\lambda_2)=2;$$
$$g_{12}(\lambda_1)=g_{21}(\lambda_2)=1.$$

По формуле имеем

$$x(t) = (3x_0 + x_0')e^{-2t} + \int_0^t e^{-2(t-\tau)}36\tau\, d\tau - (2x_0 + x_0')\, e^{-3t} -$$
$$- \int_0^t e^{-3(t-\tau)}36\tau\, d\tau = (3x_0 + x_0' + 9)\, e^{-2t} -$$
$$- (2x_0 + x_0' + 4)\, e^{-3t} + 6t - 5,$$

откуда при заданных начальных значениях получаем решение задачи Коши:

$$x(t) = 20e^{-2t} - 13e^{-3t} + 6t - 5.$$

Алгебраические дополнения последней строки характеристической матрицы $F(\lambda)$, которые получены в (п.11), являются по определению элементами последнего столбца $g^{(n)}$ присоединенной матрицы $G(\lambda)$, т.е. $g_{jn} = \Delta_{nj}(\lambda) = \lambda^{j-1}$ (j=1, 2, ... n). Они же приняты в качестве общего выражения элементов столбца модальной матрицы H. Поэтому можно записать присоединенную матрицу как $G(\lambda) = g^{(n)}(\lambda)\; g_1(\lambda)$. Строки $g_{(i)}(\lambda_1)$, $g_{(1)}(\lambda_2)$.....$g_{(1)}(\lambda_n)$ образуют матрицу N и, следовательно, обратная модальная матрица может быть получена из соотношение:

$$H^{-1} = D^{-1}N = \begin{bmatrix} \frac{1}{\Delta'(\lambda_1)} g_{(1)}(\lambda_1) \\ \cdot\; \cdot\; \cdot\; \cdot\; \cdot\; \cdot \\ \frac{1}{\Delta'(\lambda_n)} g_{(1)}(\lambda_n) \end{bmatrix}.$$

Так, для нашего примера

$$H = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -3 \end{bmatrix};$$

$$H^{-1} = \begin{bmatrix} \frac{1}{\Delta'(\lambda_1)} g_{11}(\lambda_1) & \frac{1}{\Delta'(\lambda_1)} g_{12}(\lambda_1) \\ \frac{1}{\Delta'(\lambda_2)} g_{11}(\lambda_2) & \frac{1}{\Delta'(\lambda_2)} g_{12}(\lambda_2) \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -2 & -1 \end{bmatrix}.$$

Действительно, проверка подтверждает правильность полученного результата

$$HH^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

13. Приведение системы уравнений к нормальной форме. Если система дифференциальных уравнений первого порядка задана в произвольной форме

$$\left.\begin{aligned} q_{11}\frac{dx_1}{dt}+\cdots+q_{1n}\frac{dx_n}{dt} &= w_{11}x_1+\cdots+w_{1n}x_n+ \\ &+u_{11}v_1(t)+\cdots+u_{1m}v_m(t) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ q_{n1}\frac{dx_1}{dt}+\cdots+q_{nn}\frac{dx_n}{dt} &= w_{n1}x_1+\cdots+w_{nn}x_n+ \\ &+u_{n1}v_1(t)+\cdots+u_{nm}v_m(t) \end{aligned}\right\},$$

где $v_1(t), \ldots, v_m(t)$ — m задающих функций времени.

В матричной записи эта система имеет вид:

$$Q(dx/dt) = Wx + Uv.$$

Приведение ее к нормальной форме означает преобразование матрицы Q в единичную, что можно осуществить умножением уравнения слева на Q^{-1}, если $\det Q \neq 0$. Тогда находим

$$dx/dt = Ax + Bv,$$

где $A = Q^{-1}W$ и $B = Q^{-1}U$.

В общем случае более целесообразно применить процедуру исключения Гаусса-Жордана (п. 4.3) к матрице $[Q, W, U]$ по столбцам матрицы Q. В результате получим матрицу $[1, A, B]$, которая и определяет нормальную форму исходной системы дифференциальных уравнений.

Пример:

$$\begin{bmatrix} 1 & 1 & 2 \\ 3 & 2 & 4 \\ 5 & 3 & 8 \end{bmatrix}\frac{dx}{dt} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 2 \\ -1 & -6 & 4 \end{bmatrix}x(t) + \begin{bmatrix} 0 & -1 \\ 4 & 0 \\ -2 & 1 \end{bmatrix}v(t).$$

Преобразование матрицы (Q, W, U) выполняется следующим образом:

$$\left[\begin{array}{ccc|ccc|cc} 1 & 1 & 2 & 1 & -2 & 0 & 0 & -1 \\ 3 & 2 & 4 & 0 & 1 & 2 & 4 & 0 \\ 5 & 3 & 8 & -1 & -6 & 4 & -2 & 1 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc|cc} 1 & 1 & 2 & 1 & -2 & 0 & 0 & -1 \\ 0 & -1 & -2 & -3 & 7 & 2 & 4 & 3 \\ 0 & -2 & -2 & -6 & 4 & 4 & -2 & 6 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc|cc} 1 & 0 & 0 & -2 & 5 & 2 & 4 & 2 \\ 0 & 1 & 2 & 3 & -7 & -2 & -4 & -3 \\ 0 & 0 & 2 & 0 & -10 & 0 & -6 & 0 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc|cc} 1 & 0 & 0 & -2 & 5 & 2 & 4 & 2 \\ 0 & 1 & 0 & 3 & 3 & -2 & 2 & -3 \\ 0 & 0 & 1 & 0 & -5 & 0 & -3 & 0 \end{array}\right].$$

Эквивалентная система уравнений в нормальной форме

$$\frac{dx}{dt} = \begin{bmatrix} -2 & 5 & 2 \\ 3 & 3 & -2 \\ 0 & -5 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 4 & 2 \\ 2 & -3 \\ -3 & 0 \end{bmatrix} v(t).$$

Если матрица Q особенная, то в процессе исключения в этой матрице образуется нулевая строка, которая указывает на алгебраическую зависимость переменных $x_1, x_2, ..., x_n$. Производная одной из них может быть исключенная. Для этого алгебраическое уравнение, соответствующее нулевой строке матрицы Q, дифференцируется, из него определяется производная исключаемой переменной, которая, наряду с самой переменной, подставляется в остальные уравнения. Затем процедура исключения продолжается.

Видоизменим, например, матрицу Q так, чтобы она была особенной и проведем исключение

$$\left[\begin{array}{ccc|ccc|cc} 1 & 1 & 2 & 1 & -2 & 0 & 0 & -1 \\ 3 & 2 & 4 & 0 & 1 & 2 & 4 & 0 \\ 5 & 4 & 8 & -1 & -6 & 4 & -2 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc|cc} 1 & 1 & 2 & 1 & -2 & 0 & 0 & -1 \\ 0 & -1 & -2 & -3 & 7 & 2 & 4 & 3 \\ 0 & -1 & -2 & -6 & 4 & 4 & -2 & 6 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc|cc} 1 & 0 & 0 & -2 & 5 & 2 & 4 & 2 \\ 0 & 1 & 2 & 3 & -7 & -2 & -4 & -3 \\ 0 & 0 & 0 & -3 & -3 & 2 & -6 & 3 \end{array}\right].$$

Преобразованная система имеет вид:

$$\left.\begin{aligned}\frac{dx_1}{dt} &= -2x_1 + 5x_2 + 2x_3 + 4v_1 + 2v_2\\ \frac{dx_2}{dt} + 2\frac{dx_3}{dt} &= 3x_1 - 7x_2 - 2x_3 - 4v_1 - 3v_2\\ 0 &= -3x_1 - 3x_2 + 2x_3 - 6v_1 + 3v_2\end{aligned}\right\}.$$

Последнее уравнение не содержит производных и указывает на зависимость переменных x_1, x_2 и x_3. Выразим из него одну из переменных (например, x_2)

$$x_2 = -x_1 + (2/3)\,x_3 - 2v_1 + v_2$$

и продифференцируем это выражение

$$\frac{dx_2}{dt} = -\frac{dx_1}{dt} + \frac{2}{3}\frac{dx_3}{dt} - 2\frac{dv_1}{dt} + \frac{dv_2}{dt}.$$

Подставляя значение x_2 и ее производной в остальные уравнения, получаем:

$$\left.\begin{aligned}\frac{dx_1}{dt} &= -7x_1 + \frac{16}{3}x_3 - 6v_1 + 7v_2\\ -\frac{dx_1}{dt} + \frac{8}{3}\frac{dx_3}{dt} &= 10x_1 - \frac{20}{3}x_3 + 10v_1 - 10v_2 + 2\frac{dv_1}{dt} - \frac{dv_2}{dt}\end{aligned}\right\},$$

или в матричной записи:

$$\begin{bmatrix} 1 & 0\\ -1 & \frac{8}{3}\end{bmatrix}\begin{bmatrix}\frac{dx_1}{dt}\\ \frac{dx_3}{dt}\end{bmatrix} = \begin{bmatrix}-7 & \frac{16}{3}\\ 10 & -\frac{20}{3}\end{bmatrix}\begin{bmatrix}x_1\\ x_3\end{bmatrix} + \begin{bmatrix}-6 & 7\\ 10 & -10\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix} + \begin{bmatrix}0 & 0\\ 2 & -1\end{bmatrix}\begin{bmatrix}\frac{dv_1}{dt}\\ \frac{dv_2}{dt}\end{bmatrix}.$$

Как видим, в правой части системы появились производные задаючих функций и уравнение принимают более общую форму:

$$Q\frac{dx}{dt} = Wx + Uv + U'\frac{dv}{dt}.$$

Процедура исключения продолжается над матрицей [Q, W, U, U'] по столбцам матрицы Q:

$$\left[\begin{array}{cc|cc|cc|cc} 1 & 0 & -7 & \frac{16}{3} & -6 & 7 & 0 & 0\\ -1 & \frac{8}{3} & 10 & -\frac{20}{3} & 10 & -10 & 2 & -1\end{array}\right] \to$$

$$\to \left[\begin{array}{cc|cc|cc|cc} 1 & 0 & -7 & \frac{16}{3} & -6 & 7 & 0 & 0\\ 0 & 1 & \frac{9}{8} & -\frac{1}{2} & \frac{3}{2} & -\frac{9}{8} & \frac{3}{4} & -\frac{3}{8}\end{array}\right].$$

Таким образом, приходим к системе в нормальной форме:

$$\left.\begin{aligned}\frac{dx_1}{dt} &= -7x_1 + \frac{16}{3}x_3 - 6v_1 + 7v_2\\ \frac{dx_3}{dt} &= \frac{9}{8}x_1 - \frac{1}{2}x_3 + \frac{3}{2}v_1 - \frac{9}{8}v_2 + \frac{3}{4}\frac{dv_1}{dt} - \frac{3}{8}\frac{dv_2}{dt}\end{aligned}\right\}.$$

Если в процессе исключения появляется нулевая строка во всей матрице [Q, *W*, *U*], то это свидетельствует о неопределенности системы. Если же в какой-либо строке элементы матриц *Q* и *W* нулевые, а некоторые из элементов матрицы *U* в этой строке отличны от нуля, то система несовместна. Очевидно, условием совместности является равенство рангов матриц [Q, *W*] и [*Q, W, U*].

В общем случае система линейных уравнений может включать и производные высших порядков. Устранение высших производных осуществляется заменой переменных и введением дополнительных уравнений подобно тому, как это делалось в (п. 12) при приведении к нормальной форме уравнения *n*-го порядка.

В этом микромодули рассмотрен простейший класс дифференциальных уравнений при условии, что все нули характеристического многочлена различны. В случае кратных корней структура решения усложняется. Соответствующий аппарат удобно рассматривать на языке теории функций от матриц, основы которой излагаются в следующем микромодуле.

Микромодуль 7
Индивидуальные тестовые задачи

1. Для данной матрицы

$$A = \begin{bmatrix} -3 & 4 & -2 \\ 1 & 0 & 1 \\ 6 & -6 & 5 \end{bmatrix}$$

а) запишите характеристическую матрицу $F(\lambda) = (\lambda E - A)$;

б) выразите алгебраические дополнения матрицы $F(\lambda)$ и запишите присоединенную матрицу $G(\lambda) = \mathrm{Adj}(\lambda E - A)$;

в) представьте определитель $\Delta(\lambda)=\det(\lambda E - A)$ как многочлен от λ;

г) из решения уравнения $\Delta(\lambda) = 0$ найдите собственные значения λ_1, λ_2, λ_3 матрицы *A*.

2. Для матрицы из задачи 1:

а) определите собственные векторы $h^{(1)}$, $h^{(2)}$, $h^{(3)}$, используя алгебраические дополнения элементов какой-либо строки матрицы;

б) запишите модальную матрицу *H* и найдите обратную к ней H^{-1};

в) проверьте соотношение $\Lambda = H^{-1}AH$ и $A = H\Lambda H^{-1}$, где $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \lambda_3]$.

3. Запишите в матричной форме однородную систему дифференциальных уравнений

$$\left.\begin{aligned} \frac{dx_1}{dt} &= -3x_1 + 4x_2 - 2x_3 \\ \frac{dx_2}{dt} &= x_1 + x_3 \\ \frac{dx_3}{dt} &= 6x_1 - 6x_2 + 5x_3 \end{aligned}\right\}.$$

Воспользовавшись результатами предыдущей задачи, найдите:

а) общее решение этой системы в векторной форме $x=H\varphi(t)c$ и выражения для переменных $x_1(t), x_2(t), x_3(t)$;

б) фундаментальную матрицу $\Phi(t) = H\varphi(t)H^{-1}$;

в) решение системы при начальных значениях $x_{10} = 3$, $x_{20} = -2$, $x_{30} = 0$.

4. Решите систему дифференциальных уравнений из задачи 3 с помощью преобразования переменных $x = Hy$. Запишите решения для векторов y и x, а также выразите переменные y_1, y_2, y_3 при начальных значениях x_{10}, x_{20}, x_{30}.

5. Запишите решение системы дифференциальных уравнений из задачи 3 с помощью формулы Коши.

6. Решите неоднородную систему дифференциальных уравнений

$$\left.\begin{aligned} \frac{dx_1}{dt} &= 2x_1 + x_2 + 2e^{t} \\ \frac{dx_2}{dt} &= x_1 + 2x_2 - 3e^{4t} \end{aligned}\right\}$$

при начальных значениях $x_{10}=2$ и $x_{20}=-1$ с помощью

а) фундаментальной матрицы; б) формулы Коши.

7. Приведите однородное дифференциальное уравнение третьего порядка

$$\frac{d^3x}{dt^3} + 6\frac{d^2x}{dt^2} + \frac{dx}{dt} + 6x = 0$$

к нормальной форме и покажите, что решение полученной системы при начальных значениях $x(0)=1$, $x'(0)=x''(0)=0$ имеет вид:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x(t) \\ x'(t) \\ x''(t) \end{bmatrix} = \begin{bmatrix} 3e^{-t} - 3e^{-2t} + e^{-3t} \\ -3e^{-t} + 6e^{-2t} - 3e^{-3t} \\ 3e^{-t} - 12e^{-2t} + 9e^{-3t} \end{bmatrix}.$$

Как записать решение исходного уравнения?

8. Покажите, что уравнение

$$\frac{d^2x}{dt^2} + x = 4e^t$$

при начальных условиях $x(0) = 4$ и $x'(0) = -3$ имеет решение

$$x = 2\cos t - 5\sin t + 2e^t.$$

9. Дана система дифференциальных уравнений первого порядка

$$\begin{bmatrix} 3 & -2 & 0 & -1 \\ 0 & 2 & 2 & 1 \\ 1 & -2 & -3 & -2 \\ 0 & 1 & 2 & 1 \end{bmatrix} \frac{dx}{dt} = \begin{bmatrix} 33 & 20 & 18 & 7 \\ -24 & -16 & -9 & -9 \\ 43 & 28 & 11 & 20 \\ -21 & -14 & -4 & -11 \end{bmatrix} x + \begin{bmatrix} -2 \\ 0 \\ 3 \\ -1 \end{bmatrix} e^{3t}.$$

а) Покажите, что данная система приводится к нормальной форме:

б) Определите фундаментальную матрицу $\Phi(t) = H\varphi(t)H^{-1}$ нормальной системы и найдите решение, которое удовлетворяет начальным условиям $x_0=(1;\ 0;\ 0;\ 0)$.

в) Получите решение системы при тех же начальных условиях по формуле Коши.

Микромодуль 8
Функции от матриц

3.8. Определение функции от матрицы

Функции от матриц - это один из важнейших и, наверное, наиболее сложных разделов теории матриц.

В предыдущем микромодуле на примере дифференциальных уравнений с постоянными коэффициентами был выяснен смысл экспоненциальной функции от матрицы и установлена ее связь с проблемой собственных значений для простого случая, когда все они различны. В интересах многочисленных приложений следует обобщить понятие функции от матрицы и снять ограничения на характер собственных значений.

Подобно поэтому, как всякая аналитическая функция $f(x)$ может быть представленная сходящимся рядом (многочленом) от x

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots = \sum_{s=0}^{n} a_s x^s \ \ldots$$

функция от матрицы $f(X)$ представима в виде *многочлена от матрицы*, который формально получается заменой скалярной переменной x матрицей X:

$$f(X) = a_0 + a_1 X + a_2 X^2 + \cdots = \sum_{s=1}^{\infty} a_s X^s.$$

Например,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots; \quad \sin X = X - \frac{X^3}{3!} + \frac{X^5}{5!} - \cdots$$

Если ряд сходится достаточно быстро, то функцию от матрицы можно вычислить суммированием членов. Однако часто такой путь связан с огромными вычислительными трудностями. Кроме того, из поля зрения ускользают общие свойства и закономерности, которые нередко имеют первостепенное значение. Аппарат аналитической теории функций от матриц содержит методы их компактного представления, которые могут использоваться для вычисления таких функций.

Различные способы представления и определение функции от матрицы *f(X)* сводятся в основном к двум подходам:

1) разложение функции *f(X)* в ряд приводится к более простому виду, для которого можно найти эффективные методы ее определения;

2) матрица *X* преобразуется к некоторой другой матрице $\tilde{X}$, для которой *f(X)* выражается просто через скалярные функции (например, в (п.6) предыдущего микромодуля было использовано преобразование к диагональной матрице).

Пусть даны квадратная матрица $A = \| a_{ik} \|^n_1$ и функция $f(\lambda)$ скалярного аргумента λ. Требуется определить, что следует понимать под $f(A)$, т.е. требуется распространить функцию $f(\lambda)$ и на матричные значения аргумента.

Решение этой задачи нам известно для простейшего случая, когда $f(\lambda) = \gamma_0\lambda^l + \gamma_1\lambda^{l-1} + \ldots + \gamma_l$ - многочлен относительно λ. В этом случае $f(A) = \gamma_0 A^l + \gamma_1 A^{l-1} + \ldots + \gamma_l E$... Исходя из этого часного случая, постараемся определить *f(A)* в общем случае.

Обозначим через

$$\psi(\lambda) = (\lambda - \lambda_1)^{m1} (\lambda - \lambda_2)^{m2} \ldots (\lambda - \lambda_s)^{ms} \tag{3.71}$$

минимальный многочлен матрицы *A* (здесь $\lambda_1, \lambda_2, \ldots, \lambda_S$ — все различные характеристические числа матрицы *A*). Степень этого многочлена

$$m = \sum_{k=1}^{s} m_k .$$

Пусть два многочлена $g(\lambda)$ и $h(\lambda)$ таковы, что

$$G(A) = h(A). \qquad (3.72)$$

Тогда разность $d(\lambda) = g(\lambda) - h(\lambda)$, будучи аннулирующим многочленом для матрицы A, делится на $\psi(\lambda)$ без остатка, что запишем так:

$$g(\lambda) \equiv h(\lambda)(\mathrm{mod}\ \psi(\lambda)). \qquad (3.73)$$

Отсюда в силу (3.71)

$$d(\lambda_k) = 0,\quad d'(\lambda_k) = 0, \ldots, d^{(mk-1)}(\lambda_k) = 0 \quad (k = 1, 2, \ldots, s),$$

т.е.

$$g(\lambda_k)=h(\lambda_k),\ g'(\lambda_k)= h'(\lambda_k), \ldots,\ g^{(mk-1)}(\lambda_k)=h^{(mk-1)}(\lambda_k)\ (k=1,2,\ldots, s). \qquad (3.74)$$

m чисел

$$f(\lambda_k), f'(\lambda_k), \ldots, f^{(mk-1)}(\lambda_k)(k=1,2,\ldots, s) \qquad (3.75)$$

будем условно называть *значениями функции* $f(\lambda)$ *на спектре матрицы* A и совокупность этих значений символически обозначать через $f(\Lambda_A)$. Если для функции $f(\lambda)$ существуют (т.е. имеют смысл) значения (3.75), то мы будем говорить, что *функция* $f(\lambda)$ *определена на спектре матрицы* A.

Равенства (3.74) показывают, что многочлены $g(\lambda)$ и $h(\lambda)$ имеют одни и те же значения на спектре матрицы A. В символической записи

$$g(\Lambda_A)= h(\Lambda_A).$$

Рассуждения наши обратимы: из (3.74) вытекают (3.73) и, следовательно, (3.72).

Таким образом, если задана матрица A, то значения многочлена $g(\lambda)$ на спектре матрицы A целиком определяют матрицу $g(A)$, т.е. все многочлены $g(\lambda)$, принимающие одни и те же значения на спектре матрицы A, имеют одно и то же матричное значение $g(A)$.

Мы потребуем, чтобы определение $f(A)$ в общем случае подчинялось такому же принципу: *значения функции* $f(\lambda)$ *на спектре матрицы* A *должны полностью определять* $f(A)$, *т.е. все функции* $f(\lambda)$, *имеющие одни и те же значения на спектре матрицы* A, *должны иметь одно и то же матричное значение* $f(A)$ (кроме того, предполагается что общее определение для $f(A)$ в частном случае, когда $f(\lambda)$ — многочлен, должно давать тот же результат, что и непосредственная подстановка в многочлен вместо λ матрицы A).

Но тогда, наверное, для определения $f(A)$ в общем случае достаточно подыскать такой многочлен $g(\lambda)$, который принимал бы те же значения на спектре матрицы A, что и $f(\lambda)$ (в 3.9 будет показано, что такой интерполяционный многочлен всегда существует, и будет дан алгоритм для вычисления коэффициентов интерполяционного многочлена наименьшей степени), и положить:

$$f(A)=g(A).$$

Таким образом, приходим к следующему определению:

Определение 1. Если функция $f(\lambda)$ определена на спектре матрицы *A,* то

$$f(A)=g(A),$$

где $g(A)$ — любой многочлен, принимающий на спектре матрицы A те же значения, что и $f(\lambda)$:

$$f(\Lambda_A)=g(\Lambda_A).$$

Среди всех многочленов с комплексными коэффициентами, которые принимает те же значения на спектре, что и $f(\lambda)$, имеется один и только один многочлен $r(\lambda)$ степени $< m$ (этот многочлен получается из любого другого многочлена, имеющего те же спектральные значения, что и остаток от деления на $\psi(\lambda)$). Этот многочлен $r(\lambda)$ однозначно определяется интерполяционными условиями:

$$r(\lambda_k)=f(\lambda_k),\ r'(\lambda_k)=f'(\lambda_k),\ \ldots,\ r^{(m_k-1)}(\lambda_k)=f^{(m_k-1)}(\lambda_k) \quad (k=1,2,\ldots,s). \tag{3.76}$$

Многочлен $r(\lambda)$ называют *интерполяционным многочленом Лагранжа-Сильвестра* для функции $f(\lambda)$ на спектре матрицы A. Определение 1 можно еще сформулировать так:

Определение 2. Пусть $f(\lambda)$ — функция, которая определена на спектре матрицы *A,* а $r(\lambda)$ — соответствующий интерполяционный многочлен Лагранжа-Сильвестра. Тогда

$$f(A)=r(A).$$

Замечание. Если минимальный многочлен $\psi(\lambda)$ матрицы A не имеет кратных корней [в равенстве (3.71) $m_1 = m_2 = \ldots = m_s = 1$; $s = m$], то для того, чтобы $f(A)$ имело смысл, достаточно, чтобы функция $f(\lambda)$ была определена в характеристических точках $\lambda_1, \lambda_2, \ldots, \lambda_m$. Если же $\psi(\lambda)$ имеет кратные корни, то в некоторых характеристических точках должны быть определены и производные от $f(\lambda)$ до известного порядка.

Пример 1. Рассмотрим матрицу

$$H=\overbrace{\begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{Vmatrix}}^{n}.$$

Для нее минимальным многочленом будет λ^n. Поэтому значениями $f(\lambda)$ на спектре H будут числа $f(0)$, $f'(0)$, ..., $f^{n-1}(0)$, и многочлен $r(\lambda)$ будет иметь вид

$$r(\lambda) = f(0) + \frac{f'(0)}{1!}\lambda + \ldots + \frac{f^{(n-1)}(0)}{(n-1)!}\lambda^{n-1}.$$

Таким образом,

$$f(H) = f(0)E + \frac{f'(0)}{1!}H + \ldots + \frac{f^{(n-1)}(0)}{(n-1)!}H^{n-1} = \left\| \begin{array}{cccc} f(0) & \frac{f'(0)}{1!} & \cdots & \frac{f^{(n-1)}(0)}{(n-1)!} \\ 0 & f(0) & \ddots & \vdots \\ \vdots & & \ddots & \frac{f'(0)}{1!} \\ 0 & 0 & \cdots & f(0) \end{array} \right\|.$$

Пример 2. Рассмотрим матрицу

$$J = \overbrace{\left\| \begin{array}{ccccc} \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_0 \end{array} \right\|}^{n}.$$

Заметим, что $J = \lambda_0 E + H$ и, следовательно, $J - \lambda_0 E = H$. Минимальным многочленом для J, наверное, будет $(\lambda - \lambda_0)^n$. Интерполяционный многочлен $r(\lambda)$ для функции $f(\lambda)$ определится равенством

$$r(\lambda) = f(\lambda_0) + \frac{f'(\lambda_0)}{1!}(\lambda - \lambda_0) + \ldots + \frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(\lambda - \lambda_0)^{n-1}.$$

Поэтому

$$f(J) = r(J) = f(\lambda_0)E + \frac{f'(\lambda_0)}{1!}H + \ldots + \frac{f^{(n-1)}(\lambda_0)}{(n-1)!}H^{n-1} =$$

$$= \left\| \begin{array}{cccc} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \cdots & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & f(\lambda_0) & \ddots & \vdots \\ \vdots & & \ddots & \frac{f'(\lambda_0)}{1!} \\ 0 & 0 & \cdots & f(\lambda_0) \end{array} \right\|.$$

Отметим три свойства функции от матрицы.

1° Если $\lambda_1, \lambda_2, ..., \lambda_n$ — характеристические числа матрицы n-го порядка A, то $f(\lambda_1), f(\lambda_2), .. . , f(\lambda_n)$ — полная система характеристических чисел матрицы $f(A)$.

В частном случае, когда $f(\lambda)$ — многочлен, это предложение было доказано раньше. Доказательство для общего случая сводится к этому частному случаю, поскольку (в силу определения 2) $f(A)=r(A)$ и $f(\lambda_i) = r(\lambda_i)$ (i = 1, ..., n), где $r(\lambda)$ — интерполяционный многочлен Лагранжа-Сильвестра для функции $f(\lambda)$.

2° *Если две матрицы A и B подобны и матрица T преобразует A в B:*

$$B= T^{-1}AO,$$

то матрицы f(A) и f(B) подобны и та же матрица T преобразует f(A) в f(B)

$$f(B) = T^{-1}f(A)T.$$

Действительно, две подобные матрицы имеют одинаковые минимальные многочлены (из $B = T^{-1}AO$ следует $B^k = T^{-1}A^kT$ (k — 0,1, 2, ...); отсюда для любого многочлена $g(\lambda)$ имеем $g(B)= T^{-1}g(A)T$. Поэтому из $g(A) = 0$ вытекает $g(B)=0$ и наоборот) и, следовательно, функция $f(\lambda)$ принимает одни и те же значения как на спектре матрицы A, так и на спектре матрицы B. Поэтому существует интерполяционный многочлен $r(\lambda)$ такой, что $f(A) = r(A)$, $f(B) = r(B)$. Но тогда из равенства $r(B) = T^{-1}r(A)T$ следует:

$$f(B) = T^{-1}f(A)T.$$

3° *Если A — квазидиагональная матрица:*

$$A = (A_1, A_2, ..., A_u),$$

это

$$f(A) = \{ f(A_1), f(A_2),..., f(A_u)\}.$$

Обозначим через $r(\lambda)$ интерполяционный многочлен Лагранжа-Сильвестра для функции $f(\lambda)$ на спектре матрицы A. Тогда, как легко видеть,

$$f(A) =r(A) = \{r(A_1),r(A_2), ..., r(A_u)\}. \qquad (3.77)$$

С другой стороны, минимальный многочлен $\psi(\lambda)$ для A является аннулирующим многочленом для каждой из *матриц* $A_1,A_2, ..., A_u$. Поэтому из равенства

$$f(\Lambda_{A)} = r(\Lambda_{A)}$$

следует:

$$f(\Lambda_{A1}) = r(\Lambda_{A1}), ... , f(\Lambda_{Au}) = r(\Lambda_{Au}).$$

Поэтому

$$f(A_1) = r(A_1), ... , f(A_u) = r(A_u).$$

и равенство (3.77) можно записать так:

$$f(A) = \{f(A_1),f(A_2,..., f(A_u)\}. \qquad (3.78)$$

Пример 1. Если матрица простой структуры

$$A= T\{(\lambda_1, \lambda_2,...., \lambda_n)\}\ T^{-1},$$

то

$$f\,(A)= T\{\ f\,(\lambda_1),\ f\,(\lambda_2),....,\ f\,(\ \lambda_n)\}\ T^{-1}.$$

f(A) имеет смысл, если функция *f*(λ) определена в точках λ_1, λ_2,...., λ_n.
Пример 2. Матрица *J* имеет следующий квазидиагональный вид:

$$J=\left\|\begin{matrix} \overbrace{\begin{matrix}\lambda_1 & 1 & 0 & \ldots & 0\\ 0 & \lambda_1 & 1 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \ldots & 1\\ 0 & 0 & 0 & \ldots & \lambda_1\end{matrix}}^{\nu_1} & & \\ & \ddots & \\ & & \overbrace{\begin{matrix}\lambda_u & 1 & 0 & \ldots & 0\\ 0 & \lambda_u & 1 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 \ldots \lambda_u & & 1\\ 0 & 0 & 0 \ldots 0 & & \lambda_u\end{matrix}}^{\nu_u}\end{matrix}\right\|.$$

В недиагональных блоках все элементы равны нулю. Согласно формуле (3.78)

$$f(J)=\left\|\begin{matrix} \begin{matrix} f(\lambda_1) & \frac{f'(\lambda_1)}{1!} & \ldots & \frac{f^{(\nu_1-1)}(\lambda_1)}{(\nu_1-1)!}\\ 0 & f(\lambda_1) & \ddots & \vdots\\ \vdots & & \ddots & \frac{f'(\lambda_1)}{1!}\\ 0 & 0 & \ldots & f(\lambda_1)\end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} f(\lambda_u) & \frac{f'(\lambda_u)}{1!} & \ldots & \frac{f^{(\nu_u-1)}(\lambda_u)}{(\nu_u-1)!}\\ 0 & f(\lambda_u) & \ddots & \vdots\\ \vdots & & \ddots & \frac{f'(\lambda_u)}{1!}\\ 0 & 0 & \ldots & f(\lambda_u)\end{matrix}\end{matrix}\right\|.$$

Здесь, как и в матрице J, все элементы в недиагональных блоках равны нулю. В дальнейшем будет установлено, что произвольная матрица $A=\| a_{ik} \|^n_1$ всегда подоба некоторой матрице вида J: $A= TJT^{-1}$. Поэтому (см. 2°) всегда $f(A) = Tf(J)\, T^{-1}$.

3.9. Интерполяционный многочлен Лагранжа-Сильвестра

1. Рассмотрим сначала тот случай, когда характеристическое уравнение $|\lambda E-A|=0$ не имеет кратных корней. Корни этого уравнения — характеристические числа матрицы A — обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_n$. Тогда

$$\psi(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n),$$

и условия (3.76) записываются так:

$$r(\lambda_k)=f(\lambda_k) \quad (k=1,2,\ldots,n).$$

В этом случае $r(\lambda)$ есть обычным интерполяционным многочленом A Лагранжа для функции $f(\lambda)$ в точках $\lambda_1, \lambda_2, \ldots, \lambda_n$:

$$r(\lambda) = \sum_{k=1}^{n} \frac{(\lambda-\lambda_1)\ldots(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})\ldots(\lambda-\lambda_n)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_n)} f(\lambda_k).$$

Согласно определению 2

$$f(A) = r(A) = \sum_{k=1}^{n} \frac{(A-\lambda_1 E)\ldots(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)\ldots(A-\lambda_n E)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_n)} f(\lambda_k).$$

2. Допустим теперь, что характеристический многочлен имеет кратные корни, но минимальный многочлен, являющийся делителем характеристического, имеет только простые корни:

$$\psi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_m).$$

В этом случае (как и в предыдущему) все показатели m_k в (3.71) равны единице, и равенства (3.76) принимают вид

$$r(\lambda_k)=f(\lambda_k) \quad (k=1, 2, \ldots,m).$$

$r(\lambda)$ снова есть обычным интерполяционным многочленом Лагранжа и

$$f(A) = \sum_{k=1}^{m} \frac{(A-\lambda_1 E)\ldots(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)\ldots(A-\lambda_m E)}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_m)} f(\lambda_k).$$

3. Рассмотрим общий случай:

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s} \quad (m_1+m_2+\ldots+m_s=m).$$

Представим правильно-дробную функцию

$$\frac{r(\lambda)}{\psi(\lambda)}$$

в виде суммы простых дробей:

$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s}\left[\frac{\alpha_{k1}}{(\lambda-\lambda_k)^{m_k}} + \frac{\alpha_{k2}}{(\lambda-\lambda_k)^{m_k-1}} + \ldots + \frac{\alpha_{k,m_k}}{\lambda-\lambda_k}\right], \tag{3.79}$$

где α_{kj} $(j$ — $1, 2, ..., m_k;$ $k = 1, 2, ..., s)$ — некоторые числа.

Для определения числителей простых дробей α_{kj} умножим обе части этого равенства на $(\lambda-\lambda_k)^{m_k}$ и обозначим через $\psi^k(\lambda)$ многочлен

$$\frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}}.$$

Получим:

$$\frac{r(\lambda)}{\psi^k(\lambda)} = \alpha_{k1} + \alpha_{k2}(\lambda-\lambda_k) + \ldots + \alpha_{km_k}(\lambda-\lambda_k)^{m_k-1} + (\lambda-\lambda_k)^{m_k}\rho(\lambda) \quad (k=1, 2, \ldots, s), \tag{3.80}$$

где $\rho(\lambda)$ — рациональная функция, регулярная при $\lambda = \lambda_k$ (т.е. не оборащающаяся в ∞ при $\lambda = \lambda_k$).

Отсюда

$$\left.\begin{aligned} \alpha_{k1} &= \left[\frac{r(\lambda)}{\psi^k(\lambda)}\right]_{\lambda=\lambda_k}, \\ \alpha_{k2} &= \left[\frac{r(\lambda)}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k} = r(\lambda_k)\left[\frac{1}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k} + r'(\lambda_k)\frac{1}{\psi^k(\lambda_k)}, \ \ldots \quad (k=1, 2, \ldots, s). \end{aligned}\right\} \tag{3.81}$$

Формулы (3.81) показывают, что числители α_{kj} в правой части равенства (3.79) выражаются через значения многочлена $r(\lambda)$ на спектре матрицы A, а эти значения нам известны: они равны соответствующим значениям функции $f(\lambda)$ и ее производных. Поэтому

$$\alpha_{k1} = \frac{f(\lambda_k)}{\psi^k(\lambda_k)}, \quad \alpha_{k2} = f(\lambda_k)\left[\frac{1}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k} + f'(\lambda_k)\frac{1}{\psi^k(\lambda_k)}, \ \ldots \quad (k=1, 2, \ldots, s). \tag{3.82}$$

Формулы (3.82) можно еще сокращенно записать так:

$$\alpha_{kj} = \frac{1}{(j-1)!}\left[\frac{f(\lambda)}{\psi^k(\lambda)}\right]^{(j-1)}_{\lambda=\lambda_k} \quad (j=1, 2, \ldots, m_k; \quad k=1, 2, \ldots, s). \tag{3.83}$$

После того как все α_{kj} найдены, мы определяем $r(\lambda)$ из следующей формулы, которая полусается умножением обеих частей равенства (3.79) на $\psi(\lambda)$:

$$r(\lambda) = \sum_{k=1}^{s} [\alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \ldots + \alpha_{km_k}(\lambda - \lambda_k)^{m_k - 1}]\, \psi^k(\lambda). \tag{3.84}$$

В этой формуле выражение в квадратных скобках, стоящее в качестве множителя перед $\psi^k(\lambda)$, в силу (3.83) равно сумме первых m_k членов разложения Тейлора по степеням $(\lambda - \lambda_k)$ для функции $f(\lambda)/\psi^k(\lambda)$.

Пример.

$$\psi(\lambda) = (\lambda - \lambda_1)^2 (\lambda - \lambda_2)^3 \quad (m = 5).$$

Тогда

$$\frac{r(\lambda)}{\psi(\lambda)} = \frac{\alpha}{(\lambda - \lambda_1)^2} + \frac{\beta}{\lambda - \lambda_1} + \frac{\gamma}{(\lambda - \lambda_2)^3} + \frac{\delta}{(\lambda - \lambda_2)^2} + \frac{\varepsilon}{\lambda - \lambda_2}.$$

Отсюда

$$r(\lambda) = [\alpha + \beta(\lambda - \lambda_1)](\lambda - \lambda_2)^3 + [\gamma + \delta(\lambda - \lambda_2) + \varepsilon(\lambda - \lambda_2)^2](\lambda - \lambda_1)^2$$

и, следовательно.

$$r(A) = [\alpha E + \beta(A - \lambda_1 E)](A - \lambda_2 E)^3 + [\gamma E + \delta(A - \lambda_2 E) + \varepsilon(A - \lambda_2 E)^2](A - \lambda_1 E)^2$$

α, β, γ, δ, ε найдем из следующих формул:

$$\alpha = \frac{f(\lambda_1)}{(\lambda_1 - \lambda_2)^3}, \quad \beta = -\frac{3}{(\lambda_1 - \lambda_2)^4} f(\lambda_1) + \frac{1}{(\lambda_1 - \lambda_2)^3} f'(\lambda_1),$$

$$\gamma = \frac{f(\lambda_2)}{(\lambda_2 - \lambda_1)^2}, \quad \delta = -\frac{2}{(\lambda_2 - \lambda_1)^3} f(\lambda_2) + \frac{1}{(\lambda_2 - \lambda_1)^2} f'(\lambda_2),$$

$$\varepsilon = \frac{3}{(\lambda_2 - \lambda_1)^4} f(\lambda_2) - \frac{2}{(\lambda_2 - \lambda_1)^3} f'(\lambda_2) + \frac{1}{2}\, \frac{1}{(\lambda_2 - \lambda_1)^2} f''(\lambda_2).$$

Примечание 1. Интерполяционный многочлен Лагранжа-Сильвестра может быть получен предельным переходом из интерполяционного многочлена Лагранжа.

Пусть

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s} \left(m = \sum_{k=1}^{s} m_k \right).$$

Обозначим через $L(\lambda)$ интерполяционный многочлен Лагранжа, построенный для m точек

$$\lambda_1^{(1)}, \lambda_1^{(2)}, \ldots, \lambda_1^{(m_1)};\ \lambda_2^{(1)}, \lambda_2^{(2)}, \ldots, \lambda_2^{(m_2)};\ \ldots;\ \lambda_s^{(1)}, \lambda_s^{(2)}, \ldots, \lambda_s^{(m_s)}.$$

Тогда нетрудно показать, что искомый многочлен Лагранжа-Сильвестра определяется формулой

$$r(\lambda) = \lim_{\substack{\lambda_1^{(1)}, \ldots, \lambda_1^{(m_1)} \to \lambda_1 \\ \cdots\cdots\cdots\cdots \\ \lambda_s^{(1)}, \ldots, \lambda_s^{(m_s)} \to \lambda_s}} L(\lambda).$$

Примечание 2. Пусть $A = \| a_{ik} \|$ — *вещественная матрица,* т.е. матрица с вещественными элементами. Тогда минимальный многочлен $\psi(\lambda)$ имеет вещественные коэффициенты (это следует непосредственно из определения минимального многочлена либо из формулы (3.61)) и его корени, т.е. характеристические числа λ_i, либо вещественны, либо попарно комплексно сопряжены, причем, если $\lambda_g = \overline{\lambda_h}$, то соответствующие кратности равны: $m_q = m_h$. Условимся говорить, что *функция* $f(\lambda)$ *вещественна на спектре матрицы A,* если для вещественного λ_i все ее значение на спектре $f(\lambda_i)$, $f'(\lambda_i)$, ... вещественны, а для двух комплексно сопряженных характеристических чисел λ_h и $\lambda_g = \overline{\lambda_h}$ соответствующие значения на спектре комплексно сопряжены:

$$f(\lambda_g) = \overline{f(\lambda_h)}\ ,\ f'(\lambda_g) = \overline{f'(\lambda_h)}, \ldots$$

(функция, представляемая суммой степенного ряда с вещественными коэффициентами, вещественна на спектре любой матрицы, характеристические числа которой находятся внутри круга сходимости данного ряда). В этом случае *f(A)* — *вещественная* матрица. Действительно, в этом случае согласно формулам (3.82) α_{i1}, α_{i2}, ... - вещественные числа и $\alpha_{g1} = \overline{\alpha}_{h1}$, $\alpha_{g2} = \overline{\alpha}_{h2}$,...; при этом для вещественного λ_i многочлен

$$\psi^u(\lambda) = \frac{\psi(\lambda)}{(\lambda - \lambda_i)^{m_i}}$$

имеет вещественные коэффициенты, а коэффициенты многочленов $\psi^h(\lambda)$ и $\psi^g(\lambda)$ (при $\lambda_g = \overline{\lambda_h}$) — комплексно сопряжены. Поэтому в силу формулы (3.84) интерполяционный многочлен $r(\lambda)$ имеет вещественные коэффициенты. Но тогда *r(A),* а значит и *f(A)=r(A),* — вещественная матрица.

3.10. Составляющие матрицы

Вернемся к формуле (3.84) для $r(\lambda)$. Подставляя в нее выражения (3.82) для коэффициентов α и объединяя члены, содержащие одно и то же значение функции $f(\lambda)$ или какой-либо ее производной, мы представим $r(\lambda)$ в виде

$$r(\lambda)=\sum_{k=1}^{s}[f(\lambda_k)\varphi_{k1}(\lambda)+f'(\lambda_k)\varphi_{k2}(\lambda)+\ldots+f^{(m_k-1)}(\lambda_k)\varphi_{km_k}(\lambda)].$$

Здесь $\varphi_{kj}(\lambda)$ $(j = 1, 2, ..., m_k;\ k = 1, 2, ..., s)$ — легко вычисляемые многочлены от λ степени $< m$. Эти многочлены вполне определяются заданием $\psi(\lambda)$ и не зависят от выбора функции $f(\lambda)$. Число этих многочленов равно числу значений функции $f(\lambda)$ на спектре матрицы A, т.е. равно m [m — степень минимального многочлена $\psi(\lambda)$]. Функция $\varphi_{kj}(\lambda)$ представляет собой интерполяционный многочлен Лагранжа-Сильвестра для функции, у которой все значения на спектре матрицы A равны нулю, за исключением одного $f^{(j-1)}(\lambda_k)$, равного 1.

Из формулы (3.85) следует основная формула для $f(A)$:

$$f(A)=\sum_{k=1}^{s}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ldots+f^{(m_k-1)}(\lambda_k)Z_{km_k}], \tag{3.86}$$

где

$$Z_{kj}=\varphi_{kj}(A)\quad (j=1,2,\ldots,m_k;\ k=1,2,\ldots,s). \tag{3.87}$$

Матрицы Z_{kj} вполне определяются заданием матрицы A и не зависят от выбора функции $f(\lambda)$. В правой части формулы (3.86) функция $f(\lambda)$ представлена только своими значениями на спектре матрицы A.

Матрицы Z_{kj} $(j = 1, 2, ..., m_k;\ k = 1, 2, ..., s)$ будем называть *составляющими матрицами* или *компонентами* данной матрицы A.

Компоненты матрицы A всегда линейно независимы.

Действительно, пусть

$$\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}Z_{kj}=\sum_{k=1}^{s}\sum_{j=1}^{m_k}c_{kj}\varphi_{kj}(A)=0. \tag{3.88}$$

Определим интерполяционный многочлен $r(\lambda)$ из m условий

$$r^{(j-1)}(\lambda_k)=c_{kj}\quad (j=1,2,\ldots,m_k;\ k=1,2,\ldots,s). \tag{3.89}$$

Тогда, согласно формуле (3.85),

$$r(\lambda) = \sum_{k=1}^{s} \sum_{j=1}^{m_k} c_{kj} \varphi_{kj}(\lambda).$$

(3.90)

Из сопоставления формул (3.88) и (3.89) находим

$$r(A) = 0. \qquad (3.91)$$

Но степень интерполяционного многочлена $r(\lambda)$, задаваемого формулой (3.90), меньше m, т.е. меньше степени минимального многочлена $\psi(\lambda)$. Поэтому из равенства (3.91) следует тождество

$$r(\lambda) = 0.$$

Но тогда согласно (3.79)

$$c_{kj} \ (j = 1, 2, ..., m_k; \ k = 1, 2, ..., s),$$

что и требовалось доказать. (Из доказанной линейной независимости матриц $Z_{kj}=\varphi_{kj}(A)$ следует также линейная независимость многочленов $\varphi_{kj}(\lambda)$, $j = 1, 2, ..., m_k; \ k = 1, 2, ..., s$).

Из линейной независимости составляющих матриц Z_{kj} следует, что ни одна из этих матриц не равна нулю. Заметим еще, что любые две из компонентов Z_{kj} перестановочны между собой и с матрицей A, поскольку все они суть скалярные многочлены от A.

Формулой (3.86) для $f(A)$ особенно удобно пользоваться тогда, когда приходится иметь дело с несколькими функциями от одной и той же матрицы A, либо когда функция $f(\lambda)$ зависит не только от λ, но и от некоторого параметра t. В последнем случае в правой части формулы (3.86) компоненты Z_{kj} не зависят от t и параметр t входит только в скалярные коэффициенты при этих матрицах.

В ранее рассмотренном примере, где

$$\psi(\lambda) = (\lambda - \lambda_1)^2 (\lambda - \lambda_2)^3 \quad (m = 5).$$

мы можем $r(\lambda)$ представить в виде

$$r(\lambda) = f(\lambda_1)\varphi_{11}(\lambda) + f'(\lambda_1)\varphi_{12}(\lambda) + f(\lambda_2)\varphi_{21}(\lambda) + f'(\lambda_2)\varphi_{22}(\lambda) + f''(\lambda_2)\varphi_{23}(\lambda),$$

где

$$\varphi_{11}(\lambda) = \left(\frac{\lambda - \lambda_2}{\lambda_1 - \lambda_2}\right)^3 \left[1 - \frac{3(\lambda - \lambda_1)}{\lambda_1 - \lambda_2}\right], \quad \varphi_{12}(\lambda) = \frac{(\lambda - \lambda_1)(\lambda - \lambda_2)^3}{(\lambda_1 - \lambda_2)^3},$$

$$\varphi_{21}(\lambda) = \left(\frac{\lambda - \lambda_1}{\lambda_2 - \lambda_1}\right)^2 \left[1 - \frac{2(\lambda - \lambda_2)}{\lambda_2 - \lambda_1} + \frac{3(\lambda - \lambda_2)^2}{(\lambda_2 - \lambda_1)^2}\right],$$

$$\varphi_{22}(\lambda) = \frac{(\lambda - \lambda_1)^2(\lambda - \lambda_2)}{(\lambda_2 - \lambda_1)^2}\left[1 - \frac{2(\lambda - \lambda_2)}{\lambda_2 - \lambda_1}\right],$$

$$\varphi_{23}(\lambda) = \frac{(\lambda - \lambda_1)^2(\lambda - \lambda_2)^2}{2(\lambda_2 - \lambda_1)^2}.$$

Поэтому

$$f(A)=f(\lambda_1)Z_{11}+f'(\lambda_1)Z_{12}+f(\lambda_2)Z_{21}+f'(\lambda_2)Z_{22}+f''(\lambda)Z_{23},$$

где

$$Z_{11}=\varphi_{11}(A)=\frac{1}{(\lambda_1-\lambda_2)^3}(A-\lambda_2E)^3\left[E-\frac{3}{\lambda_1-\lambda_2}(A-\lambda_1E)\right],$$

$$Z_{12}=\varphi_{12}(A)=\frac{1}{(\lambda_1-\lambda_2)^3}(A-\lambda_1E)(A-\lambda_2E)^3 \quad \text{и т. д.}$$

Если дана матрица *A,* то для конкретного нахождения компонент этой матрицы можно в основной формуле (3.86) положить

$$f(\mu)=\frac{1}{\lambda-\mu}$$

где λ - некоторый параметр. Тогда получим:

$$(\lambda E-A)^{-1}=\frac{C(\lambda)}{\psi(\lambda)}=\sum_{k=1}^{s}\left[\frac{Z_{k1}}{\lambda-\lambda_k}+\frac{1!\,Z_{k2}}{(\lambda-\lambda_k)^2}+\ldots+\frac{(m_k-1)!\,Z_{km_k}}{(\lambda-\lambda_k)^{m_k}}\right], \tag{3.92}$$

где $C(\lambda)$ — приведенная присоединенная матрица для $\lambda E-A$.

(При $f(\mu)=\dfrac{1}{\lambda-\mu}$ имеем $f(A)=(\lambda E-A)^{-1}$. Действительно, $f(A)=r(A)$, где $r(\mu)$ - интерполяционный многочлен Лагранжа-Сильвестра. Из того, что $f(\mu)$ и $r(\mu)$ совпадают на спектре матрицы *A,* следует, что на этом спектре совпадают $(\lambda-\mu)r(\mu)$ и $(\lambda-\mu)f(\mu)=1$. Отсюда

$$(\lambda E-A)\,r(A)=(\lambda E-A)\,f(A)=E).$$

Матрицы $(j-1)!Z_{kj}$ являются числителями простейших дробей в разложении (3.92), и потому по аналогии с разложением (3.79) эти числители могут быть выражены через значения $C(\lambda)$ на спектре матрицы *A* по формулам, подобным (3.81):

$$(m_k-1)!\,Z_{km_k}=\frac{C(\lambda_k)}{\psi^k(\lambda_k)},\quad (m_k-2)!\,Z_{k,m_k-1}=\left[\frac{C(\lambda)}{\psi^k(\lambda)}\right]'_{\lambda=\lambda_k},\ \ldots$$

Отсюда

$$Z_{kj}=\frac{1}{(j-1)!\,(m_k-j)!}\left[\frac{C(\lambda)}{\psi^k(\lambda)}\right]^{(m_k-j)}_{\lambda=\lambda_k}\quad (j=1,2,\ldots,m_k;\ k=1,2,\ldots,s). \tag{3.93}$$

Подставляя в (3.86) вместо составляющих матриц их выражения (3.92), мы можем основную формулу (3.86) представить в виде

$$f(A)=\sum_{k=1}^{s}\frac{1}{(m_k-1)!}\left[\frac{C(\lambda)}{\psi^k(\lambda)}f(\lambda)\right]^{(m_k-1)}_{\lambda=\lambda_k}. \tag{3.94}$$

Пример 1.

$$A = \left\| \begin{matrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{matrix} \right\| \begin{matrix} \mathit{2} \\ \mathit{2}, \\ \mathit{1} \end{matrix} \quad \lambda E - A = \left\| \begin{matrix} \lambda - 2 & 1 & -1 \\ 0 & \lambda - 1 & -1 \\ 1 & -1 & \lambda - 1 \end{matrix} \right\|.$$

В данном случае $\Delta(\lambda) = |\lambda E - A| = (\lambda - 1)^2 (\lambda - 2)$. Поскольку минор элемента [1, 2] в $\lambda E - A$ равен 1, то $D_2(\lambda) = 1$, и потому

$$\psi(\lambda) = \Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2) = \lambda^3 - 4\lambda^2 + 5\lambda - 2,$$

$$\Psi(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu^2 + (\lambda - 4)\mu + \lambda^2 - 4\lambda + 5$$

и

$$C(\lambda) = \Psi(\lambda E, A) = A^2 + (\lambda - 4) A + (\lambda^2 - 4\lambda + 5) E =$$

$$= \left\| \begin{matrix} 3 & -2 & 2 \\ -1 & 2 & 2 \\ -3 & 3 & 1 \end{matrix} \right\| \begin{matrix} \mathit{3} \\ \mathit{3} \\ \mathit{1} \end{matrix} + (\lambda - 4) \left\| \begin{matrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{matrix} \right\| + (\lambda^2 - 4\lambda + 5) \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|.$$

(Курсивом набраны элементы контрольного суммарного столбца. Умножая строки матрицы *A* на суммарный столбец матрицы *B*, получаем элементы суммарного столбца произведения *AB).*

Основная формула в данном случае имеет вид

$$f(A) = f(1) Z_1 + f'(1) Z_2 + f(2) Z_3. \tag{3.95}$$

Полагая здесь

$$f(\mu) = \frac{1}{\lambda - \mu},$$

находим:

$$(\lambda E - A)^{-1} = \frac{C(\lambda)}{\psi(\lambda)} = \frac{Z_{11}}{\lambda - 1} + \frac{Z_{12}}{(\lambda - 1)^2} + \frac{Z_{22}}{\lambda - 2},$$

откуда

$$Z_1 = -C(1) - C'(1), \quad Z_2 = -C(1), \quad Z_3 = C(2).$$

Пользуясь приведенным выше выражением для $C(\lambda)$, вычисляем Z_1, Z_2, Z_3 и подставляем полученные результаты в (3.95):

$$f(A) = f(1) \left\| \begin{matrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{matrix} \right\| + f'(1) \left\| \begin{matrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{matrix} \right\| + f(2) \left\| \begin{matrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{matrix} \right\| =$$

$$= \left\| \begin{matrix} f(1) + f'(1) & -f'(1) & f'(1) \\ f(1) + f'(1) - f(2) & -f'(1) + f(2) & f'(1) \\ f(1) - f(2) & -f(1) + f(2) & f(1) \end{matrix} \right\|.$$

Пример 2. Покажем, как можно определить *f(A),* исходя только из основной формулы. Пусть снова

$$A=\begin{Vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{Vmatrix}, \quad \psi(\lambda)=(\lambda-1)^2(\lambda-2).$$

Тогда

$$(A)=f(1)Z_1+f'(1)Z_2+f(2)Z_3. \tag{3.96}$$

Подставим в формулу (3.96) вместо $f(\lambda)$ последовательно 1, λ - 1, $(\lambda — 1)^2$:

$$Z_1+Z_3=E=\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

$$Z_2+Z_3=A-E=\begin{Vmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{Vmatrix}\begin{matrix} 1 \\ 1, \\ 0 \end{matrix}$$

$$Z_3=(A-E)^2=\begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix}\begin{matrix} 0 \\ 0. \\ 0 \end{matrix}$$

Вычитая из первых двух равенств почленно третье, определим все Z_k. Подставляя в (3.96), получим выражение для *f(A).*

Разобранные примеры иллюстрировали три способа практического нахождения *f(A).* В первом способе мы находили интерполяционный многочлен $r(\lambda)$ и полагали *f(A)=r(A).* Во втором способе мы, пользуясь разложением (3.92), выражали компоненты Z_{kj} в формуле (3.86) через значения приведенной присоединенной матрицы $C(\lambda)$ на спектре матрицы *A*. В третьем способе мы исходили из основной формулы (3.86) и подставляли у нее вместо $f(\lambda)$ последовательно некоторые простейшие многочлены; из полученных линейных уравнений определяли составляющие матрицы Z_{kj}.

Третий способ является практически наиболее удобным. В общем виде его можно сформулировать так:

В формулу (3.86) вместо $f(\lambda)$ подставляем последовательно некоторые многочлены $g_1(\lambda)$, $g_2(\lambda)$, ...,$g_m(\lambda)$:

$$g_i(A)=\sum_{k=1}^{s}\left[g_i(\lambda_k)Z_{k1}+g_i'(\lambda_k)Z_{k2}+\ldots+g_i^{(m_k-1)}(\lambda_k)Z_{k\,m_k}\right]$$
$$(i=1,2,\ldots,m). \tag{3.97}$$

Из m уравнений (3.97) определяем m матриц Z_{kj} и подставляем полученные выражения в (3.86).

Результат исключения Z_{kj} из $(m + 1)$-го равенства (3.97) и (3.86) может быть записан в виде

$$\begin{vmatrix} f(A) & f(\lambda_1) & \dots & f^{(m_1-1)}(\lambda_1) & \dots & f(\lambda_s) & \dots & f^{(m_s-1)}(\lambda_s) \\ g_1(A) & g_1(\lambda_1) & \dots & g_1^{(m_1-1)}(\lambda_1) & \dots & g_1(\lambda_s) & \dots & g_1^{(m_s-1)}(\lambda_s) \\ \cdot & \cdot & & & & & & \\ \cdot & \cdot & & & & & & \\ \cdot & \cdot & & & & & & \\ g_m(A) & g_m(\lambda_1) & \dots & g_m^{(m_1-1)}(\lambda_1) & \dots & g_m(\lambda_s) & \dots & g_m^{(m_s-1)}(\lambda_s) \end{vmatrix} = 0.$$

Раскладывая этот определитель по элементам первого столбца, мы получим искомое выражение для $f(A)$. Здесь при $f(A)$ в качестве множителя будет стоять определитель $\Delta=|\ g^{(j)}_i\ (\lambda_k)\ |$ (в i-й строке определителя Δ стоят значения многочлена $g_i(\lambda)$ на спектре матрицы A; $i = 1, 2, ..., m$). Для того чтобы можно было определить $f(A)$, нужно, чтобы $\Delta \neq 0$. Это будет иметь место, если никакая линейная комбинация (с не равными одновременно нулю коэффициентами) многочленов $g_1(\lambda)$, $g_2(\lambda)$, ..., $g_m(\lambda)$ не обращается сплошь в нуль на спектре матрицы A, т.е. не делится на $\psi(\lambda)$.

Условие $\Delta \neq 0$ всегда будет выполнено, если степени многочленов $g_1(\lambda)$, $g_2(\lambda)$,... , $g_m(\lambda)$ соответственно равны 0, 1, 2, ..., m — 1 (в последнем примере $m = 3$, $g_1(\lambda)= 1$, $g_2(\lambda)= \lambda — 1$, $g_3(\lambda)= (\lambda — 1)^2$).

В заключение отметим, что большие степени матрицы A^n удобно вычислять по основной формуле (3.86), заменяя в ней $f(\lambda)$ на $\lambda^{n.}$

Формулу (3.86) можно использовать и для вычисления обратной матрицы A^{-1}, полагая

$$f(\lambda) = \frac{1}{\lambda}$$

или, что то же, полагая $\lambda = 0$ в формуле (3.92).

Пример. Дана матрица $A = \left\| \begin{matrix} 5 & -4 \\ 4 & -3 \end{matrix} \right\|$. Нужно вычислить элементы степени A^{100}. Минимальный многочлен $\psi(\lambda)=(\lambda — 1)^2$.

Основная формула

$$f(A) = f(1)Z_1 + f'(1)Z_2.$$

Заменяя здесь $f(\lambda)$ на 1, а затем на λ - 1, получим:

$$Z_1 = E, \quad Z_2 = A — E.$$

Поэтому

$$f(A) = f(1)\, E + f'(1)(\, A — E).$$

Полагая $f(\lambda) = \lambda^{100}$, найдем:

$$A^{100} = E + 100(A - E) = \left\| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\| + 100 \left\| \begin{matrix} 4 & -4 \\ 4 & -4 \end{matrix} \right\| = \left\| \begin{matrix} 401 & -400 \\ 400 & -399 \end{matrix} \right\|.$$

3.11. Представление функций от матриц рядами

Пусть дана матрица $A = \| a_{ik} \|^n_1$ с минимальным многочленом

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_s)^{m_s} \left(m = \sum_{k=1}^{s} m_k \right).$$

Пусть, далее, функция $f(\lambda)$ и последовательность функций $f_1(\lambda), f_2(\lambda), \dots ., f_p(\lambda), \dots$ определены на спектре матрицы A.

Мы будем говорить, что *последовательность функций* $f_p(\lambda)$ *при* $p \to \infty$ *стремится к некоторому пределу на спектре матрицы A,* если существуют пределы

$$\lim_{p\to\infty} f_p(\lambda_k), \quad \lim_{p\to\infty} f'_p(\lambda_k), \dots, \lim_{p\to\infty} f_p^{(m_k-1)}(\lambda_k) \qquad (k = 1, 2, \dots, s).$$

Мы будем говорить, что *последовательность функций* $f_p(\lambda)$ *стремится при* $p \to \infty$ *к функции* $f(\lambda)$ *на спектре матрицы A,* и будем писать:

$$\lim_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A),$$

если

$$\lim_{p\to\infty} f_p(\lambda_k) = f(\lambda_k), \quad \lim_{p\to\infty} f'_p(\lambda_k) = f'(\lambda_k), \dots, \lim_{p\to\infty} f_p^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k)$$
$$(k = 1, 2, \dots, s).$$

Основная формула

$$f(A) = \sum_{k=1}^{s} [f(\lambda_k) Z_{k1} + f'(\lambda_k) Z_{k2} + \dots + f^{(m_k-1)}(\lambda_k) Z_{km_k}]$$

выражает $f(A)$ через значение $f(\lambda)$ на спектре матрицы A. Если рассматривать матрицу как вектор в пространстве n^2 измерений $\boldsymbol{R}_{n^2}$, то из основной формулы в силу линейной независимости матриц Z_{kj} следует, что все $f(A)$ (при заданном A) образуют m-мерное подпространство в $\boldsymbol{R}_{n^2}$ с базисом Z_{kj} (j — 1, 2, ..., m_k; k = 1, 2, ..., s). В этом базисе «вектор» $f(A)$ имеет своими координатами m значений функции $f(\lambda)$ на спектре матрицы A.

Эти соображения приводят к следующей теореме:

Теорема 1. *Для того чтобы последовательность матриц $f_p(A)$ при $p \to\infty$ стремилась к некоторому пределу, необходимо и достаточно, чтобы последовательность $f_p(\lambda)$ при $p\to\infty$ на спектре матрицы A стремилась к пределу, т.е. пределы*

$$\lim_{p\to\infty} f_p(A), \quad \lim_{p\to\infty} f_p(\Lambda_A)$$

всегда существуют одновременно. При этом равенство

$$\lim_{p\to\infty} f_p(\Lambda_A) = f(\Lambda_A) \tag{3.98}$$

влечет за собой равенство

$$\lim_{p\to\infty} f_p(A) = f(A) \tag{3.99}$$

и наоборот.

Доказательство. 1. Если значения $f_p(\lambda)$ на спектре матрицы A при $p \to\infty$ стремится к предельным значениям, то из формулы

$$f_p(A) = \sum_{k=1}^{s} [f_p(\lambda_k) Z_{k1} + f_p'(\lambda_k) Z_{k2} + \ldots + f_p^{(m_k-1)}(\lambda_k) Z_{km_k}] \tag{3.100}$$

следует существование предела $\lim\limits_{p\to\infty} f_p(A).$ На основании этой же формулы и формулы (3.86) из (3.98) вытекает (3.99).

2. Обратно, пусть существует $\lim\limits_{p\to\infty} f_p(A).$ Так как m составляющих матриц Z линейно независимы, то мы можем из (3.100) выразить (в виде линейных форм) m значений $f_p(\lambda)$ на спектре матрицы A через m элементов матрицы $f_p(A)$. Отсюда следует существование предела $\lim\limits_{p\to\infty} f_p(\Lambda_A),$ и равенство (3.98) имеет место при наличии равенства (3.99).

Согласно установленной теореме, если последовательность многочленов $g_p(\lambda)$ (p = 1.2, ...) стремится к функции $f(\lambda)$ на спектре матрицы A, то

$$\lim_{p\to\infty} g_p(A) = f(A).$$

Эта формула подчеркивает естественность и общность данного нами определения $f(A)$. $f(\lambda)$ всегда получается предельным переходом из $g_p(A)$ при $p \to\infty$, если только последовательность многочленов $g_p(\lambda)$ сходится к $f(\lambda)$ на спектре матрицы A. Последнее условие является необходимым для существования предела $\lim\limits_{p\to\infty} g_p(A)$ при $p \to\infty$.

Условимся говорить, что *ряд* $\sum_{p=0}^{\infty} u_p(\lambda)$ *сходится на спектре матрицы А к функции* $f(\lambda)$, и будем писать:

$$f(\Lambda_A) = \sum_{p=0}^{\infty} u_p(\Lambda_A), \tag{3.101}$$

если все фигурирующие здесь функции определены на спектре матрицы A и имеют место равенства

$$f(\lambda_k) = \sum_{p=0}^{\infty} u_p(\lambda_k), \quad f'(\lambda_k) = \sum_{p=0}^{\infty} u'_p(\lambda_k), \quad \ldots, \quad f^{(m_k-1)}(\lambda_k) = \sum_{p=0}^{\infty} u_p^{(m_k-1)}(\lambda_k)$$

$$(k = 1, 2, \ldots, s),$$

причем в правых частях этих равенств стоят сходящиеся ряды. Другими словами, если положить

$$s_p(\lambda) = \sum_{q=0}^{p} u_q(\lambda) \quad (p = 0, 1, 2, \ldots),$$

то равенство (3.101) равносильно равенству

$$f(\Lambda_A) = \lim_{p \to \infty} s_p(\Lambda_A). \tag{3.102}$$

Очевидно, что доказанной теореме можно дать следующую эквивалентную формулировку:

Теорема 1′. *Для того чтобы ряд* $\sum_{p=0}^{\infty} u_p(\lambda)$ *сходился к некоторой матрице, необходимо и достаточно, чтобы ряд* $\sum_{p=0}^{\infty} u_p(\lambda)$ *сходился на спектре матрицы А. При этом из равенства*

$$f(\Lambda_A) = \sum_{p=0}^{\infty} u_p(\Lambda_A)$$

следует равенство

$$f(A) = \sum_{p=0}^{\infty} u_p(A)$$

и наоборот.

Пусть дан степенной ряд с кругом сходимости $| \lambda - \lambda_0 | < R$ и суммой $f(\lambda)$:

$$(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p \quad (| \lambda - \lambda_0 | < R). \tag{3.103}$$

Так как степенной ряд можно почленно дифференцировать любое число раз внутри круга сходимости, то ряд (3.103) сходится на спектре любой матрицы, характеристические числа которой попадают внутрь круга сходимости.

Таким образом, имеет место теорема.

Теорема 2. *Если функция* $f(\lambda)$ *разлагается в степенной ряд в круге* $|\lambda - \lambda_0| < r$

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p, \tag{3.104}$$

то это разложение сохраняет силу, если скалярный аргумент λ *заменить любой матрицей* A*, характеристические числа которой лежат внутри круга сходимости.*

Примечание. В этой теореме можно допустить, чтобы характеристическое число λ_k матрицы A попало на периферию круга сходимости, но при этом нужно дополнительно потребовать, чтобы m_k — 1 раз почленно продифференцированный ряд (3.104) в точке $\lambda = \lambda_k$ сходился. Отсюда, как известно, уже следует сходимость j раз продифференцированного ряда (3.104) в точке λ_k к $f^{(j)}(\lambda_k)$ для $j = 0, 1, 2, ..., m_k - 1$.

Из доказанной теоремы вытекают, например, следующие разложения (разложения в первых двух строках имеют место при произвольной матрице A):

$$e^A = \sum_{p=0}^{\infty} \frac{A^p}{p!}, \quad \cos A = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p)!} A^{2p}, \quad \sin A = \sum_{p=0}^{\infty} (-1)^p \frac{A^{2p+1}}{(2p+1)!},$$

$$\operatorname{ch} A = \sum_{p=0}^{\infty} \frac{A^{2p}}{(2p)!}, \qquad \operatorname{sh} A = \sum_{p=0}^{\infty} \frac{A^{2p+1}}{(2p+1)!},$$

$$(E - A)^{-1} = \sum_{p=0}^{\infty} A^p \qquad (|\lambda_k| < 1;\ k = 1, 2, \ldots, s),$$

$$\ln A = \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} (A - E)^p \qquad (|\lambda_k - 1| < 1;\ k = 1, 2, \ldots, s)$$

(здесь под $\ln \lambda$ мы понимаем так называемое главное значение многозначной функции $\operatorname{Ln} \lambda$, т.е. ветвь, для которой $\operatorname{Ln} 1 = 0$).

Формула (3.92) позволяет легко распространить интегральную формулу Коши для аналитических функций на функции от матриц.

Рассмотрим в плоскости комплексного переменного λ правильную область, ограниченную замкнутым контуром Γ и содержащую внутри себя характеристические числа λ_1, ..., λ_n матрицы A. Возьмем произвольную аналитическую функцию $f(\lambda)$, регулярную в этой области (включая границу Г). Тогда по известным формулам Коши

$$f(\lambda_k)=\frac{1}{2\pi i}\int_\Gamma \frac{f(\lambda)}{\lambda-\lambda_k}\,d\lambda,\ f'(\lambda_k)=\frac{1}{2\pi i}\int_\Gamma \frac{f(\lambda)}{(\lambda-\lambda_k)^2}\,d\lambda,\ \ldots,\ f^{(m_k-1)}(\lambda_k)=$$

$$=\frac{(m_k-1)!}{2\pi i}\int_\Gamma \frac{f(\lambda)}{(\lambda-\lambda_k)^{m_k}}\,d\lambda$$

$$(k=1,\ \ldots,\ s).$$

Умножая обе части матричного равенства (3.92) на $\frac{f(\lambda)}{2\pi i}$ и интегрируя вдоль Г (интеграл от матрицы определяем как результат «поэлементного» интегрирования. Поэтому

$$\int_\Gamma (\lambda E-A)^{-1} f(\lambda)\,d\lambda=\left\|\int_\Gamma (\lambda E-A)^{-1}_{ik} f(\lambda)\,d\lambda\right\|_{i,k=1}^{n},$$

где

$$(\lambda E-A)^{-1}_{ik}=\frac{b_{ik}}{\Delta(\lambda)}\quad (i,k=1,\ \ldots,\ n)$$

— элементы матрицы $(\lambda E - A)^{-1}$), получим:

$$\frac{1}{2\pi i}\int_\Gamma (\lambda E-A)^{-1} f(\lambda)\,d\lambda=\sum_{k=1}^{s}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ \ldots\ +f^{(m_k-1)}(\lambda_k)Z_{km_k}],$$

что, согласно основной формуле (3.86), и дает:

$$f(A)=\frac{1}{2\pi i}\int_\Gamma (\lambda E-A)^{-1} f(\lambda)\,d\lambda. \tag{3.105}$$

Эти же срассуждения показывают, что интеграл, который стоит в правой части равенства (3.105), равен нулю (при $f(\lambda)\neq 0$!), если все характеристические числа матрицы A расположены вне контура Г, и равен

$$\sum_{k=1}^{q}[f(\lambda_k)Z_{k1}+f'(\lambda_k)Z_{k2}+\ \ldots\ +f^{(m_k-1)}(\lambda_k)Z_{km_k}]\quad (q<s),$$

если характеристические числа, $\lambda_1, \ldots, \lambda_q$ расположены внутри, а $\lambda_{q+1}, \ldots, \lambda_n$ — вне Г.

Формулу (3.105) можно принять за определение аналитической функции от матрицы.

3.12. Свойства функций от матриц

В этом пункте будет доказано несколько предложений, которые позволяют распространить тождества, связывающие функции от скалярного переменного, на матричные значения аргумента.

1° *Пусть* $G(u_1, u_2, ..., u_l)$ *— многочлен относительно* $u_1, u_2, ..., u_l$; $f_1(\lambda)$, $f_2(\lambda)$, ..., $f_l(\lambda)$ — *функции от* λ, *определенные на спектре матрицы A, и*

$$g(\lambda) \equiv G[f_1(\lambda), f_2(\lambda), \ldots, f_l(\lambda)]. \tag{3.106}$$

Тогда из

$$g(\Lambda_A) = 0$$

следует

$$G[f_1(A), f_2(A), \ldots, f_l(A)] = 0.$$

Действительно, обозначим через $r_1(\lambda)$, $r_2(\lambda)$, ..., $r_l(\lambda)$ интерполяционные многочлены Лагранжа-Сильвестра для $f_1(\lambda)$, $f_2(\lambda)$, ..., $f_l(\lambda)$ и положим:

$$h(\lambda) = G[r_1(\lambda), r_2(\lambda), \ldots, r_l(\lambda)].$$

Тогда из (3.106) вытекает:

$$h(\Lambda_A) = 0.$$

Отсюда следует, что

$$G[f_1(A), f_2(A), \ldots, f_l(A)] = G[r_1(A), r_2(A), \ldots, r_l(A)] = h(A) = 0,$$

что и требовалось доказать.

Согласно предложению 1° из тождества

$$\cos^2 \lambda + \sin^2 \lambda = 1$$

следует для любой матрицы A

$$\cos^2 A + \sin^2 A = E$$

(в данном случае: $G(u_1, u_2) = u_1^2 + u_2^2 - 1$, $f_1(\lambda) = \cos \lambda$, $f_2(\lambda) = \sin \lambda$).

Точно так же для любой матрицы A

$$e^A e^{-A} = E,$$

т.е.

$$e^{-A} = (e^A)^{-1}.$$

Далее, для любой матрицы A

$$e^{tA} = \cos A + i \sin A.$$

Пусть дано неособенная матрица A ($|A| \neq 0$). Обозначим через $\sqrt{\lambda}$ однозначную ветвь многозначной функции $\sqrt{\lambda}$, определенную в области, не содержащей нуля и удерживающей все характеристические числа матрицы A. Тогда имеет смысл $\sqrt{A}$. При этом из $(\sqrt{\lambda})^2 - \lambda = 0$ будет следовать:

$$(\sqrt{\lambda})^2 = A.$$

Пусть $f(\lambda) = 1/\lambda$, и $A = \| a_{ik} \|^n_1$ — неособенная матрица. Тогда функция $f(\lambda)$ определена на спектре матрицы A, и потому в равенстве

$$\lambda f(\lambda) = 1$$

можно заменить f на A:

$$A \cdot f(A) = E,$$

т.е.

$$f(A) = A^{-1}.$$

Обозначая через $r(\lambda)$ интерполяционный многочлен для функции $1/\lambda$, мы сможем обратную матрицу A^{-1} представить в виде многочлена от данной:

$$A^{-1} = r(A).$$

Рассмотрим рациональную функцию $\rho(\lambda) = g(\lambda)/h(\lambda)$, где $g(\lambda)$ и $h(\lambda)$ — взаимно простые многочлены относительно λ. Эта функция определена на спектре матрицы A в том и только в том случае, если характеристические числа матрицы A не являются корнями многочлена $h(\lambda)$, т.е. если $|h(A)| \neq 0$. При выполнении этого условия мы можем в тождестве

$$\rho(\lambda) h(\lambda) = g(\lambda),$$

заменить λ на A:

$$\rho(A) h(A) = g(A).$$

Отсюда

$$\rho(A) = g(A)[h(A)]^{-1} = [h(A)]^{-1} g(A).$$

2° *Если составная функция*

$$g(\lambda) \equiv h[f(\lambda)]$$

определена на спектре матрицы А, то,

$$g(A) = h[f(A)],$$

т.е. $g(A) = h(B)$, *где* $B = f(A)$. При доказанные этого предложения, как и раньше, будем предполагать, что

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s}$$

— минимальный многочлен матрицы A. Тогда значения функции $g(\lambda)$ на спектре матрицы A определяются по формулам (предполагается, что все входящие в эти формулы величины

$$f(\lambda_k), \ldots, f^{(m_k-1)}(\lambda_k), \; h(\mu_k), \ldots, h^{(m_k-1)}(\mu_k) \qquad (k = 1, \ldots, s)$$

имеют смысл)

$$g(\lambda_k) = h(\mu_k), \quad g'(\lambda_k) = h'(\mu_k) f'(\lambda_k), \ldots,$$

$$g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\mu_k)[f'(\lambda_k)]^{(m_k-1)} + \ldots + h'(\mu_k) f^{(m_k-1)}(\lambda_k),$$

(3.107)

где $\mu_k = f(\lambda_k)$ $(k = 1, ..., s)$. Многочлен

$$\chi(\mu) = (\mu - \mu_1)^{m_1} (\mu - \mu_2)^{m_2} \dots (\mu - \mu_s)^{m_s}$$

будет аннулирующим многочленом для матрицы B. Действительно, каждое число λ_k является корнем по крайней мере кратности m_k функции

$$q(\lambda) \equiv \chi[f(\lambda)] = \prod_{k=1}^{s} [f(\lambda) - f(\lambda_k)]^{m_k}.$$

Поэтому

$$q(\Lambda_A)=0$$

и согласно 1°

$$q(A)=\chi[f(A)] = \chi(B) = 0.$$

Поэтому среди значений

$$h(\mu_1),\ h'(\mu_1),\ \dots,\ h^{(m_k-1)}(\mu_k) \qquad (k = 1, \dots, s), \tag{3.108}$$

содержатся все значения функции $h(\mu)$ на спектре матрицы B. Исходя из значений (3.108), построим интерполяционный многочлен $r(\lambda)$ для функции $h(\lambda)$. Тогда, с одной стороны,

$$h(B)= r(B).$$

С другой стороны, как показывают формулы (3.107), функции $g(\lambda)$ и $g_1(\lambda) = r[f(\lambda)]$ будут равны на спектре матрицы A. Поэтому, применяя к разности $g(\lambda) - r[f(\lambda)]$ предложение 1°, получим:

$$g(A)-r[f(A)]=0;$$

но тогда

$$g(A)= r[f(A)]=r(B) = h(B) = h[f(A)],$$

что и требовалось доказать.

Комбинируя предложения 1° и 2°, приходим к следующему обобщению предложения 1°.

3. *Пусть*

$$g(\lambda) \equiv G[f_1(\lambda), f_2(\lambda), \dots, f_l(\lambda)],$$

где функции $f_1(\lambda)$, $f_2(\lambda)$, ..., $f_l(\lambda)$ определены на спектре матрицы A, а функция $G(u_1,u_2, ...,u_l)$ есть результат последовательного применения к величинам $u_1,u_2, ...,u_l$ операций сложения, умножения, умножения на число и замены величины произвольной функцией от нее. Тогда из

$$g(\Lambda_A)=0$$

следует

$$G[f_1(A),f_2(A),...,f_l(A)]=0.$$

Так, например, пусть A — неособенная матрица ($|A| \neq 0$). Обозначим через $\ln\lambda$ однозначную ветвь многозначной функции $\mathrm{Ln}\,\lambda$, определенную в некоторой области, которая не содержит числа 0 и содержащей все характеристические числа матрицы A. Тогда в скалярном тождестве

$$e^{\ln\lambda} - \lambda = 0$$

можно заменить скалярный аргумент λ на матрицу A:

$$e^{\ln A} - A = 0,$$

т.е. $e^{\ln A}=A$. Иначе говоря, матрица $X=\ln A$ удовлетворяет матричному уравнению $e^X=A$, т.е. является «натуральным логарифмом» матрицы A. Беря в качестве $\ln\lambda$ другие однозначные ветви многозначной функции $\mathrm{Ln}\,\lambda$, мы получим другие логарифмы матрицы A. Пусть $A=\|a_{ik}\|^n_1$— вещественная неособенная матрица. Дальше будут установлены необходимые и достаточные условия для того, чтобы вещественная матрица имела вещественный натуральный логарифм. Здесь же мы рассмотрим два частных случая.

1) Матрица A не имеет вещественных отрицательных характеристических чисел. Обозначим через $\ln_0\lambda$ однозначную ветвь функции $\ln\lambda$ в комплексной λ-плоскости с разрезом вдоль отрицательной действительной оси, определяемую равенством

$$\ln_0\lambda = \ln r + i\varphi, \; -\pi < \varphi < \pi \quad (\lambda = re^{i\varphi}).$$

Функция $\ln_0\lambda$ принимает вещественные значения при положительных вещественных λ и комплексно соприженные значения при комплексно сопряженных значениях λ. Поэтому функция $\ln_0\lambda$ вещественна на спектре матрицы A и $\ln_0 A$ — вещественная матрица.

2) $A = B^2$, где B - вещественная матрица (в этом случае матрица A имеет отрицательные характеристические числа, если матрица B имеет чисто мнимые характеристические числа). Наряду с функцией $\ln_0\lambda$ введем в рассмотрение две однозначные ветви функции $\ln\lambda$ в комплексной λ-плоскости с разрезом вдоль положительной действительной оси:

$$\ln_1\lambda = \ln r + i\varphi, \quad 0 \leqslant \varphi < 2\pi \qquad (\lambda = re^{i\varphi}),$$

$$\ln_2\lambda = \ln r + i\varphi, \quad -2\pi < \varphi \leqslant 0 \qquad (\lambda = re^{i\varphi}).$$

Пусть матрица B имеет различные характеристические числа λ_k $(k = 1, ..., s)$. Выберем круговые окрестности G_k точек λ_k $(k = 1, ..., s)$ так, чтобы они не пересекались и не содержали начало $\lambda=0$. В области, составленной из этих окрестностей, определим функцию $f(\lambda)$ равенством

$$f(\lambda) = \ln_0 \lambda^2, \text{ если } \lambda \in G_k \text{ и } \operatorname{Re} \lambda_k \neq 0,$$

$$f(\lambda) = \ln_1 \lambda^2, \text{ если } \lambda \in G_k \text{ и } \operatorname{Re} \lambda_k = 0, \ \operatorname{Im} \lambda_k > 0,$$

$$f(\lambda) = \ln_2 \lambda^2, \text{ если } \lambda \in G_k \text{ и } \operatorname{Re} \lambda_k = 0, \ \operatorname{Im} \lambda_k < 0.$$

Тогда функция $f(\lambda)$ представляет собой однозначную ветвь функции $\ln \lambda^2$, определенную и вещественную на спектре матрицы B. Поэтому $f(B)$ — вещественная матрица и

$$e^{f(B)} = B^2 = A,$$

т.е. матрица $f(B)$ является вещественным натуральным логарифмом матрицы A.

Примечание 1. Если $\boldsymbol{A}$ — линейный оператор в n-мерном пространстве $\boldsymbol{R}$, то $f(\boldsymbol{A})$ определяется совершенно так же, как и $f(A)$:

$$f(\boldsymbol{A}) = r(\boldsymbol{A}),$$

где $r(\lambda)$— интерполяционный многочлен Лагранжа-Сильвестра для $f(\lambda)$ на спектре оператора $\boldsymbol{A}$ (спектр оператора $\boldsymbol{A}$ определяется минимальным аннулирующим многочленом $\psi(\lambda)$ оператора $\boldsymbol{A}$).

Согласно этому определению, если оператору $\boldsymbol{A}$ соответствует матрица $A = \| a_{ik} \|^n_1$ в некотором базисе пространства, то оператору $f(\boldsymbol{A})$ в том же базисе соответствует матрица $f(A)$. Все утверждения и формулировки этого раздела, в которых фигурирует матрица A, остаются в силе и после замены матрицы A оператором $\boldsymbol{A}$.

Примечание 2. Можно определить функцию от матрицы $f(A)$, исходя из характеристического многочлена

$$\Delta(\lambda) = \prod_{k=1}^{s} (\lambda - \lambda_k)^{n_k},$$

заменяя им минимальный многочлен

$$\psi(\lambda) = \prod_{k=1}^{s} (\lambda - \lambda_k)^{m_k}.$$

При этом полагают $f(A)=g(A)$, где $g(\lambda)$ — интерполяционный многочлен степени $< n$ по mod $\Delta(\lambda)$ для функции $f(\lambda)$ (многочлен $g(\lambda)$ не определяется однозначно равенством $f(A) = g(A)$ и условием «степень $< n$»). Формулы (3.86), (3.92) и (3.94) заменяются формулами

$$f(A) = \sum_{k=1}^{s} [f(\lambda_k) \widehat{Z}_{k1} + f'(\lambda_k) \widehat{Z}_{k2} + \ldots + f^{(n_k-1)}(\lambda_k) \widehat{Z}_{kn_k}], \tag{3.109}$$

$$(\lambda E - A)^{-1} = \frac{B(\lambda)}{\Delta(\lambda)} = \sum_{k=1}^{s} \left[\frac{\widehat{Z}_{k1}}{\lambda - \lambda_k} + \frac{1!\, \widehat{Z}_{k2}}{(\lambda - \lambda_k)^2} + \ldots + \frac{(n_k - 1)!\, \widehat{Z}_{kn_k}}{(\lambda - \lambda_k)^{n_k - 1}} \right], \tag{3.110}$$

$$f(A)=\sum_{k=1}^{s}\frac{1}{(n_k-1)!}\left[\frac{B(\lambda)}{\Delta_k(\lambda)}f(\lambda)\right]_{\lambda=\lambda_k}^{(n_k-1)}$$

(3.111)

(Частный случай формулы (3.111), когда $f(\lambda)\equiv\lambda^h$, иногда называют формулой Перона).

где

$$\Delta_k(\lambda)=\frac{\Delta(\lambda)}{(\lambda-\lambda_k)^{n_k}}\quad (k=1, 2, \ldots, s).$$

Однако в формуле (3.109) значение

$$f^{(m_k)}(\lambda_k),\ f^{(m_k+1)}(\lambda_k),\ \ldots,\ f^{(n_k-1)}(\lambda_k)$$

входят лишь фиктивно, поскольку из сопоставления (3.92) с (3.110) следует:

$$\widehat{Z}_{k1}=Z_{k1},\ \ldots,\ \widehat{Z}_{km_k}=Z_{km_k},\quad \widehat{Z}_{k,m_k+1}=\ldots=\widehat{Z}_{kn_k}=0.$$

3.13. Применение функций от матрицы к интегрированию системы линейных дифференциальных уравнений с постоянными коэффициентами

1. Рассмотрим сначала систему однородных линейных дифференциальных уравнений с постоянными коэффициентами первого порядка:

$$\left.\begin{aligned}\frac{dx_1}{dt}&=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n,\\ \frac{dx_2}{dt}&=a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \frac{dx_n}{dt}&=a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n;\end{aligned}\right\}$$

(3.112)

здесь t — независимая переменная, $x_1, x_2, \ldots, x_n$ — неизвестные функции переменной t, a_{ik} ($i, k = 1, 2, \ldots, n$) - комплексные числа.

Введем в рассмотрение квадратную матрицу $A=\|a_{ik}\|^n_1$, составленную из коэффициентов, и столбцевую матрицу $x=(x_1, x_2, \ldots, x_n)$. Тогда система уравнений (3.112) может быть записана в виде одного матричного дифференциального уравнения

$$\frac{dx}{dt}=Ax.$$

(3.113)

Здесь и в дальнейшем под *производной от матрицы* мы понимаем матрицу, получающуюся из данной, путем замены их производными. Поэтому dx/dt — столбцевая матрица с элементами

$$\frac{dx_1}{dt},\ \frac{dx_2}{dt},\ \ldots,\ \frac{dx_n}{dt}.$$

Будем искать решение системы дифференциальных уравнений, удовлетворяющее заданным начальным условиям:

$$x_1|_{t=0} = x_{10},\quad x_2|_{t=0} = x_{20},\ \ldots,\quad x_n|_{t=0} = x_{n0}$$

или в сокращенной записи:

$$x|_{t=0} = x_0 \qquad (3.114)$$

Разложим искомый столбец x в ряд Маклорена по степеням t.

$$x = x_0 + \dot{x}_0 t + \ddot{x}_0 \frac{t^2}{2!} + \ldots \left(\dot{x}_0 = \left.\frac{dx}{dt}\right|_{t=0},\ \ddot{x}_0 = \left.\frac{d^2x}{dt^2}\right|_{t=0} \ \text{и т. д.} \right). \qquad (3.115)$$

Но из (3.113) почленнымм дифференцированием находим:

$$\frac{d^2x}{dt^2} = A\frac{dx}{dt} = A^2x,\quad \frac{d^3x}{dt^3} = A\frac{d^2x}{dt^2} = A^3x,\ \ldots \qquad (3.116)$$

Подставляя в (3.113) и (3.116) значение $t = 0$, получим:

$$\dot{x}_0 = Ax_0,\quad \ddot{x}_0 = A^2 x_0,\ \ldots$$

Теперь ряд (3.115) запишется так:

$$x = x_0 + tAx_0 + (t^{2/2}!)A^2x_0 + \ldots = e^{At}x_0. \qquad (3.117)$$

Непосредственной подстановкой в (3.113) убеждаемся

$$\left(\frac{d}{dt}(e^{At}) = \frac{d}{dt}\left(E + At + \frac{A^2t^2}{2!} + \ldots\right) = A + A^2t + \frac{A^3t^2}{2!} + \ldots = Ae^{At}. \right)$$

в том, что (3.117) является решением дифференциального уравнения (3.113). Полагая в (3.117) $t = 0$, найдем: $x|_{t=0} = x_0$.

Таким образом, формула (3.117) дает нам решение данной системы дифференциальных уравнений, удовлетворяющее начальным условиям (3.114).

Положим в (3.86) $f(\lambda) = e^{\lambda t}$. Тогда

$$e^{At} = \| q_{ik}(t) \|_1^n = \sum_{k=1}^{s} (Z_{k1} + Z_{k2}t + \ldots + Z_{km_k}t^{m_k-1})\, e^{\lambda_k t}. \qquad (3.118)$$

Теперь решение (3.117) может быть записано в следующей форме:

$$\left.\begin{aligned} x_1 &= q_{11}(t)\,x_{10} + q_{12}(t)\,x_{20} + \ldots + q_{1n}(t)\,x_{n0}, \\ x_2 &= q_{21}(t)\,x_{10} + q_{22}(t)\,x_{20} + \ldots + q_{2n}(t)\,x_{n0}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= q_{n1}(t)\,x_{10} + q_{n2}(t)\,x_{20} + \ldots + q_{nn}(t)\,x_{n0}; \end{aligned}\right\} \tag{3.119}$$

здесь x_{10}, x_{20}, ..., x_{n0} — произвольные постоянные, равные начальным значением неизвестных функций x_1, x_2, ..., x_n.

Таким образом, интегрирование данной системы дифференциальных уравнений сведено к вычислению элементов матрицы e^{At}.

Если в качестве начального значения аргумента взять значение $t=t_0$, то формула (3.117) заменится формулой

$$x = e^{A(t-t_0)}x_0. \tag{3.120}$$

Пример.

$$\frac{dx_1}{dt} = 3x_1 - x_2 + x_3,$$

$$\frac{dx_2}{dt} = 2x_1 + x_3,$$

$$\frac{dx_3}{dt} = x_1 - x_2 + 2x_3.$$

Матрица коэффициентов:

$$A = \begin{Vmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{Vmatrix}.$$

Составим характеристический определитель

$$\Delta(\lambda) = -\begin{vmatrix} 3-\lambda & -1 & 1 \\ 2 & -\lambda & 1 \\ 1 & -1 & 2-\lambda \end{vmatrix} = (\lambda-1)(\lambda-2)^2.$$

Наибольший общий делитель миноров второго порядка этого определителя $D_2(\lambda) = 1$. Поэтому

$$\psi(\lambda) = \Delta(\lambda) = (\lambda-1)(\lambda-2)^2.$$

Основная формула:

$$f(A) = f(1)\,Z_1 + f(2)\,Z_2 + f'(2)\,Z_3.$$

Возьмем вместо $f(\lambda)$ последовательно 1, $\lambda - 2$, $(\lambda - 2)^2$. Получим:

$$Z_1 + Z_2 = E = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

$$-Z_1 + Z_3 = A - 2E = \begin{Vmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{Vmatrix} \begin{matrix} 1 \\ 1, \\ 0 \end{matrix}$$

$$Z_1 = (A - 2E)^2 = \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix} \begin{matrix} 0 \\ 0. \\ 0 \end{matrix}$$

Определенные отсюда Z_1, Z_2 и Z_3 подставляем в основную формулу:

$$f(A) = f(1) \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix} + f(2) \begin{Vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{Vmatrix} + f'(2) \begin{Vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{Vmatrix}.$$

Заменяя здесь $f(\lambda)$ на $e^{\lambda t}$, будем иметь:

$$e^{At} = e^t \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{Vmatrix} + e^{2t} \begin{Vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{Vmatrix} + te^{2t} \begin{Vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{Vmatrix} = \begin{Vmatrix} (1+t)\,e^{2t} & -te^{2t} & te^{2t} \\ -e^t + (1+t)\,e^{2t} & e^t - te^{2t} & te^{2t} \\ -e^t + e^{2t} & e^t - e^{2t} & e^{2t} \end{Vmatrix}.$$

Таким образом,

$$x_1 = C_1 (1 + t)\, e^{2t} - C_2 te^{2t} + C_3 te^{2t},$$
$$x_2 = C_1 [-e^t + (1 + t)\, e^{2t}] + C_2 (e^t - te^{2t}) + C_3 te^{2t},$$
$$x_3 = C_1 (-e^t + e^{2t}) + C_2 (e^t - e^{2t}) + C_3 e^{2t},$$

где

$$C_1 = x_{10}, \quad C_2 = x_{20}, \quad C_3 = x_{30}.$$

2. Рассмотрим теперь систему неоднородных линейных дифференциальных уравнений с постоянными коэффициентами:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + f_1(t), \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + f_2(t), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + f_n(t), \end{aligned}\right\} \qquad (3.121)$$

где $f_i(t)(i = 1, 2, ..., n)$ — непрерывные функции в интервале $t_0 \le t \le t_1$. Обозначая через $f_i(t)$ столбцевую матрицу с элементами $f_1(t), f_2(t), ..., f_n(f)$ и снова полагая $A = \|a_{ik}\|^n_1$, мы систему (3.121) запишем так:

$$\frac{dx}{dt} = Ax + f(t). \qquad (3.122)$$

Введем вместо x новый столбец неизвестных функций z, связанный с x соотношением

$$x = e^{At}z. \qquad (3.123)$$

Дифференцируя почленно (3.123) и подставляя полученное выражение для $\frac{dx}{dt}$ в (3.122), найдем:

$$e^{At}\frac{dz}{dt} = f(t). \qquad (3.124)$$

Отсюда

$$z(t) = c + \int_{t_0} e^{-A\tau} f(\tau)\, d\tau \qquad (3.125)$$

(Как уже указывалось для частного случая, если дана матричная функция от скалярного аргумента $B(\tau) = \| b_{ik}(\tau)\|$ ($i = 1,2, ...,m;$ $k= 1,2, ... n$; $t_1 \le \tau \le t_2$), то *интеграл*

$$\int_{t_1}^{t_2} B(\tau)\, d\tau$$

определяется естественным образом:

$$\int_{t_1}^{t_2} B(\tau)\, d\tau = \left\| \int_{t_1}^{t_2} b_{ik}(\tau)\, d\tau \right\| \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n).$$

и потому согласно (3.123)

$$x = e^{At}\left[c + \int_{t_0}^{t} e^{-A\tau} f(\tau)\, d\tau\right] = e^{At}c + \int_{t_0} e^{A(t-\tau)} f(\tau)\, d\tau; \qquad (3.126)$$

здесь c — столбец с произвольными постоянными элементами.

Давая в (3.126) аргументу t значение t_0, найдем:

$$c = e^{-At_0}x_0,$$

и, следовательно, решение (3.126) может быть записано так:

$$x = e^{A(t-t_0)}x_0 + \int_{t_0} e^{A(t-\tau)} f(\tau)\, d\tau. \qquad (3.127)$$

Полагая $e^{At}=\|q_{tj}(t)\|^n{}_1$, мы решение (3.127) можем записать в развернутом виде так:

$$\left.\begin{aligned} x_1 &= q_{11}(t-t_0)\,x_{10} + \ldots + q_{1n}(t-t_0)\,x_{n0} + \\ &\qquad + \int\limits_{t_0} [q_{11}(t-\tau) f_1(\tau) + \ldots + q_{1n}(t-\tau) f_n(\tau)]\,d\tau, \\ &\qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= q_{n1}(t-t_0)\,x_{10} + \ldots + q_{nn}(t-t_0)\,x_{n0} + \\ &\qquad + \int\limits_{t_0}^{t} [q_{n1}(t-\tau) f_1(\tau) + \ldots + q_{nn}(t-\tau) f_n(\tau)]\,q\tau. \end{aligned}\right\} \tag{3.128}$$

3. Рассмотрим в качестве примера движение весомой материальной точки в пустоте близ поверхности Земли с учетом движения Земли. В этом случае, как известно, ускорение точки относительно Земли определяется постоянной силой веса $m\boldsymbol{g}$ и инерционной кориолисовой силой $—2m\boldsymbol{\omega}\times\boldsymbol{v}$ ($\boldsymbol{v}$ — скорость точки относительно Земли, $\boldsymbol{\omega}$ — постоянная угловая скорость Земли) (здесь жирные буквы означают векторы, а символ «×» означает векторное умножение). Поэтому дифференциальное уравнение движения точки имеет вид

$$\frac{d\boldsymbol{v}}{dt} = \boldsymbol{g} - 2\boldsymbol{\omega}\times\boldsymbol{v}. \tag{3.129}$$

Определим линейный оператор $\boldsymbol{A}$ в трехмерном евклидовом пространстве равенством

$$\boldsymbol{A}\boldsymbol{x} = -2\boldsymbol{\omega}\times\boldsymbol{x} \tag{3.130}$$

и напишем вместо (3.129):

$$\frac{d\boldsymbol{v}}{dt} = \boldsymbol{A}\boldsymbol{v} + \boldsymbol{g}. \tag{3.131}$$

Сопоставляя (3.131) с (3.122), легко найдем по формуле (3.127):

$$\boldsymbol{v} = e^{\boldsymbol{A}t}\,\boldsymbol{v}_0 + \int\limits_0^t e^{\boldsymbol{A}t}\,dt \cdot \boldsymbol{g} \qquad (\boldsymbol{v}_0 = \boldsymbol{v}_{t=0}).$$

Интегрируя почленно, определим радиус-вектор движущейся точки:

$$\boldsymbol{r} = \boldsymbol{r}_0 + \int\limits_0 e^{\boldsymbol{A}t}\,dt\,\boldsymbol{v}_0 + \int\limits_0 dt \int\limits_0^t e^{\boldsymbol{A}t}\,dt\,\boldsymbol{g}, \tag{3.132}$$

где

$$\boldsymbol{r}_0 = \boldsymbol{r}_{t=0} \quad \text{и} \quad \boldsymbol{v}_0 = \boldsymbol{v}_{t=0}$$

Подставляя вместо e^{At} сумму ряда

$$\boldsymbol{E}+\boldsymbol{A}\frac{t}{1!}+\boldsymbol{A}^2\frac{t^2}{2!}+\ldots$$

и заменяя оператор ***A*** его выражением с (3.130), будем иметь:

$$\boldsymbol{r}=\boldsymbol{r}_0+\boldsymbol{v}_0t+\frac{1}{2}\boldsymbol{g}t^2-\boldsymbol{\omega}\times\left(\boldsymbol{v}_0t^2+\frac{1}{3}\boldsymbol{g}t^3\right)+\boldsymbol{\omega}\times\left[\boldsymbol{\omega}\times\left(\frac{2}{3}\boldsymbol{v}_0t^3+\frac{1}{6}\boldsymbol{g}t^4\right)\right]+\ldots$$

Считая угловую скорость ω малой величиной (для Земли $\omega\approx 7{,}3\cdot 10^{-5}$ 1/сек.) и отбрасывая члены, которые содержат вторую и высшие степени ω, мы для дополнительного отклонения точки, вызванного вращением Земли, получим приближенную формулу

$$\boldsymbol{d}\approx-\boldsymbol{\omega}\times\left(\boldsymbol{v}_0t^2+\frac{1}{3}\boldsymbol{g}t^3\right).$$

Возвращаясь к точному решению (3.132), вычислим e^{At}. Предварительно установим, что минимальный многочлен оператора ***A*** имеет вид

$$\psi(\lambda)=\lambda(\lambda^2+4\omega^2).$$

Действительно, из (3.130) находим:

$$\boldsymbol{A}^2\boldsymbol{x}=4\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{x})=4(\boldsymbol{\omega x})\boldsymbol{\omega}-4\omega^2\boldsymbol{x},$$

$$\boldsymbol{A}^3\boldsymbol{x}=-2\boldsymbol{\omega}\times\boldsymbol{A}^2\boldsymbol{x}=8\omega^2(\boldsymbol{\omega}\times\boldsymbol{x}).$$

Отсюда и из (3.130) следует, что операторы ***E, A,*** $\boldsymbol{A}^2$ линейно независимы, а

$$\boldsymbol{A}^3+4\omega^2\boldsymbol{A}=0.$$

Минимальный многочлен ψ(λ) имеет просте корней 0, $2\omega i$, — $2\omega i$. Интерполяционный многочлен Лагранжа для e^{At} имеет вид

$$1+\frac{\sin 2\omega t}{2\omega}\lambda+\frac{1-\cos 2\omega t}{4\omega^2}\lambda^2.$$

Тогда

$$e^{At}=\boldsymbol{E}+\frac{\sin 2\omega t}{2\omega}\boldsymbol{A}+\frac{1-\cos 2\omega t}{4\omega^2}\boldsymbol{A}^2.$$

Подставляя это выражение для e^{At} в (3.137) и заменяя оператор ***A*** его выражением (3.130), найдем:

$$\boldsymbol{r}=\boldsymbol{r}_0+\boldsymbol{v}_0t+\boldsymbol{g}\frac{t^2}{2}-\boldsymbol{\omega}\times\left(\frac{1-\cos 2\omega t}{2\omega^2}\boldsymbol{v}_0+\frac{2\omega t-\sin 2\omega t}{4\omega^3}\boldsymbol{g}\right)+$$
$$+\boldsymbol{\omega}\times\left[\boldsymbol{\omega}\times\left(\frac{2\omega t-\sin 2\omega t}{2\omega^3}\boldsymbol{v}_0+\frac{-1+2\omega^2t^2+\cos 2\omega t}{4\omega^4}\boldsymbol{g}\right)\right].$$

(3.133)

Рассмотрим частный случай $\boldsymbol{v}_0=0$. Тогда, раскрывая тройное векторное произведение (по формуле ***a*** × ***(b*** × ***c)*** = ***b (acc)* — *c (ab);*** при этом следует учесть, что вектор ***g*** направлен к центру Земли, откуда

$\boldsymbol{\omega g}$ = —ωg sinφ), получим:

$$r = r_0 + g\frac{t^2}{2} + \frac{2\omega t - \sin 2\omega t}{4\omega^3}(g \times \omega) - \frac{\cos 2\omega t - 1 + 2\omega^2 t^2}{4\omega^3}(g \sin\varphi\omega + \omega g),$$

где φ - географическая широта в данном месте Земли. Член

$$\frac{2\omega t - \sin 2\omega t}{4\omega^3}(g \times \omega)$$

представляет собой отклонение, направленное перпендикулярно к плоскости меридиана на восток, а последнее слагаемое в правой части последней формулы дает отклонение, лежащей в плоскости меридиана и направленное от земной оси (перпендикулярно к ней).

4. Пусть теперь дана следующая система линейных дифференциальных уравнений второго порядка:

$$\left.\begin{aligned} &\frac{d^2x_1}{dt^2} + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ &\frac{d^2x_2}{dt^2} + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ &\frac{d^2x_n}{dt^2} + a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0, \end{aligned}\right\} \qquad (3.134)$$

где a_{ik} $(i, k — 1, 2, ..., n)$ — постоянные коэффициенты. Вводя снова столбец $x=(x_1, x_2, ..., x_n)$ и квадратную матрицу $A=\| a_{ik} \|^n_1$, мы перепишем систему (3.134) в матричном виде

$$\frac{d^2x}{dt^2} + Ax = 0.$$

Рассмотрим сначала случай, когда $|A|\neq 0$. Если $n=1$,т.е. x и A скаляры и $|A|\neq 0$, общее решение уравнения (3.134) может быть записано в виде

$$x = \cos(\sqrt{A}\,t)\,x_0 + (\sqrt{A})^{-1}\sin(\sqrt{A}\,t)\,\dot{x}_0, \qquad (3.135)$$

где

$$x_0 = x_{t=0} \text{ и } \dot{x}_0 = \left(\frac{dx}{dt}\right)_{t=0}.$$

Непосредственной проверкой убеждаемся, что (3.135) представляет собой решение уравнения (3.134) при любом n, когда x — столбец, а A — неособенная квадратная матрица (под $\sqrt{A}$ мы здесь понимаем произвольную матрицу, квадрат которой равен A; $\sqrt{A}$ существует при $| A | \neq 0$).

При этом мы опираемся на формулы

$$\left.\begin{aligned} \cos(\sqrt{A}\,t) &= E - \frac{1}{2!}At^2 + \frac{1}{4!}A^2t^4 - \ldots, \\ (\sqrt{A})^{-1}\sin(\sqrt{A}\,t) &= t - \frac{1}{3!}At^3 + \frac{1}{5!}A^2t^5 - \ldots \end{aligned}\right\} \tag{3.136}$$

Формула (3.135) охватывает все решения системы (3.134), поскольку начальные значения x_0 и $\dot{x}_0$ могут быть выбраны произвольно.

В формулах (3.136) правые части имеют смысл и при $|A| = 0$. Поэтому (3.135) представляет собой общее решение данной системы дифференциальных уравнений и в случае, когда $|A| = 0$, если только под функциями $\cos(\sqrt{A}t)$ и $(\sqrt{A})^{-1}\sin(\sqrt{A}t)$, которые входят в состав этого выражения, понимать правые части формул (3.136).

Предоставляем читателю проверить, что общее решение неоднородной системы

$$\frac{d^2x}{dt^2} + Ax = f(t), \tag{3.137}$$

которое удовлетворяет начальным условиям $x_{t=0} = x_0$ и

$$\left(\frac{dx}{dt}\right)_{t=0} = \dot{x}_0,$$

может быть записанное в виде

$$x = \cos(\sqrt{A}\,t)\,x_0 + (\sqrt{A})^{-1}\sin(\sqrt{A}\,t)\,\dot{x}_0 + \\ + (\sqrt{A})^{-1}\int_0^t \sin[\sqrt{A}(t-\tau)]\,f(\tau)\,d\tau. \tag{3.138}$$

Если в качестве начального момента времени берется $t=t_0$, то в формулах (3.135) и (3.138) следует заменить $\cos(\sqrt{A}t)$ и $\sin(\sqrt{A}t)$ на

$\cos\sqrt{A}(t - t_0)$ и $\sin\sqrt{A}(t - t_0)$, а $\int_0^t$ на $\int_{t_0}^t$.

3.14. Устойчивость движения в случае линейной системы

Пусть $x_1, x_2, \ldots, x_n$ — параметры, которые характеризуют отклонение «возмущенного» движения данной механической системы от исследуемого движения (в этих параметрах исследуемое движение характеризуется постоянными нулевыми значениями $x_1=0, x_2=0,\ldots,x_n=0$. Поэтому при математической трактовке вопроса говорят об устойчивости нулевого решения системы дифференциальных уравнений (3.139)), и пусть эти параметры удовлетворяют системе дифференциальных уравнений первого порядка:

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n, t) \qquad (i = 1, 2, \ldots, n); \tag{3.139}$$

здесь независимая переменная t — время, правые части $f_i(x_1, x_2, \ldots, x_n, t)$ — непрерывные функции в некоторой области значений $x_1, x_2, \ldots, x_n$ (содержащей точку $x_1 = 0$, $x_2 = 0$, ..., $x_n = 0$) при всех $t > t_0$ (t_0 — начальный момент времени).

Введем определение устойчивости движения по Ляпунову.

Исследуемое движение называется *устойчивым*, если для любого числа $\varepsilon > 0$ можно указать число $\delta>0$ такое, что при любых начальных (при $t = t_0$) значениях параметров $x_{10}, x_{20},\ldots, x_{n0}$, меньших по модулю числа δ, параметры $x_1, x_2, \ldots, x_n$ во все время движения ($t \geq t_0$) по модулю меньше ε, т.е. для любого $\varepsilon > 0$ можно указать такое $\delta > 0$, что из

$$|x_{i0}| < \delta \qquad (i = 1,2, \ldots,n) \tag{3.140}$$

следует:

$$|x_i(t)| < \varepsilon \quad (t \geq t_0). \tag{3.141}$$

Если при этом дополнительно при некоторому $\delta_0>0$ всегда $\lim\limits_{t \to +\infty} x_i(t) = 0$ ($i = 1,2, \ldots, n$), сколь скоро $|x_{i0}| < \delta$ ($i = 1,2, \ldots,n$), то исследуемое движение называется *асимптотически устойчивым.*

Рассмотрим теперь линейную систему, т.е. тот частный случай, когда система (3.139) является системой линейных однородных дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\, x_k, \tag{3.142}$$

где $p_{ik}(t)$ — непрерывные функции при $t \geq t_0$ ($i, k = 1,2, \ldots, n$).

В матричной записи система (3.142) запишется так:

$$\frac{dx}{dt} = P(t)\, x; \tag{3.143}$$

здесь x — столбцевая матрица с элементами $x_1, x_2, \ldots, x_n$, а $P(t) = \| p_{ik}(t) \|^n_1$ — матрица коэффициентов.

Обозначим через

$$q_{1j}(t),\ q_{2j}(t)\ ,\ldots,q_{nj}(t) \quad (j=1,2,\ldots, n) \tag{3.144}$$

n линейно независимых решений системы (3.142) (здесь второй индекс j обозначает номер решения).

Матрицу $Q(t)=||q_{ti}||^n{}_1$, столбцами которой являются эти решения, называют *интегральной матрицей* системы (3.142).

Произвольное решение системы линейных однородных дифференциальных уравнений получается как линейная комбинация с постоянными коэффициентами из n линейно независимых решений:

$$x_i = \sum_{j=1}^{n} c_j q_{ij}(t) \qquad (t = 1, 2, \ldots, n),$$

или в матричной записи:

$$x=Q(t)c, \qquad (3.145)$$

где c — столбцевая матрица, элементами которой являются произвольные постоянные $c_1, c_2, \ldots, c_n$.

Выберем теперь специальную интегральную матрицу, для которой

$$Q(t_0) = E; \qquad (3.146)$$

Другими словами, при выборе n линейно независимых решений (3.144) будем исходить из следующих специальных начальных условий (любые начальные условия определяют, и притом однозначно, некоторое решение данной системы):

$$q_{ij}(t_0) = \delta_{ij} = \begin{cases} 0 & (i \neq j), \\ 1 & (i = j) \end{cases} \qquad (i, j = 1, 2, \ldots, n).$$

Тогда, полагая в формуле (3.145) $t — t_0$, из (3.146) найдем:

$$x_0=c,$$

и потому формула (3.145) примет вид

$$x=Q(t)x_0, \qquad (3.147)$$

или в развернутом виде

$$x_i = \sum_{j=1}^{n} q_{ij}(t)\, x_{j0} \qquad (i = 1, 2, \ldots, n). \qquad (3.148)$$

Рассмотрим три случая:

1. $Q(t)$— *ограниченная матрица* в интервале $(t_0, +\infty)$, т.е. существует такое число M, что

$$|q_{ij}(t)| \leqslant M \qquad (t \geqslant t_0;\ i, j = 1, 2, \ldots, n).$$

В этом случае из (3.148) следует:

$$|x_i(t)| \leqslant nM \max |x_{j0}|.$$

Условие устойчивости выполняется. (Достаточно в (3.140), (3.141) взять $\delta<\varepsilon/nM$) .

Движение, характеризуемое нулевым решением $x_1=0$, $x_2=0$, ..., $x_n=0$, *устойчиво.*

2. $\lim\limits_{t\to+\infty} Q(t) = 0.$

В этом случае матрица $Q(t)$ ограничена в интервале $(t_0, +\infty)$, и потому, как уже было выяснено, движение устойчиво. Кроме того, их (3.147) следует, что

$$\lim_{t\to+\infty} x(t) = 0$$

при любому x_0. *Движение асимптотически устойчиво.*

3. $Q(t)$ — *неограниченная матрица* в интервале $(t_0, +\infty)$. Это означает, что по крайней мере одна из функций $q_{tj}(t)$, например $q_{hk}(t)$, не ограничена в интервале $(t_0, +\infty)$. Возьмем начальные условия $x_{10}=0, \ldots, x_{k0} \neq 0, \ldots, x_{n0}= 0$. Тогда

$$x_h(t) = q_{hk}(t)x_{k0}.$$

Каким бы малым по модулю не было x_{k0}, функция $x_h(f)$ будет не ограничена. Условие (3.141) не будет выполняться ни при одном δ. *Движение неустойчиво.*

Рассмотрим теперь частный случай, когда коэффициенты в системе (3.142) - постоянные числа:

$$P(t) = P = \text{const}. \qquad (3.149)$$

В этом случае

$$x = e^{P(t-t_0)} x_0. \qquad (3.150)$$

Сопоставляя (3.150) с (3.147), находим, что в этом случае

$$Q(t) = e^{P(t-t_0)}. \qquad (3.151)$$

Обозначим через

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s}$$

минимальный многочлен матрицы коэффициентов P.

Для исследования интегральной матрицы (3.151) воспользуемся формулой (3.86). В этом случае

$$f(\lambda) = e^{\lambda(t-t_0)}$$

(t рассматривается как параметр),

$$f^{(j)}(\lambda_k) = (t-t_0)^j e^{\lambda_k(t-t_0)}.$$

Формула (3.86) дает:

$$e^{P(t-t_0)} = \sum_{k=1}^{s}[Z_{k1} + Z_{k2}(t-t_0) + \ldots + Z_{km_k}(t-t_0)^{m_k-1}]\, e^{\lambda_k(t-t_0)}. \qquad (3.152)$$

Рассмотрим три случая:

1. Re $\lambda_k \leq 0$ (k = 1, 2, ..., s), причем для тех λ_k, для которых Re $\lambda_k = 0$, соответствующее m_k=1 (т.е. чисто мысленные характеристические числа являются простыми корнями минимального многочлена).

2. Re $\lambda_k<0$ (k = 1,2, ..., s).

3. При некотором k имеем Re $\lambda_k > 0$ или Re $\lambda_k = 0$, но $m_k > 1$.

Из формулы (3.152) следует, что в случае 1 матрица $Q(t)=e^{P(t-t0)}$ ограничена в интервале $(t_0, +\infty)$, в случае 2 $e^{P(t-t0)}\to 0$ при $t\to +\infty$ и в случае 3 матрица $e^{P(t-t0)}$ не ограничена в интервале $(t_0, +\infty)$.

Здесь особого рассмотрения требует лишь тот случай, когда в выражении (3.152) для $e^{P(t-t0)}$ имеется несколько слагаемых максимального роста (при $t\to +\infty$), т.е. с максимальным Re $\lambda_k=\alpha_0\geq 0$ и (при данному Re $\lambda_k=\alpha_0$) с максимальным значением $m_k=m_0$. Тогда выражение (3.152) можно представить в виде

$$e^{P(t-t_0)} = e^{\alpha_0(t-t_0)}(t-t_0)^{m_0-1}\left[\sum_{j=1}^{r} Z_{k_j m_0} e^{i\beta_j(t-t_0)} + (*)\right], \tag{3.153}$$

где β_1, β_2, ..., β_r — различные действительные числа, а (*) обозначает матрицу, которая стремится к нулю при $t\to+\infty$. Из этого представления вытекает, что матрица $e^{P(t-t0)}$ не ограничена при $\alpha_0+ m_0 - 1 > 0$ (иначе говоря, при $\alpha_0 > 0$ либо при $\alpha_0 = 0$, но $m_0 > 1$), поскольку матрица

$$\sum_{j=1}^{r} Z_{k_j m_0} e^{i\beta_j(t-t_0)}$$

не может стремиться к нулю при $t\to +\infty$. В последнему мы убедимся, если докажем, что функция

$$f(t) = \sum_{j=1}^{r} c_j e^{i\beta_j t}, \tag{3.154}$$

где c_j — комплексные числа, а β_j — вещественные и различные между собой числа, может стремиться к нулю при $t\to +\infty$ только в случае $f(t)$ = 0. Но, действительно,

$$\overline{f(t)} = \sum_{j=1}^{r} \bar{c}_j e^{-i\beta_j t}. \tag{3.155}$$

Перемножая почленно равенства (3.154) и (3.155) и интегрируя по t в пределах от 0 до T, получим:

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T |f(t)|^2\,dt = \sum_{j=1}^{r} |c_j|^2. \tag{3.156}$$

Но из

$$\lim_{t \to +\infty} f(t) = 0$$

вытекает, что и

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T |f(t)|^2 dt = 0.$$

Поэтому из равенства (3.156) находим, что $c_1 = c_2 = ... = c_r = 0$, т.е. $f(t) = 0$.

Поэтому в случае 1 движение $(x_1 = 0, x_2 = 0, ..., x_n = 0)$ устойчиво, в случае 2 — асимптотически устойчиво и в случае 3 — неустойчиво. Результаты исследования могут быть сформулированы в виде следующей теоремы:

Теорема 3. *Нулевое решение линейной системы* (3.142) *при* P=const *является устойчивым по Ляпунову, если* 1) *все характеристические числа матрицы* P *имеют отрицательные или нулевые вещественные части,* 2) *все характеристические числа с нулевыми вещественными частями, т.е. чисто мнимые характеристические числа (если таковые имеются), являются простыми корнями минимального многочлена матрицы* P, *и неустойчивым, если хотя бы одно из условий* 1), 2) *не выполняется.*

Нулевое решение линейной системы (3.142) *является асимптотически устойчивым в том и только в том случае, когда все характеристические числа матрицы* P *имеют отрицательные вещественные части.*

Приведенные высшее соображение позволяют высказать суждение о характере интегральной матрицы $e^{P(t-t_0)}$ в общем случае при произвольных характеристических числах постоянной матрицы P.

Теорема 4. *Интегральная матрица* $e^{P(t-t_0)}$ *линейной системы* (3.142) *при* P = const *всегда представима в виде*

$$e^{P(t-t_0)} = Z_{-}(t) + Z_0(t) + Z_{+}(t), \tag{3.157}$$

где 1) $\lim_{t \to +\infty} Z_{-}(t) = 0$, 2) $Z_0(t)$ *либо* =const, *либо является ограниченной матрицей в интервале* $(t_0, t \to +\infty)$, *не имеющей предела при* $t \to +\infty$, 3) $Z_+(t)$ *либо* $\equiv 0$, *либо является неограниченной матрицей в интервале* $(t_0, +\infty)$.

Доказательство. Разобьем в правой части равенства (3.152) все m слагаемых на три группы. Обозначим через $Z_{-}(t)$ сумму всех слагаемых, которые содержат множитель $e^{\lambda_k(t-t_k)}$ с $\text{Re}\,\lambda_k < 0$. Через $Z_+(t)$ обозначим сумму слагаемых, в которых либо $\text{Re}\,\lambda_k > 0$, либо

Re λ_k = 0 при наличии множителя $(t — t_0)^\nu$ при $\nu > 0$. Через $Z_0(t)$ обозначим сумму всех других слагаемых. Приведенные ранее соображения показывают, что

$$\lim_{t \to +\infty} Z_-(t) = 0,$$

а функция $Z_+(t)$ не ограничена, если только она не равна тождественно нулю. Функция же $Z_0(t)$ ограничена. Покажем, что из существования предела

$$\lim_{t \to +\infty} Z_0(t) = B$$

следует, что $Z_0(t)$ = const. Действительно, разность $Z_0(t) — B$ может быть представленная в виде суммы

$$\sum_{j=1}^{r} Z_{k_j m_0} e^{i\beta_j(t-t_0)}$$

из равенства (3.153). Относительно же суммы такого вида выше было показано, что она может иметь предел 0 при $t \to +\infty$ лишь тогда, когда она тождественно равнанулю.

Теорема 4 доказана.

Сводка методов определения функции от матрицы. Рассмотренные методы определения аналитической функции $f(A)$ от матрицы A сводятся к следующим:

1) Подстановка матрицы A в степенной ряд соответствующей скалярной функции $f(\lambda)$ и вычисление частичной суммы членов ряда для $f(A)$, дающей достаточно точное приближение .

2) Использование интерполяционного многочлена $r(\lambda)$ и соотношения $f(A)=r(A)$, вытекающего из теоремы Кэли-Гамильтона.

3) Использование интерполяционного многочлена Лагранжа-Сильвестра с определением числовых коэффициентов μ_{kj} разложения на простые дроби (5).

4) Определение компонент Z_{kj} матрицы A через значения приведенной присоединенной матрицы $C(\lambda)$ на спектре данной матрицы с использованием теоремы Сильвестра.

5) Определение компонент Z_{kj} из системы m уравнений, получаемых на основе теоремы Сиільвестра для m простейших многочленов.

При использовании всех этих методов, кроме первого, необходимо знать собственные значения характеристической матрицы $F(\lambda)=[\lambda E—A]$. Вычисление собственных значений требует решения алгебраического уравнения n-й степени и представляет собой самостоятельную задачу высшей алгебры, которая непосредственно не связана с теорией матриц.

Прямое использование теоремы Сильвестра связано с определением значений присоединенной матрицы на спектре данной матрицы, которые могут быть получены с помощью алгоритма Фаддеева.

При определении компонент Z_{kj} вместо минимального многочлена $\psi(\lambda)$ можно пользоваться характеристическим многочленом $\Delta(\lambda)$, однако при этом некоторые компоненты будут равны нулевым матрицам. Таким образом, тот же результат потребует большего объема вычислений.

3.15. Системы линейных дифференциальных уравнений с переменными коэффициентами.

Пусть дана система линейных однородных дифференциальных уравнений первого порядка:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\, x_k \qquad (i = 1, 2, \ldots, n), \tag{3.158}$$

где $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ — комплексные функции вещественного аргумента t, непрерывные в некотором (конечном или бесконечном) интервале изменения t' (все соотношения этого пункта, в которые входят функции от t, имеют место для данного интервала изменения t).

Полагая

$$P(t) = \| p_{ik}(t) \|_1^n \text{ и } x = (x_1, x_2, \ldots, x_n),$$

систему (3.158) запишем так:

$$\frac{dx}{dt} = P(t)\, x. \tag{3.159}$$

Интегральной матрицей системы (3.158) мы будем называть квадратную матрицу $X(t) = \| x_{ik}(t) \|_1^n$, столбцами которой являются n линейно независимых решений системы.

Так как каждый столбец матрицы X удовлетворяет уравнению (159), то и интегральная матрица X удовлетворяет уравнению

$$\frac{dX}{dt} = P(t)\, X. \tag{3.160}$$

В дальнейшем мы вместо системы (3.158) будем рассматривать матричное уравнение (3.160).

Как мы знаем, из теоремы о существовании и единственности решения системы линейных дифференциальных уравнений следует, что интегральная матрица *X(t)* однозначно определяется, если задано значение этой матрицы при некотором («начальном») значении $t = t_0$, (предполагается, что t_0 принадлежит данному интервалу изменения *t*), $X(t_0) = X_0$. В качестве матрицы X_0 можно взять любую неособенную квадратную матрицу *n*-го порядка. В частном случае, когда $X(t_0) = E$, интегральную матрицу *X(t)* будем называть *нормированной.*

Продифференцируем определитель матрицы *X,* дифференцируя последовательно строки определителя и используя при этом дифференциальные соотношения

$$\frac{dx_{ij}}{dt} = \sum_{k=1}^{n} p_{ik}x_{kj} \qquad (i, j = 1, 2, \ldots, n).$$

Тогда получим:

$$\frac{d\,|X|}{dt} = (p_{11} + p_{22} + \ldots + p_{nn})\,|X|.$$

Отсюда следует известное тождество Якоби

$$|X| = ce^{\int_{t_0}^{t} \mathrm{Sp}\, P\, dt}, \tag{3.161}$$

где *c* — постоянная, а

$$\mathrm{Sp}\, P = p_{11} + p_{22} + \ldots + p_{nn}$$

— след матрицы *P (t).*

Так как определитель | *X* | не может тождественно равняться нулю, то $c \neq 0$. Но тогда из тождества Якоби следует, что определитель | *X* | при любом значении аргумента отличен от нуля

$$|X| \neq 0,$$

т. е. *интегральная матрица при любом значении аргумента является неособенной.*

Если $\tilde{X}(t)$ — неособенное $(|\tilde{X}| \neq 0)$ частное решение уравнения (3.160), то общее решение этого уравнения определяется формулой

$$X = \tilde{X}C, \tag{3.162}$$

где *C* — произвольная постоянная матрица.

Действительно, умножая обе части равенства

$$\frac{d\tilde{X}}{dt} = P\tilde{X} \qquad (3.163)$$

справа на C, убеждаемся, что и матрица $\tilde{X}C$ удовлетворяет уравнению (3.160). С другой стороны, если X — произвольное решение уравнения (3.160), то из (3.163) следует:

$$\frac{dX}{dt} = \frac{d}{dt}(\tilde{X}\cdot\tilde{X}^{-1}X) = \frac{d\tilde{X}}{dt}\tilde{X}^{-1}X + \tilde{X}\frac{d}{dt}(\tilde{X}^{-1}X) = PX + \tilde{X}\frac{d}{dt}(\tilde{X}^{-1}X),$$

откуда в силу (3.160)

$$\frac{d}{dt}(\tilde{X}^{-1}X) = 0$$

и

$$\tilde{X}^{-1}X = \mathrm{const} = C,$$

т. е. имеет место (3.162).

Все интегральные матрицы X системы (158) получаются по формуле (3.162) при

$$|C| \neq 0.$$

Рассмотрим частный случай:

$$\frac{dX}{dt} = AX, \qquad (3.164)$$

где A — постоянная матрица. При этом $\tilde{X} = e^{At}$ есть частное неособенное решение уравнения (3.164) (почленно дифференцируя ряд $e^{At} = \sum_{k=0}^{\infty}\frac{A^k}{A!}t^k$, находим: $\frac{d}{dt}\,e^{At} = Ae^{At}$) и потому общее решение этого уравнения имеет вид

$$X = e^{At}C, \qquad (3.165)$$

где C — произвольная постоянная матрица.

Полагая в (3.165) $t = t_0$, найдем: $X_0 = e^{At_0}C$. Отсюда

$$C = e^{-At_0}X_0$$

и потому формулу (3.165) можно представить в виде

$$X = e^{A(t-t_0)}X_0. \qquad (3.166)$$

Эта формула эквивалентна выведенной ранее формуле (3.120) микромодуля 3.

Рассмотрим еще так называемую систему Коши:

$$\frac{dX}{dt} = \frac{A}{t-a}X \quad (A \text{ — постоянная матрица}). \qquad (3.167)$$

Этот случай сводится к предыдущему заменой аргумента:

$u = \ln(t - a)$.

Поэтому общее решение системы (3.167) выглядит так:

$$X = e^{A \ln(t-a)} C = (t - a)^A C. \qquad (3.168)$$

Функции e^{At} и $(t - a)^A$, встречающиеся в формулах (3.165) и (3.168), могут быть представлены в виде

$$e^{At} = \sum_{k=1}^{s} (Z_{k1} + Z_{k2}t + \ldots + Z_{km_k} t^{m_k - 1}) e^{\lambda_k t}, \qquad (3.169)$$

$$(t-a)^A = \sum_{k=1}^{s} (Z_{k1} + Z_{k2} \ln(t-a) + \ldots + Z_{km_k} [\ln(t-a)]^{m_k - 1}) (t-a)^{\lambda_k}. \qquad (3.170)$$

Здесь

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_s)^{m_s}$$

$$(\lambda_i \neq \lambda_k \text{ при } i \neq k;\ i, k = 1, 2, \ldots, s)$$

— минимальный многочлен матрицы *A,* а $Z_{kj}\ (j = 1, 2, \ldots, m_k;\ k = 1, 2, \ldots, s)$ — линейно независимые постоянные матрицы, являющиеся многочленами от *A* (в правой части формулы (3.169) каждое слагаемое $X_k = (Z_{k1} + Z_{k2}t + \ldots + Z_{km_k} t^{m_k - 1}) e^{\lambda_k t}$ $(k = 1, 2, \ldots, s)$ является решением уравнения (3.164). Действительно, этому уравнению удовлетворяет произведение $g(A) e^{At}$ при любой функции $g(\lambda)$. Но $X_k = f(A) = g(A) e^{At}$, если $f(\lambda) = g(\lambda) e^{\lambda t}$ и $g(\lambda_k) = 1$, а все остальные $m - 1$ значений функции $g(\lambda)$ на спектре матрицы *A* равны нулю).

Замечание. Иногда в качестве интегральной матрицы системы дифференциальных уравнений (3.158) берут матрицу *W,* у которой строки являются линейно независимыми решениями системы. Очевидно, матрица *C* будет транспонированной матрицей для *X:*

$$W = X'.$$

Переходя в обеих частях равенства (160) к транспонированным матрицам, мы вместо (3.160) получим следующее уравнение для *W:*

$$\frac{dW}{dt} = W P(t). \qquad (3.171)$$

В правой части этого уравнения матрица *W* стоит первым множителем, а не вторым, как *X* в уравнении (3.160).

3.16. Преобразование Ляпунова

Допустим теперь, что в системе (3.158) [и в уравнении (3.160)] матрица коэффициентов $P(t) = \| p_{ik}(t) \|_1^n$ — непрерывная ограниченная функция от t в интервале $[t_0, \infty)$ (это означает, что каждая из функций $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ непрерывна и ограничена в интервале $[t_0, \infty)$, т. е. при $t \geq t_0$).

Введем вместо неизвестных функций $x_1, x_2, \ldots, x_n$ новые неизвестные функции $y_1, y_2, \ldots, y_n$ при помощи преобразования

$$x_i = \sum_{k=1}^{n} l_{ik}(t) y_k \qquad (i = 1, 2, \ldots, n). \tag{3.172}$$

На матрицу преобразования

$$L(t) = \| l_{ik}(t) \|_1^n$$

наложим следующие ограничения:

1° $L(t)$ имеет непрерывную производную $\frac{dL}{dt}$ в интервале $[t_0, \infty)$;

2° $L(t)$ и $\frac{dL}{dt}$ ограничены в интервале $[t_0, \infty)$;

3° существует постоянная m такая, что

$$0 < m < \operatorname{mod} | L(t) | \qquad (t \geq t_0),$$

т. е. определитель $| L(t) |$ ограничен по модулю снизу положительной постоянной m.

Преобразование (3.172), в котором матрица коэффициентов $L(t) = \| l_{ik}(t) \|_1^n$ удовлетворяет условиям 1° — 3°, мы будем называть *преобразованием Ляпунова,* а соответствующую матрицу $L(t)$ — *матрицей Ляпунова.*

Примеры. 1. Если $L = \text{const}$ и $| L | \neq 0$, то матрица L удовлетворяет условиям 1°—3°. Следовательно, неособенное преобразование с постоянными коэффициентами всегда является преобразованием Ляпунова.

2. Если $D = \| d_{ik} \|_1^n$ — матрица простой структуры с чисто мнимыми характеристическими числами, то матрица

$$L(t) = e^{Dt}$$

удовлетворяет условиям 1°—3° и потому является матрицей Ляпунова; при этом в формуле (3.169) все $m_k = 1$, а $\lambda_k = i\varphi_k$ (φ_k вещественны, $k = 1, 2, \ldots, s$).

Легко проверить, что из свойств 1° — 3° матрицы *L(t)* следует, что существует обратная матрица $L^{-1}(t)$ и что она удовлетворяет тем же условиям 1° — 3°, т. е. обратное преобразование для преобразования Ляпунова снова является преобразованием Ляпунова. Точно так же проверяется, что два последовательных преобразования Ляпунова в результате снова дают преобразование Ляпунова. Таким образом, преобразования Ляпунова образуют группу. Преобразования Ляпунова обладают следующим важным свойством:

Если при преобразовании (3.172) *система уравнений* (3.158) *переходит в систему*

$$\frac{dy_i}{dt} = \sum_{k=1}^{n} q_{ik}(t)\, y_k, \tag{3.173}$$

нулевое решение которой является устойчивым, асимптотически устойчивым или неустойчивым по Ляпунову, то таким же свойством обладает и нулевое решение исходной системы (3.158).

Другими словами, преобразования Ляпунова не изменяют характеристики нулевого решения (в отношении устойчивости). Поэтому эти преобразования могут быть использованы при исследовании устойчивости для упрощения исходной системы уравнений.

Преобразование Ляпунова устанавливает одно-однозначное соответствие между решениями систем (3.158) и (3.173), при этом линейно независимые решения остаются таковыми и после преобразования. Поэтому преобразование Ляпунова переводит интегральную матрицу *X* системы (3.58) в некоторую интегральную матрицу *Y* системы (3.173), при этом

$$X = L(t)\, Y. \tag{3.174}$$

В матричной записи система (173) имеет вид

$$\frac{dY}{dt} = Q(t)\, Y, \tag{3.175}$$

где $Q(t) = \| q_{ik}(t) \|_1^n$ — матрица коэффициентов системы (3.173).

Подставляя в (3.160) вместо *X* произведение *LY* и сопоставляя полученное уравнение с (3.175), легко найдем следующую формулу, выражающую матрицу *Q* через матрицы *P* и *L:*

$$Q = L^{-1}PL - L^{-1}\frac{dL}{dt}\,. \tag{3.176}$$

Две системы (3.158) и (3.173) или, что то же, (3.160) и (3.175) мы будем называть *эквивалентными* (в смысле Ляпунова), если они переводятся друг в друга преобразованием Ляпунова. Матрицы коэффициентов P и Q эквивалентных систем всегда связаны между собой формулой (3.176), в которой матрица L удовлетворяет условиям 1° — 3°.

3.17. Приводимые системы

Среди систем линейных дифференциальных уравнений первого порядка наиболее простыми и наиболее изученными являются системы с постоянными коэффициентами. Поэтому представляют интерес системы, которые при помощи преобразования Ляпунова могут быть приведены к системам с постоянными коэффициентами. Такие системы А. М. Ляпунов называл *приводимыми.*
Пусть дана приводимая система

$$\frac{dX}{dt} = PX. \tag{3.177}$$

Тогда некоторое преобразование Ляпунова

$$X = L(t)\,Y \tag{3.178}$$

переводит ее в систему

$$\frac{dY}{dt} = AY, \tag{3.179}$$

где A — постоянная матрица. Поэтому система (177) имеет частное решение

$$\tilde{X} = L(t)\,e^{At}. \tag{3.180}$$

Легко видеть, что, и обратно, всякая система (3.177), имеющая частное решение вида (3.180), где $L(t)$ — матрица Ляпунова, а A — постоянная матрица, является приводимой и при этом она приводится к виду (3.179) при помощи преобразования Ляпунова (3.178).

Следуя А. М. Ляпунову, покажем, что *всякая система* (3.177) *с периодическими коэффициентами приводима.*

Пусть в данной системе (3.177) $P(t)$ — непрерывная функция в интервале $(-\infty, +\infty)$ с периодом τ:

$$P(t+\tau) = P(t). \tag{3.181}$$

Заменяя в (3.177) t на $t+\tau$ и используя (181), получим:

$$\frac{dX(t+\tau)}{dt} = P(t)\,X(t+\tau).$$

Таким образом, $X(t+\tau)$, как и $X(t)$, является интегральной матрицей системы (3.177). Поэтому $X(t+\tau)=X(t)V$, где V — некоторая постоянная неособенная матрица. Поскольку $|V|\neq 0$, то можно определить

$$V^{\frac{t}{\tau}}=e^{\frac{t}{\tau}\ln V}.$$

(Здесь $\ln V=f(V)$, где $f(\lambda)$ — какая-либо однозначная ветвь функции ln λ в односвязной области G, содержащей все характеристические числа матрицы V и не содержащей числа 0.)

Эта матричная функция от t, как и $X(t)$, умножается справа на V, если к аргументу прибавить τ. Поэтому «частное»

$$L(t)=X(t)V^{-\frac{t}{\tau}}=X(t)e^{-\frac{t}{\tau}\ln V}$$

является непрерывной периодической функцией с периодом τ:

$$L(t+\tau)=L(t)$$

и с определителем $|L(t)|\neq 0$. Матрица $L(t)$ удовлетворяет условиям 1° — 3° предыдущего параграфа и, следовательно, является матрицей Ляпунова.

С другой стороны, поскольку решение X системы (3.177) представимо в виде

$$X=L(t)e^{\frac{\ln V}{\tau}t},$$

то система (3.177) приводима.

В данном случае преобразование Ляпунова

$$X=L(t)Y,$$

приводящее систему (3.177) к виду

$$\frac{dY}{dt}=\frac{1}{\tau}\ln V\cdot Y,$$

имеет периодические коэффициенты с периодом τ.

А. М. Ляпуновым был установлен важный критерий устойчивости и неустойчивости по первому линейному приближению для нелинейных систем

$$\frac{dx_i}{dt}=\sum_{k=1}^{n}a_{ik}x_k+(**)\quad (i=1,2,\ldots,n), \qquad (3.182)$$

где в правых частях стоят сходящиеся степенные ряды относительно $x_1, x_2, \ldots\ \ldots, x_n$, а (**) обозначает сумму членов этих рядов второго порядка и выше относительно $x_1, x_2, \ldots, x_n$; коэффициенты в a_{ik} $(i,\ k = 1, 2, \ldots, n)$ линейных членах постоянны (коэффициенты при нелинейных членах могут зависеть от *t*. На эти функциональные коэффициенты налагаются известные ограничения).

Критерий Ляпунова. *Нулевое решение системы* (3.182) *будет устойчивым* (*и притом асимптотически), если матрица коэффициентов первого линейного приближения* $A = \| a_{ik} \|_1^n$ *имеет все характеристические числа с отрицательными вещественными частями, и неустойчивым, если хотя бы одно из этих характеристических чисел имеет положительную вещественную часть.*

Приведенные выше рассуждения позволяют использовать этот критерий для систем с периодическими коэффициентами в линейных членах:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\, x_k + (**) \qquad (i = 1, \ldots, n). \quad (3.183)$$

Действительно, на основании предыдущих рассуждений можно при помощи преобразования Ляпунова систему (3.183) привести к виду (3.182), где

$$A = \| a_{ik} \|_1^n = \frac{1}{\tau} \ln V,$$

а *V* — постоянная матрица, на которую умножается интегральная матрица соответствующей линейной системы (3.177) при сдвиге аргумента на τ. Не нарушая общности, можем считать τ > 0. В силу свойств преобразования Ляпунова нулевое решение исходной системы и нулевое решение преобразованной одновременно являются устойчивыми, асимптотически устойчивыми или неустойчивыми. Но характеристические числа λ_i и ν_i $(i = 1, 2, \ldots, n)$ матриц *A* и *V* связаны между собой формулой

$$\lambda_i = \frac{1}{\tau} \ln \nu_i \qquad (i = 1, 2, \ldots, n).$$

Поэтому, применяя критерий Ляпунова к приведенной системе, найдем:

Нулевое решение системы (3.183) *будет асимптотически устойчивым, если все характеристические числа* $\nu_1, \nu_2, \ldots, \nu_n$ *матрицы V по модулю* <1, *и неустойчивым, если хотя бы одно из этих чисел по модулю* >1.

А. М. Ляпунов установил свой критерий устойчивости по линейному приближению для значительно более широкого класса систем, а именно для систем вида (3.182), у которых система линейного приближения не обязательно система с постоянными коэффициентами, но принадлежит к классу систем, названных Ляпуновым правильными.

Класс правильных линейных систем содержит в себе как часть все приводимые системы.

Критерий неустойчивости для случая, когда первое линейное приближение является правильной системой, был установлен Н. Г. Четаевым.

3.18. Каноническая форма приводимой системы. Теорема Еругина

Пусть даны приводимая система (3.177) и эквивалентная ей (в смысле Ляпунова) система

$$\frac{dY}{dt} = AY,$$

где A — постоянная матрица.

Стоит вопрос, *в какой степени матрица A определяется данной системой* (3.177). Этот вопрос можно еще сформулировать так:
В каком случае две системы

$$\frac{dY}{dt} = AY \text{ и } \frac{dZ}{dt} = BZ,$$

где A и B — постоянные матрицы, являются эквивалентными по Ляпунову, т. е. переводятся друг в друга преобразованием Ляпунова?

Для того чтобы ответить на этот вопрос, введем понятие о матрицах, имеющих одну и ту же вещественную часть спектра.

Мы будем говорить, что две матрицы A и B n-го порядка *имеют одну и ту же вещественную часть спектра* в том и только в том случае, когда элементарные делители матриц A и B имеют соответственно вид

$$(\lambda-\lambda_1)^{m_1},\ (\lambda-\lambda_2)^{m_2},\ \ldots,\ (\lambda-\lambda_s)^{m_s} \text{ и } (\lambda-\mu_1)^{m_1},\ (\lambda-\mu_2)^{m_2},\ \ldots,\ (\lambda-\mu_s)^{m_s},$$

где

$$\operatorname{Re}\lambda_k = \operatorname{Re}\mu_k \qquad (k = 1, 2, \ldots, s).$$

Имеет место следующая теорема, установленная Н. П. Еругиным:
Теорема 1 (Еругина). *Две системы*

$$\frac{dY}{dt} = AY \text{ и } \frac{dZ}{dt} = BZ \tag{3.184}$$

(A и B — постоянные матрицы n-го порядка) эквивалентны в смысле Ляпунова в том и только в том случае, если матрицы A и B имеют одну и ту же вещественную часть спектра.

Доказательство. Пусть даны системы (3.184). Приведем матрицу A к нормальной жордановой форме (E_k — единичная матрица, в H_k элементы первой наддиагонали равны единице, а остальные элементы равны нулю; порядок E_k, H_k равен степени k-го элементарного делителя матрицы A, т. е. m_k $(k = 1, 2, \ldots, s)$.)

$$A = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2,\ \ldots,\ \lambda_s E_s + H_s\}\, T^{-1}, \qquad (3.185)$$

где

$$\lambda_k = \alpha_k + i\beta_k \quad (\alpha_k,\ \beta_k \text{ — вещественные числа; } k = 1, 2, \ldots, s). \qquad (3.186)$$

В соответствии с (185) и (186) положим:

$$\left.\begin{aligned} A_1 &= T\{\alpha_1 E_1 + H_1, \alpha_2 E_2 + H_2, \ldots, \alpha_s E_s + H_s\}\, T^{-1}, \\ A_2 &= T\{i\beta_1 E_1, i\beta_2 E_2, \ldots, i\beta_s E_s\}\, T^{-1}. \end{aligned}\right\} \qquad (3.187)$$

Тогда

$$A = A_1 + A_2, \quad A_1 A_2 = A_2 A_1. \qquad (3.188)$$

Определим матрицу $L(t)$ равенством $A(t) = e^{A_2 t}$. $L(t)$— матрица Ляпунова.

Но частное решение первой из систем (3.184) в силу (3.188) имеет вид

$$e^{At} = e^{A_2 t} e^{A_1 t} = L(t)\, e^{A_1 (t)}.$$

Отсюда следует, что первая из систем (3.184) эквивалентна системе

$$\frac{dU}{dt} = A_1 U, \qquad (3.189)$$

где согласно (3.187) матрица A_1 имеет вещественные характеристические числа и ее спектр совпадает с вещественной частью спектра матрицы A.

Аналогично вторую из систем (3.184) заменим эквивалентной

$$\frac{dV}{dt} = B_1 V, \qquad (3.190)$$

где матрица B_1 имеет вещественные характеристические числа и ее спектр совпадает с вещественной частью спектра матрицы B.

Теорема будет доказана, если мы покажем, что *две системы* (3.189) *и* (3.190), *в которых матрицы A_1 и B_1 — постоянные матрицы с вещественными характеристическими числами, могут быть* *эквивалентны лишь в том случае, когда матрицы A_1 и B_1 подобны.* Из этого положения следует теорема 1, поскольку эквивалентность систем (3.189) и (3.190) означает эквивалентность систем (3.184), а подобие матриц A_1 и B_1 означает, что эти матрицы имеют одинаковые элементарные делители и потому матрицы A и B имеют одну и ту же вещественную часть спектра.

Пусть преобразование Ляпунова $U = L_1V$ переводит (3.189) в (3.190). Тогда матрица L_1 удовлетворяет уравнению

$$\frac{dL_1}{dt} = A_1L_1 - L_1B_1. \qquad (3.191)$$

Это матричное уравнение относительно L_1 эквивалентно системе n^2 дифференциальных уравнений относительно n^2 элементов матрицы L_1. Правая часть в (3.191) представляет собой линейную операцию над «вектором» L_1 в пространстве n^2 измерений

$$\frac{dL_1}{dt} = \widehat{F}(L_1) \quad [\widehat{F}(L_1) = A_1L_1 - L_1B_1]. \qquad (3.192)$$

Любое характеристическое число линейного оператора $\hat{F}$ (и соответствующей ему матрицы порядка n^2) представляется в виде разности γ—δ , где γ — характеристическое число матрицы A_1, а δ — характеристическое число матрицы B_1. (В самом деле, пусть Λ_0 — какое-либо характеристическое число оператора $\hat{F}$. Тогда существует матрица $L \neq 0$ такая, что $\widehat{F}(L) = \Lambda_0 L$ или

$$(A_1 - \Lambda_0 E)L = LB_1. \qquad (*)$$

Матрицы

$$A_1 - \Lambda_0 E \text{ и } B_1$$

имеют хотя бы одно общее характеристическое число, так как в противном случаесуществовал бы такой многочлен g (λ), что

$$g(A_1 - \Lambda_0 E) = 0, \quad g(B_1) = E,$$

а это невозможно, поскольку из (*) следует: $g(A_1 - \Lambda_0 E)\cdot L = L \cdot g(B_1)$ и $L \neq 0$. Но если матрицы A_1—$\Lambda_0 E$ и B_1 имеют общее характеристическое число, то Λ_0=γ—δ, где γ и δ — характеристические числа соответственно матриц A_1 и B_1.)

Отсюда следует, что оператор $\hat{F}$ имеет только вещественные характеристические числа.

Обозначим через $\widehat{\psi}(\lambda) = (\lambda - \widehat{\lambda}_1)^{\widehat{m}_1}(\lambda - \widehat{\lambda}_2)^{\widehat{m}_2} \ldots (\lambda - \widehat{\lambda}_u)^{\widehat{m}_u}$ ($\widehat{\lambda}_i$ вещественны: $\widehat{\lambda}_i \neq \widehat{\lambda}_j$ при $i \neq j;\ i, j = 1, 2, \ldots, u$) минимальный многочлен для $\overset{J}{F}$. Тогда решение $L_1(t) = e^{\widehat{F}t}L^{(0)}$ системы (3.192) в силу формулы (169) запишется так:

$$L_1(t) = \sum_{k=1}^{u} \sum_{j=0}^{\widehat{m}_k - 1} L_{kj} t^j e^{\widehat{\lambda}_k t}, \tag{3.193}$$

где L_{kj} — постоянные матрицы n-го порядка. Поскольку матрица $L_1(t)$ ограничена в интервале (t_0, ∞), то как для любого $\widehat{\lambda}_k > 0$, так и при $\widehat{\lambda}_k = 0$ и $j > 0$ соответствующие матрицы $L_{kj} = 0$. Обозначим через $L_-(t)$ сумму всех слагаемых в (3.193), в которых $\widehat{\lambda}_k < 0$. Тогда

$$L_1(t) = L_-(t) + L_0, \tag{3.194}$$

где

$$\lim_{t\to+\infty} L_-(t) = 0, \quad \lim_{t\to+\infty} \frac{dL_-(t)}{dt} = 0, \; L_0 = \text{const}. \tag{3.195}$$

Тогда согласно (3.194) и (3.195)

$$\lim_{t\to+\infty} L_1(t) = L_0,$$

откуда следует, что

$$|L_0| \neq 0,$$

поскольку определитель $|L_1(t)|$ ограничен по модулю снизу.

Подставляя в (3.191) вместо $L_1(t)$ сумму $L_-(t) + L_0$, получим:

$$\frac{dL_-(t)}{dt} - A_1 L_-(t) + B_1 L_-(t) = A_1 L_0 - B_1 L_0,$$

откуда в силу (3.195)

$$A_1 L_0 - L_0 B_1 = 0$$

и, следовательно,

$$B_1 = L_0^{-1} A_1 L_0. \tag{3.196}$$

Обратно, если имеет место (3.196), то преобразование Ляпунова

$$U = L_0 V$$

переводит систему (3.189) в систему (3.190). Теорема доказана.

Из доказанной теоремы вытекает, что *всякая приводимая система* (3.177) *при помощи преобразования Ляпунова X = LY может быть приведена к виду*

$$\frac{dY}{dt} = JY,$$

где J — жорданова матрица с вещественными характеристическими числами. Эта каноническая форма системы заданием матрицы *P(t)* определяется однозначно с точностью до порядка диагональных клеток в *J*.

3.19. Матрицант

Рассмотрим систему дифференциальных уравнений

$$\frac{dX}{dt} = P(t)X, \tag{3.197}$$

где $P(t) = \| p_{ik}(t) \|_1^n$ — непрерывная матричная функция в некотором интервале *(a, b)* изменения аргумента *t ((a, b)* — произвольный интервал (конечный или бесконечный). Все элементы $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ матрицы *P(t)* — комплексные функции вещественного аргумента *t*, непрерывные в интервале *(a, b)*. Все последующее сохраняет силу, если вместо непрерывности потребовать лишь ограниченность и интегрируемость по Риману [в любом конечном подинтервале интервала (*a*, *b*)] всех функций $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$).

Воспользуемся методом последовательных приближений для определения нормированного решения системы (3.197), т. е. решения, обращающегося в единичную матрицу при $t = t_0$ [t_0 — фиксированное число из интервала *(a, b)*]. Последовательные приближения X_k (k = 0, 1, 2, ...) будем находить из рекуррентных соотношений

$$\frac{dX_k}{dt} = P(t)X_{k-1} \quad (k = 1, 2, \ldots),$$

выбирая в качестве приближения X_0 единичную матрицу *E*.

Полагая $X_k(t_0) = E$ $(k = 0, 1, 2, \ldots)$, мы сможем X_k представить в виде

$$X_k = E + \int_{t_0}^{t} P(t)X_{k-1}\,dt \quad (k = 1, 2, \ldots).$$

Таким образом,

$$X_0 = E,\ X_1 = E + \int_{t_0}^{t} P(t)\,dt,\quad X_2 = E + \int_{t_0}^{t} P(t)\,dt + \int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt, \ldots,$$

т. е. X_k (k = 0, 1, 2, ...) есть сумма первых k+1 членов матричного ряда

$$E+\int_{t_0}^{t} P(t)\,dt+\int_{t_0}^{t} P(t)\,dt\int_{t_0}^{t} P(t)\,dt+\dots \qquad (3.198)$$

Для того чтобы доказать, что этот ряд абсолютно и равномерно сходится в любой замкнутой части интервала (a, b) и определяет искомое решение уравнения (3.197), мы построим мажорантный ряд.

Определим неотрицательные функции $g(t)$ и $h(t)$ в интервале (a, b) равенствами (по определению значение функции $g(t)$ при каком-либо из значений t равно наибольшему из n^2 модулей значений $p_{ik}(t)$ $(i, k = 1, 2, \dots, n)$ при том же значении t)

$$g(t)=\max\{|p_{11}(t)|,\ |p_{12}(t)|,\dots,|p_{nn}(t)|\},\ h(t)=\left|\int_{t_0} g(t)\,dt\right|. \qquad (*)$$

Легко проверяется, что функции $g(t)$, а следовательно, и $h(t)$ непрерывны в интервале (a, b)(непрерывность функции $g(t)$ в любой точке t_1 интервала (a, b) следует из того, что разность $g(t) — g(t_1)$ при t, достаточно близком к t_1, всегда совпадает с одной из n^2 разностей $|p_{ik}(t)|-|p_{ik}(t_1)|$ $(i, k = 1, 2, \dots, n)$.)

Каждый из n^2 скалярных рядов, на которые распадается матричный ряд (3.198), мажорируется рядом

$$1+h(t)+\frac{nh^2(t)}{2!}+\frac{n^2h^3(t)}{3!}+\dots \qquad (3.199)$$

Действительно,

$$\left|\left\{\int_{t_0}^{t} P(t)\,dt\right\}_{i,k}\right|=\left|\int_{t_0}^{t} p_{ik}(t)\,dt\right|\leqslant\left|\int_{t_0} g(t)\,dt\right|=h(t),$$

$$\left|\left\{\int_{t_0}^{t} P(t)\,dt\int_{t_0}^{t} P(t)\,dt\right\}_{i,k}\right|=\left|\sum_{j=1}^{n}\int_{t_0}^{t} p_{ij}\,dt\int_{t_0}^{t} p_{jk}(t)\,dt\right|\leqslant$$

$$\leqslant n\left|\int_{t}^{t} g(t)\,dt\int_{t_0}^{t} g(t)\,dt\right|=\frac{nh^2(t)}{2},$$

и т. д.

Ряд (3.199) сходится в интервале (a, b), причем сходится равномерно в любой замкнутой части этого интервала. Отсюда вытекает, что и матричный ряд (3.198) сходится в (a, b) и притом абсолютно и равномерно в любом замкнутом интервале, входящем в (a, b).

Почленным дифференцированием проверяем, что сумма ряда (3.198) представляет собой решение уравнения (3.197); это решение обращается в E при $t = t_0$. Почленное дифференцирование ряда (3.198) допустимо, поскольку ряд, получающийся после дифференцирования, отличается множителем P от ряда (3.198) и, следовательно, как и ряд (3.198), является равномерно сходящимся в любой замкнутой части интервала *(a, b).*

Таким образом, доказана теорема о существовании нормированного решения уравнения (3.197). Это решение будем обозначать через $\Omega_{t_0}^t(P)$ или просто $\Omega_{t_0}^t$. Любое другое решение, как было показано в п. 3.15, имеет вид

$$X = \Omega_{t_0}^t C,$$

где C — произвольная постоянная матрица. Из этой формулы следует, что любое решение, и в частности нормированное, однозначно определяется своим значением при $t = t_0$.
Нормированное решение $\Omega_{t_0}^t$ уравнения (3.197) часто называют *матрицантом.*

Мы показали, что матрицант представим в виде ряда (представление матрицанта в виде такого ряда было впервые получено Пеано)

$$\Omega_{t_0}^t = E + \int_{t_0}^{t} P(t)\,dt + \int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt + \ldots, \qquad (3.200)$$

который сходится абсолютно и равномерно в любом замкнутом интервале, в котором функция P (t) непрерывна.

Отметим некоторые формулы для матрицанта.

1° $$\Omega_{t_0}^t = \Omega_{t_1}^t \Omega_{t_0}^{t_1} \quad (t_0, t_1, t \in (a, b)).$$

Действительно, поскольку $\Omega_{t_0}^t$ и $\Omega_{t_1}^t$ — два решения уравнения (3.197), то

$$\Omega_{t_0}^t = \Omega_{t_1}^t C \quad (C - \text{постоянная матрица}).$$

Полагая здесь $t = t_1$, получим $C = \Omega_{t_0}^{t_1}$.

2° $$\Omega_{t_0}^t(P + Q) = \Omega_{t_0}^t(P)\,\Omega_{t_0}^t(S), \text{ где } S = [\Omega_{t_0}^t(P)]^{-1} Q\,\Omega_{t_0}^t(P).$$

Для вывода этой формулы положим:

$$X = \Omega_{t_0}^t(P), \quad Y = \Omega_{t_0}^t(P + Q)$$

и

$$Y = XZ. \qquad (3.201)$$

Дифференцируя почленно (3.201), найдем:

$$(P + Q)XZ = PXZ + X\frac{dZ}{dt}.$$

Отсюда

$$\frac{dZ}{dt} = X^{-1}QXZ$$

и, следовательно, поскольку из (3.201) следует, что $Z(t_0) = E$,

$$Z = \Omega_{t_0}^t (X^{-1}QX).$$

Подставляя в (3.201) вместо *X, Y, Z* соответствующие матриканты, получаем формулу 2°.

3° $$\ln |\Omega_{t_0}^t (P)| = \int_{t_0}^{t} \mathrm{Sp}\, P\, dt.$$

Эта формула следует из тождества Якоби (3.161), если в него вместо *X (t)* подставить $\Omega_{t_0}^t (P)$.

4° *Если* $A = \| a_{ik} \|_1^n = \mathrm{const}$, *то*

$$\Omega_{t_0}^t (A) = e^{A(t-t_0)}.$$

Введем следующие обозначения. Если $P = \| p_{ik} \|_1^n$, то через mod *P* будем обозначать матрицу

$$\mathrm{mod}\, P = \| \, |p_{ik}| \, \|_1^n.$$

Кроме того, если $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$ — две вещественные матрицы и

$$a_{ik} \leqslant b_{ik} \quad (i, k = 1, 2, \ldots, n),$$

то мы будем писать:

$$A \leqslant B.$$

Тогда из представления (3.40) следует:

5° *Если* $\mathrm{mod}\, P(t) \leqslant Q(t)$, *то*

$$\mathrm{mod}\, \Omega_{t_0}^t (P) \leqslant \Omega_{t_0}^t (Q) \quad (t > t_0).$$

В дальнейшем матрицу *n*-го порядка, у которой все элементы равны единице, будем обозначать через *I*:

$$I = \| 1 \|.$$

Рассмотрим функцию *g*(*t*), определенную в (*). Тогда

$$\mathrm{mod}\, P(t) \leqslant g(t)\, I.$$

Отсюда в силу 5°

$$\mathrm{mod}\, \Omega_{t_0}^t (P) \leqslant \Omega_{t_0}^t (g(t)\, I) \quad (t > t_0). \qquad (3.202)$$

Но $\Omega_{t_0}^t (g(t)\, I)$ есть нормированное решение уравнения

$$\frac{dX}{dt} = g(t)IX.$$

Следовательно, в силу 4° (используя замену независимой переменной t переменной $h = \int_{t_0}^{t} g(t)\,dt.$)

$$\Omega_{t_0}^{t}(g(t)I) = e^{h(t)I} \leqslant \left(1 + h(t) + \frac{nh^2(t)}{2!} + \frac{n^2h^3(t)}{3!} + \dots\right) I,$$

где

$$h(t) = \int_{t_0}^{t} g(t)\,dt.$$

Поэтому из (3.202) следует:

$$6^\circ \qquad \operatorname{mod}\Omega_{t_0}^{t}(P) \leqslant \left(\frac{1}{n}e^{nh(t)} + \frac{n-1}{n}\right) I \leqslant e^{nh(t)} I \quad (t > t_0),$$

где

$$h(t) = \int_{t_0}^{t} g(t)\,dt, \quad g(t) = \max_{1 \leqslant i,\, k \leqslant n} \{\,|\,p_{ik}(t)\,|\,\}.$$

Покажем теперь, как при помощи матриканта выражается общее решение системы линейных дифференциальных уравнений с правыми частями:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\,x_k + f_i(t) \qquad (i = 1, 2, \dots, n); \qquad (3.203)$$

$p_{ik}(t), f_i(t)\ (i, k = 1, 2, \dots, n)$ — непрерывные функции в интервале изменения аргумента t.

Вводя столбцевые матрицы («векторы») $x = (x_1, x_2, \dots, x_n)$ и $f = (f_1, f_2, \dots, f_n)$ и квадратную матрицу $P = \|\,p_{ik}\,\|_1^n$, запишем эту систему так:

$$\frac{dx}{dt} = P(t)\,x + f(t). \qquad (3.204)$$

Будем искать решение этого уравнения в виде

$$x = \Omega_{t_0}^{t}(P)\,z, \qquad (3.205)$$

где z — неизвестный столбец, зависящий от t. Подставим это выражение для x в (204), получим:

$$p\Omega_{t_0}^{t}(P)\,z + \Omega_{t_0}^{t}(P)\frac{dz}{dt} = P\Omega_{t_0}^{t}(P)\,z + f(t),$$

откуда

$$\frac{dz}{dt} = [\Omega_{t_0}^t(P)]^{-1} f(t).$$

Интегрируя, находим:

$$z = \int_{t_0}^{t} [\Omega_{t_0}^{\tau}(P)]^{-1} f(\tau)\, d\tau + c,$$

где c — произвольный постоянный вектор. Подставим это выражение в (3.205), получим:

$$x = \Omega_{t_0}^t(P) \int_{t_0}^{t} [\Omega_{t_0}^{\tau}(P)]^{-1} f(\tau)\, d\tau + \Omega_{t_0}^t(P)\, c. \qquad (3.206)$$

Давая t значение t_0, найдем: $x(t_0)=c$. Поэтому формула (3.206) принимает вид

$$x = \Omega_{t_0}^t(P)\, x(t_0) + \int_{t_0}^{t} K(t, \tau) f(\tau)\, d\tau, \qquad (3.207)$$

где

$$K(t, \tau) = \Omega_{t_0}^t(P)\, [\Omega_{t_0}^{\tau}(P)]^{-1}$$

— так называемая матрица Коши.

3.21. Мультипликативный интеграл.

Рассмотрим матрицант $\Omega_{t_0}^t(P)$. Разобьем основной интервал (t_0, t) на n частей, введя промежуточные точки $t_1, t_2, ..., t_{n-1}$, и положим

$$\Delta t_k = t_k - t_{k-1} \quad (k = 1, 2, \ldots, n;\ t_n = t).$$

Тогда на основании свойства 1° матрицанта (см. предыдущий параграф)

$$\Omega_{t_0}^t = \Omega_{t_{n-1}}^t \ldots \Omega_{t_1}^{t_2} \Omega_{t_0}^{t_1}. \qquad (3.208)$$

Выберем в интервале (t_{k-1}, t_k) промежуточную точку $\tau_k\ (k = 1, 2, \ldots, n)$. Тогда, считая Δt_k малыми величинами первого порядка, при вычислении $\Omega_{t_{k-1}}^{t_k}$ с точностью до малых второго порядка можно принять $P(t) \approx \mathrm{const} = P(\tau_k)$.

Тогда

$$\Omega_{t_{k-1}}^{t_k} = e^{P(\tau_k)\Delta t_k} + (**) = E + P(\tau_k)\,\Delta t_k + (**); \tag{3.209}$$

здесь символом (**) мы обозначаем сумму членов, начиная со второго порядка малости.

Из (3.208) и (3.209) находим:

$$\Omega_{t_0}^{t} = e^{P(\tau_n)\Delta t_n} \ldots e^{P(\tau_2)\Delta t_2} e^{P(\tau_1)\Delta t_1} + (*) \tag{3.210}$$

и

$$\Omega_{t_0}^{t} = [E + P(\tau_n)\,\Delta t_n] \ldots [E + P(\tau_2)\,\Delta t_2]\,[E + P(\tau_1)\,\Delta t_1] + (*). \tag{3.211}$$

Переходя к пределу при неограниченном увеличении числа интервалов разбиения и стремлении к нулю длин этих интервалов (при предельном переходе малые члены (*) исчезают (эти рассуждения могут быть уточнены путем оценки членов, обозначенных нами через (*)), получаем точные предельные формулы

$$\Omega_{t_0}^{t}(P) = \lim_{\Delta t_k \to 0} [e^{P(\tau_n)\Delta t_n} \ldots e^{P(\tau_2)\Delta t_2} e^{P(\tau_1)\Delta t_1}] \tag{3.212}$$

и

$$\Omega_{t_0}^{t}(P) = \lim_{\Delta t_k \to 0} [E + P(\tau_n)\,\Delta t_n] \ldots [E + P(\tau_2)\,\Delta t_2]\,[E + P(\tau_1)\,\Delta t_1]. \tag{3.213}$$

Выражение, стоящее под знаком предела в правой части последнего равенства, представляет собой *интегральное произведение* (аналог интегральной суммы для обычного интеграла). Предел его мы назовем *мультипликативным интегралом* и обозначим символом

$$\widehat{\int}_{t_0}^{t} [E + P(t)\,dt] = \lim_{\Delta t_k \to 0} [E + P(\tau_n)\,\Delta t_n] \qquad [E + P(\tau_1)\,\Delta t_1]. \tag{3.214}$$

Формула (3.213) дает представление матриканта в виде мультипликативного интеграла

$$\Omega_{t_0}^{t}(P) = \widehat{\int}_{t_0}^{t} (E + P\,dt), \tag{3.215}$$

а равенства (3.210) и (3.211) могут быть использованы для приближенного вычисления матриканта.

Мультипликативный интеграл ввел впервые Волътерра в 1887 г. На базе этого понятия Вольтерра построил своеобразное инфинитезимальное исчисление для матричных функций.

Вся специфика мультипликативного интеграла связана с неперестановочностью между собой различных значений подинтегральной матричной функции P (r). В частном случае, когда все эти значения перестановочны между собой

$$P(t')\,P(t'') = P(t'')\,P(t') \qquad (t', t'' \in (t_0, t)),$$

мультипликативный интеграл, как это видно из (3.212) и (3.215), сводится к матрице

$$e^{\int_{t_0}^{t} P(t)\,dt}.$$

Введем теперь *мультипликативную производную*

$$D_t X = \frac{dX}{dt} X^{-1}. \qquad (3.216)$$

Операции взаимно обратны:

$$D_t \text{ и } \widehat{\int}_{t_0}^{t}$$

Если

$$D_t X = P,$$

то

$$X = \widehat{\int}_{t_0}^{t} (E + P\,dt) \cdot C \qquad (C = X(t_0))$$

и наоборот (здесь произвольная постоянная матрица C является аналогом аддитивной произвольной постоянной в обычном неопределенном интеграле).

Последняя формула может быть еще записана так (аналог формулы

$$\int_{t_0}^{t} P\,dt = X(t) - X(t_0)$$

в случае, когда $\frac{dX}{dt} = P$.)

$$\widehat{\int}_{t_0}^{t} (E + P\,dt) = X(t)\,X(t_0)^{-1}. \qquad (3.217)$$

Предлагаем читателю проверить справедливость следующих дифференциальных и интегральных формул:

Дифференциальные формулы

$$\text{I.}\quad D_t(XY) = D_t(X) + XD_t(Y)X^{-1},$$
$$D_t(XC) = D_t(X),$$
$$D_t(CY) = CD_t(Y)C^{-1} \qquad \begin{pmatrix} C \text{ — постоянная} \\ \text{матрица} \end{pmatrix}.$$

$$\text{II.}\quad D_t(X') = X'(D_tX)'X'^{-1}.$$

$$\text{III.}\quad D_t(X^{-1}) = -X^{-1}D_t(X)X = -(D_t(X'))',$$
$$D_t(X'^{-1}) = -(D_t(X))'.$$

Интегральные формулы

$$\text{IV.}\qquad \widehat{\int_{t_0}^{t}}(E + P\,dt) = \widehat{\int_{t_1}^{t}}(E + P\,dt)\widehat{\int_{t_0}^{t_1}}(E + P\,dt).$$

$$\text{V.}\qquad \widehat{\int_{t_0}^{t}}(E + P\,dt) = \left[\widehat{\int_{t}^{t_0}}(E + P\,dt)\right]^{-1}.$$

$$\text{VI.}\qquad \widehat{\int_{t_0}^{t}}(E + CPC^{-1}\,dt) = C\widehat{\int_{t_0}^{t}}(E + P\,dt)C^{-1} \qquad \begin{pmatrix} C \text{ — постоянная} \\ \text{матрица} \end{pmatrix}.$$

$$\text{VII.}\qquad \widehat{\int_{t_0}^{t}}[E + (Q + D_tX)\,dt] = X(t)\widehat{\int_{t_0}^{t}}(E + X^{-1}QX\,dt)\,X(t_0)^{-1}.$$

Выведем еще важную формулу, дающую оценку модуля разности между двумя мультипликативными интегралами:

$$\text{VIII.}\quad \operatorname{mod}\left[\widehat{\int_{t_0}^{t}}(E + P\,dt) - \widehat{\int_{t_0}^{t}}(E + Q\,dt)\right] \leqslant \frac{1}{n}e^{nq(t-t_0)}(e^{nd(t-t_0)} - 1)I\ (t > t_0),$$

если

$$\operatorname{mod} Q \leqslant qI, \qquad \operatorname{mod}(P - Q) \leqslant d\cdot I, \qquad I = \|1\|$$

(q, d — неотрицательные числа, n — порядок матриц P и Q). Обозначим через D разность $P - Q$. Тогда

$$P = Q + D, \qquad \operatorname{mod} D \leqslant d\cdot I.$$

Рассматривая мультипликативный интеграл как матрицант и пользуясь разложением (3.200) матрицанта в ряд, найдем:

$$\widehat{\int_{t_0}^{t}}[E + (Q + D)\,dt] - \widehat{\int_{t_0}^{t}}(E + Q\,dt) =$$
$$= \int_{t_0}^{t} D\,dt + \int_{t_0}^{t} D\,dt\int_{t_0}^{t} Q\,dt + \int_{t_0}^{t} Q\,dt\int_{t_0}^{t} D\,dt + \int_{t_0}^{t} D\,dt\int_{t_0}^{t} D\,dt + \ldots$$

Из этого разложения видно, что

$$\operatorname{mod}\left\{\widehat{\int_{t_0}^{t}}[E+(Q+D)\,dt]-\widehat{\int_{t_0}^{t}}(E+Q\,dt)\right\}\leqslant$$

$$\leqslant\widehat{\int_{t_0}^{t}}E+(\operatorname{mod}Q+\operatorname{mod}D)\,dt-\widehat{\int_{t_0}^{t}}(E+\operatorname{mod}Q)\,dt\leqslant$$

$$\leqslant\widehat{\int_{t_0}^{t}}[E+(q+d)I\,dt]-\widehat{\int_{t_0}^{t}}[E+qI]\,dt=e^{(q+d)I(t-t_0)}-e^{qI(t-t_0)}=$$

$$=e^{qI(t-t_0)}\left(e^{d\cdot I(t-t_0)}-E\right)\leqslant\frac{1}{n}\,e^{nq(t-t_0)}[e^{nd(t-t_0)}-1]\,I.$$

Пусть теперь матрицы P и Q зависят от некоторого параметра α:

$$P=P(t,\alpha),\quad Q=Q(t,\alpha),$$

и пусть

$$\lim_{\alpha\to\alpha_0}P(t,\alpha)=\lim_{\alpha\to\alpha_0}Q(t,\alpha)=P_0(t),$$

причем стремление к пределу равномерно относительно t в рассматриваемом интервале $(t_0,\ t)$. Допустим, что, кроме того, при $\alpha\to\alpha_0$ матрица $Q(t,\alpha)$ по модулю ограничена матрицей qI, где q — положительная постоянная. Тогда, полагая

$$d(\alpha)=\max_{\substack{1\leqslant i,\,k\leqslant n\\ t_0\leqslant\tau\leqslant t}}\left|p_{ik}(\tau,\alpha)-q_{ik}(\tau,\alpha)\right|,$$

будем иметь:

$$\lim_{\alpha\to\alpha_0}d(\alpha)=0$$

Поэтому из формулы VIII следует:

$$\lim_{\alpha\to\alpha_0}\left[\widehat{\int_{t_0}^{t}}(E+P\,dt)-\widehat{\int_{t_0}^{t}}(E+Q\,dt)\right]=0.$$

В частности, если Q не зависит от α $[Q(t,\alpha)=P_0(t)]$, мы получаем:

$$\lim_{\alpha\to\alpha_0}\widehat{\int_{t_0}^{t}}[E+P(t,\alpha)\,dt]=\widehat{\int_{t_0}^{t}}[E+P_0(t)\,dt],$$

где

$$P_0(t)=\lim_{\alpha\to\alpha_0}P(t,\alpha).$$

3.22. Дифференциальные системы в комплексной области.

Рассмотрим систему дифференциальных уравнений

$$\frac{dx_i}{dz} = \sum_{k=1}^{n} p_{ik}(z)\, x_k. \tag{3.218}$$

Здесь данные функции $p_{ik}(z)$ и искомые функции $x_i(z)$ $(i, k = 1, 2, \ldots, n)$ предполагаются однозначными аналитическими функциями комплексного аргумента z, регулярными в некоторой области *G* комплексной z-плоскости.

Вводя квадратную матрицу $P(z) = \| p_{ik}(z) \|_1^n$ и столбцевую матрицу $x = (x_1, x_2, \ldots, x_n)$, мы, как и в случае вещественного аргумента, можем записать систему (218) в виде

$$\frac{dx}{dz} = P(z)\, x. \tag{3.219}$$

Обозначая через *X* интегральную матрицу, т. е. матрицу, столбцами которой являются *n* линейно независимых решений системы (3.218), мы вместо (3.219) можем написать:

$$\frac{dX}{dz} = P(z)\, X. \tag{3.220}$$

Формула Якоби имеет место и при комплексном аргументе z.

$$|X| = c e^{\int_{z_0}^{z} \operatorname{Sp} P\, dz}. \tag{3.221}$$

При этом предполагается, что z_0 и все точки пути, вдоль которого берется $\int_{z_0}^{z}$, являются регулярными точками для однозначной аналитической функции

$$\operatorname{Sp} P(z) = p_{11}(z) + p_{22}(z) + \ldots + p_{nn}(z)$$

(Здесь и в дальнейшем в качестве путей интегрирования берутся кусочно гладкие кривые).

Специфичность рассматриваемого случая комплексного аргумента заключается в том, что при однозначной функции P(z) интегральная матрица X (z) может быть многозначной функцией от z.
В качестве примера рассмотрим систему Коши

$$\frac{dX}{dz} = \frac{U}{z-a} X \quad (U \text{ — постоянная матрица}). \qquad (3.222)$$

Одним из решений этой системы, как и в случае вещественного аргумента, является интегральная матрица

$$X = e^{U \ln (z-a)} = (z-a)^U. \qquad (3.223)$$

В качестве области G возьмем всю z-плоскость за исключением точки z = a. Все точки этой области являются регулярными точками матрицы коэффициентов

$$P(z) = \frac{U}{z-a}.$$

Если $U \neq 0$, то точка $z = a$ является особой точкой (полюсом первого порядка) для матричной функции $P(z) = \frac{U}{z-a}$.

Элемент интегральной матрицы (3.223) при однократном обходе в положительном направлении точки z=a возвращается с новым значением, которое получается из старого умножением справа на постоянную матрицу

$$V = e^{2\pi i U}.$$

Для общей системы (3.220) теми же рассуждениями, что и в случае вещественного аргумента, убеждаемся в том, что два однозначных решения X и $\tilde{X}$ в некоторой части области G всегда связаны формулой

$$X = \tilde{X}C,$$

где C — некоторая постоянная матрица. Эта формула сохранится при любом аналитическом продолжении функций $X(z)$ и $\tilde{X}(z)$ в области G.

Теорема о существовании и единственности (при заданных начальных значениях) решения системы (3.218) может быть доказана аналогично вещественному случаю.

Рассмотрим *односвязную* и притом *звездообразную относительно точки* z_0 область G_1 (область называется *звездообразной относительно точки* z_0, если любой отрезок, соединяющий произвольную точку z области с точкой z_0, целиком лежит в данной области), составляющую часть области G, и пусть матричная функция P(z) регулярна (то есть все элементы $p_{ik}(z)$ $(i, k = 1, 2, \ldots, n)$ матрицы P(z) являются регулярными функциями в области G_1) ряда в области G_1. Составим ряд

$$E+\int_{z_0}^{z} P\,dz+\int_{z_0}^{z} P\,dz\int_{z_0}^{z} P\,dz+\ldots \tag{3.224}$$

Из односвязности области G_1 следует, что каждый встречающийся в ряду (3.224) интеграл не зависит от пути интегрирования и представляет собой регулярную функцию в области G_1. Поскольку область G_1 звездообразна относительно z_0, то при оценке модулей этих интегралов мы можем считать, что все интегралы берутся вдоль прямолинейного отрезка, соединяющего точки z_0 и z.

Абсолютная и равномерная в любой замкнутой части области G_1, содержащей точку z_0, сходимость ряда (3.224) вытекает из сходимости мажорантного ряда

$$1+lM+\frac{n}{2!}l^2M^2+\frac{n^2}{3!}l^3M^3+\ldots$$

Здесь M — верхняя граница для модуля матрицы P (z), а l — верхняя граница расстояний точки z от точки z_0, причем обе границы относятся к рассматриваемой замкнутой части области G_1.

Путем почленного дифференцирования проверяется, что сумма ряда (224) представляет собой решение уравнения (3.220). Это решение нормировано, поскольку оно при $z=z_0$ обращается в единичную матрицу E. Однозначное нормированное решение системы (2.220), как и в вещественном случае, будем называть матрицантом и будем обозначать через $\Omega_{z_0}^{z}(P)$. Таким образом, мы получили представление матриканта в области G_1 в виде ряда

$$\Omega_{z_0}^{z}(P)=E+\int_{z_0}^{z} P\,dz+\int_{z_0}^{z} P\,dz\int_{z_0}^{z} P\,dz+\ldots \tag{3.225}$$

(Приведенное доказательство существования нормированного решения и представления его в области G_1 рядом (3.225) сохраняет свою силу, если вместо звездообразности области сделать более общее допущение: для любой замкнутой части области G_1 существует такое положительное число l, что любую точку z этой замкнутой части можно соединить с z_0 путем, длина которого $\le l$).

Свойства 1° — 4° *матрицанта, установленные* в п. 3.20, *автоматически переносятся и на случай комплексного аргумента.*

Произвольное решение уравнения (3.220), регулярное в области G и обращающееся при $z = z_0$ в матрицу X_0, представится в виде

$$X=\Omega_{z_0}^{z}(P)\cdot C \qquad (C=X_0). \tag{3.226}$$

Формула (3.226) охватывает все однозначные решения, регулярные в окрестности точки z_0 [z_0 — регулярная точка для матрицы коэффициентов $P(z)$]. Эти решения, будучи аналитически продолжены в область G, дадут все решения уравнения (3.220), т. е. уравнение (3.220) не может иметь решений, для которых z_0 была бы особой точкой.

Для аналитического продолжения матрицанта в область G удобно пользоваться мультипликативным интегралом.

3.23. Мультипликативный интеграл в комплексной области

Мултипликативный интеграл вдоль некоторой кривой в комплексной плоскости определяется следующим образом.

Пусть даны некоторый путь L и матричная функция P(z), непрерывная на кривой L. Разобьем путь L на n частей $(z_0, z_1)\ (z_1, z_2)\ \ldots\ (z_{n-1}, z)$; здесь z_0 — начало, $z_n = z$ — конец пути, а $z_1, z_2, \ldots, z_{n-1}$ — промежуточные точки разбиения. На отрезке пути (z_{k-1}, z_k) выберем произвольную точку ξ_k и введем обозначения $\Delta z_k = z_k - z_{k-1};\ k = 1, 2, \ldots, n.$ Тогда по определению

$$\widehat{\int_L}[E + P(z)\,dz] = \lim_{\Delta z_k \to 0}[E + P(\xi_n)\Delta z_n]\ldots[E + P(\xi_1)\Delta z_1].$$

Сопоставляя это определение с приведенным ранее определением, мы видим, что новое определение совпадает с прежним в том частном случае, когда путь L является отрезком вещественной оси. Однако и в общем случае, когда путь L произвольно расположен в комплексной плоскости, новое определение может быть сведено к старому при помощи замены переменной интегрирования.

Если

$$z = z(t)$$

есть параметрическое уравнение пути, причем z(t) — непрерывная функция в интервале $(t_0,\ t)$, имеющая в этом интервале кусочно непрерывную производную $\frac{dz}{dt}$,то, как легко видеть,

$$\widehat{\int_L}[E + P(z)\,dz] = \widehat{\int_{t_0}^{t}}\left\{E + P[z(t)]\frac{dz}{dt}\,dt\right\}.$$

Эта формула показывает, что мультипликативный интеграл вдоль произвольного пути существует, если подинтегральная матрица $P(z)$ непрерывна вдоль этого пути. Даже в случае, когда $P(z)$ — непрерывная функция вдоль L, функция $P[z(t)]\frac{dz}{dt}$ может быть кусочно непрерывной. В этом случае мы можем интервал (t_0, t) разбить на частичные интервалы, в каждом из которых производная $\frac{dz}{dt}$ непрерывна, и под интегралом от t_0 до t понимать сумму интегралов вдоль этих частичных интервалов.

Мультипликативная производная определяется прежней формулой

$$D_z X = \frac{dX}{dz} X^{-1}.$$

При этом предполагается, что $X(z)$ — аналитическая функция.

Все дифференциальные формулы (I — III) предыдущего параграфа переносятся без изменения на случай комплексного аргумента. Что же касается интегральных формул IV — VI, то несколько видоизменяется их внешняя запись:

$$\text{IV}'. \qquad \widehat{\int\limits_{(L'+L'')}} (E + P\,dz) = \widehat{\int\limits_{L''}} (E + P\,dz) \widehat{\int\limits_{L'}} (E + P\,dz).$$

$$\text{V}'. \qquad \widehat{\int\limits_{-L}} (E + P\,dz) = \left[\widehat{\int\limits_{L}} (E + P\,dz)\right]^{-1}.$$

$$\text{VI}'. \qquad \widehat{\int\limits_{L}} (E + CPC^{-1}\,dz) = C \widehat{\int\limits_{L}} (E + P\,dz)\, C^{-1} \qquad \begin{pmatrix} C \text{ — постоянная} \\ \text{матрица} \end{pmatrix}.$$

В формуле IV' мы через $L' + L''$ обозначили составной путь, получающийся при прохождении сначала пути L', а затем пути L''. В формуле V′ — L обозначает путь, отличающийся от пути L только направлением.

Формула VII принимает теперь вид

$$\text{VII}'. \qquad \widehat{\int\limits_{L}} [E + (Q + D_z X)\,dz] = X(z) \widehat{\int\limits_{L}} (E + X^{-1} Q X\,dz)\, X(z_0)^{-1}.$$

Здесь $X(z_0)$ и $X(z)$ в правой части обозначают соответственно значения $X(z)$ в начале и в конце пути L. /

Формула VIII заменится теперь формулой

$$\text{VIII}'. \quad \operatorname{mod}\left[\int\limits_L (E + P\,dz) - \int\limits_L (E + Q\,dz)\right] \leqslant \frac{1}{n} e^{nql}\,(e^{nd\cdot l} - 1)\,I,$$

Где $\operatorname{mod} Q \leqslant qI$, $\operatorname{mod}(P - Q) \leqslant d \cdot I$, $I = \| 1 \|$, а l — длина пути L.

Формула VIII' получается сразу из формулы VIII, если в последней сделать преобразование переменной, выбрав в качестве новой переменной интегрирования длину дуги s вдоль пути L $\left(\text{при этом} \left|\frac{dz}{ds}\right| = 1\right)$.

Как и в случае вещественного аргумента, существует тесная связь мультипликативного интеграла с матрицантом.

Пусть дана однозначная аналитическая матричная функция P(z), регулярная в области G, и пусть G_0 — односвязная область, содержащая точку z_0 и составляющая часть области G. Тогда матрицант

$$\Omega_{z_0}^{z}(P)$$

будет регулярной функцией от z в области G_0.

Соединим точки z_0 и z произвольным путем L, целиком лежащим в G_0, и выберем на L промежуточные точки z_1, z_2, ..., z_{n-1}. Тогда, пользуясь равенством

$$\Omega_{z_0}^{z} = \Omega_{z_{n-1}}^{z} \ldots \Omega_{z_1}^{z_2}\Omega_{z_0}^{z_1},$$

совершенно так же, как в п. 3.21, предельным переходом получим:

$$\Omega_{z_0}^{z}(P) = \widehat{\int\limits_L}(E + P\,dz) = \widehat{\int}_{z_0}^{z}(E + P\,dz). \qquad (3.227)$$

Из этой формулы видно, что мультипликативный интеграл не зависит от формы пути, а зависит только от начала и конца пути, если весь путь интегрирования лежит в односвязной области G_0, в которой подинтегральная функция $P(z)$ регулярна. В частности, для замкнутого контура L, лежащего в односвязной области G_0, имеем:

$$\oint (E + P\,dz) = E. \qquad (3.228)$$

Эта формула представляет собой аналог известной теоремы Коши, согласно которой обычный (немультипликативный) интеграл вдоль замкнутого контура равен нулю, если этот контур лежит в односвязной области, в которой подинтегральная функция регулярна.

Представление матрицанта в виде мультипликативного интеграла (3.227) может быть использовано для аналитического продолжения матрицанта вдоль произвольного пути L в области G. В этом случае формула

$$X = \widehat{\int_{z_0}^{z}} (E + P\,dz)\,X_0 \tag{3.229}$$

задает все ветви многозначной интегральной матрицы X дифференциального уравнение $\frac{dX}{dz} = PX$, обращающейся в X_0 на одной из ветвей при z = z_0. Различные ветви получаются за счет различных путей, соединяющих точки z_0 и z.

Согласно формуле Якоби (3.221)

$$|X| = |X_0|\, e^{\int_{z_0}^{z} \mathrm{Sp}\, P\,dz}$$

и, в частности, при $X_0 = E$

$$\left|\widehat{\int_{z_0}^{z}} (E + P\,dz)\right| = e^{\int_{z_0}^{z} \mathrm{Sp}\, P\,dz}. \tag{3.230}$$

Из этой формулы следует, что мультипликативный интеграл всегда представляет собой неособенную матрицу, если только путь интегрирования целиком лежит в области, в которой функция P (z) регулярна.

Если L — произвольный замкнутый путь в G и G — неодносвязная область, то равенство (3.228) может и не иметь места. Более того, в этом случае значение интеграла

$$\widehat{\oint} (E + P\,dz)$$

не определяется заданием подинтегральной функции и замкнутого пути интегрирования L, а зависит еще и от выбора начальной точки интегрирования z_0 на кривой L. Действительно, выберем на замкнутой кривой L две точки z_0 и z_1 и обозначим участки пути от z_0 до z_1 и от z_1 до z_0 (в направлении интегрирования) соответственно через L_1 и L_2. Тогда согласно формуле IV (здесь мы для сокращения обозначений опускаем подинтегральное выражение $E + Pdz$, одно и то же во всех интегралах).

$$\widehat{\oint_{z_0}} = \widehat{\int_{L_2}} \cdot \widehat{\int_{L_1}}, \quad \widehat{\oint_{z_1}} = \widehat{\int_{L_1}} \cdot \widehat{\int_{L_2}}$$

и, следовательно,

$$\widehat{\oint\limits_{z_1}} = \widehat{\int\limits_{L_1}} \cdot \widehat{\oint\limits_{z_0}} \cdot \widehat{\int\limits_{L_1}}^{-1}. \qquad (3.231)$$

Формула (3.231) показывает, что символ $\widehat{\oint}(E + P\,dz)$ определяет некоторую матрицу с точностью до преобразования подобия, т. е. определяет только элементарные делители некоторой матрицы.

Рассмотрим элемент X (z) решения (3.229) в окрестности точки z_0. Пусть L — произвольный замкнутый путь в G, начинающийся и кончающийся в точке z_n. После аналитического продолжения вдоль L элемент X (z) перейдет в некоторый элемент $\tilde{X}(z)$. При этом новый элемент $\tilde{X}(z)$ будет удовлетворять тому же дифференциальному уравнению (3.220), поскольку P (z) — однозначная функция в G.
Поэтому $\tilde{X} = XV$, где V — некоторая неособенная постоянная матрица. Из формулы (3.229) следует, что

$$\tilde{X}(z_0) = \widehat{\oint\limits_{z_0}}(E + P\,dz)\,X_0.$$

Сопоставляя это равенство с предыдущим, найдем:

$$V = X_0^{-1}\widehat{\oint\limits_{z_0}}(E + P\,dz)\,X_0. \qquad (3.232)$$

В частности, для матрицанта $X = \Omega_{z_0}^{z}$ имеем $X_0 = E$, и тогда

$$V = \widehat{\oint\limits_{z_0}}(E + P\,dz). \qquad (3.233)$$

3.24. Изолированная особая точка

Исследуем поведение решения (интегральной матрицы) в окрестности изолированной особой точки a.

Пусть матричная функция P(z) регулярна для значений z, удовлетворяющих неравенствам

$$0 < |z - a| < R.$$

Совокупность этих значений образует двусвязную область G. Матричная функция P(z) в области G разлагается в ряд Лорана

$$P(z) = \sum_{n=-\infty}^{+\infty} P_n (z-a)^n. \qquad (3.234)$$

Элемент $X(z)$ интегральной матрицы после однократного обхода в положительном направлении вокруг a вдоль пути L перейдет в элемент

$$X^+(z) = X(z)\,V,$$

где V — некоторая постоянная неособенная матрица.

Пусть U — постоянная матрица, связанная с матрицей V равенством

$$V = e^{2\pi i U}. \qquad (3.235)$$

Тогда матричная функция $(z-a)^U$ после обхода вдоль L также переходит в $(z-a)^U V$. Поэтому аналитическая в области G матричная функция

$$F(z) = X(z)\,(z-a)^{-U} \qquad (3.236)$$

при аналитическом продолжении вдоль L переходит сама в себя (остается неизменной) (отсюда уже следует, что функция $F(z)$ при обходе вдоль любого другого замкнутого пути в G возвращается к исходному значению.) Поэтому матричная функция $F(z)$ регулярна в G и разлагается в G в ряд Лорана

$$F(z) = \sum_{n=-\infty}^{+\infty} F_n (z-a)^n. \qquad (3.237)$$

Из (3.236) следует:

$$X(z) = F(z)\,(z-a)^U. \qquad (3.238)$$

Таким образом, каждая интегральная матрица $X(z)$ может быть представлена в виде (3.238), где однозначная функция $F(z)$ и постоянная матрица U зависят от матрицы коэффициентов $P(z)$. Однако алгоритмическое определение матрицы U и коэффициентов F_n ряда (3.237) по коэффициентам P_n ряда (3.234) в общем случае представляет собой сложную задачу.

Частный случай этой задачи, когда

$$P(z) = \sum_{n=-1}^{\infty} P_n (z-a)^n,$$

будет разобран в п. 3.25. В этом случае точка a называется *регулярной особой точкой* системы (3.220). Если разложение (3.234) имеет вид

$$P(z) = \sum_{n=-q}^{\infty} P_n (z-a)^n \qquad (q > 1;\ P_{-q} \neq 0),$$

то точка a называется *иррегулярной особой точкой типа полюса.* Наконец, если в ряду (3.234) имеется бесчисленное множество отличных от нуля матричных коэффициентов P_n при отрицательных степенях $z — a$, то точка a называется *существенной особой точкой* данной дифференциальной системы.

Из формулы (3.238) следует, что интегральная матрица $X(z)$ при любом однократном обходе в положительном направлении (вдоль некоторого замкнутого пути L) умножается справа на одну и ту же матрицу

$$V = e^{2\pi i U}.$$

Если этот обход начинается (и кончается) в точке z_0, то согласно (3.232)

$$V = X(z_0)^{-1} \overset{\frown}{\oint_{z_0}} (E + P\,dz)\, X(z_0). \tag{3.239}$$

Если вместо интегральной матрицы $X(z)$ мы рассмотрим любую другую интегральную матрицу $\widehat{X}(z) = X(z)\,C$ (C — постоянная матрица, $|C| \neq 0$), то, как видно из (3.239), матрица V заменится подобной матрицей

$$\widehat{V} = C^{-1} V C.$$

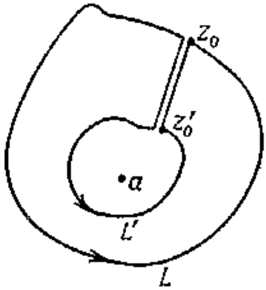
Таким образом, «интегральные подстановки» V данной системы образуют класс подобных между собой матриц.

Из формулы (239) также следует, что интеграл

$$\overset{\frown}{\oint_{z_0}} (E + P\,dz) \tag{3.240}$$

определяется начальной точкой обхода z_0 и не зависит от формы кривой обхода (конечно, при условии, что путь интегрирования обходит точку a однократно в положительном направлении.) Если же мы меняем точку z_0, то получающиеся при этом различные значения интеграла (3.240) подобны между собой (это вытекает из формулы (3.239), а также из формулы (3.231).

В этих свойствах интеграла (3.240) можно убедиться и непосредственно. Действительно, пусть L и L' — два замкнутых пути в G вокруг точки $z=a$ с начальными точками обхода z_0 и z'_0 (см. приведенный ниже рисунок).



Двусвязная область, заключенная между L и L', может быть сделана односвязной, если провести разрез от z_0 до z'_0. Интеграл вдоль разреза мы обозначим через

$$T = \widehat{\int_{z_0}^{z'_0}} (E + P\,dz).$$

Поскольку мультипликативный интеграл вдоль замкнутого контура односвязной области равен E, то

$$\widehat{\int_{L'}} T \widehat{\int_{L}}^{-1} T^{-1} = E,$$

откуда

$$\widehat{\int_{L'}} = T \widehat{\int_{L}} T^{-1}.$$

Таким образом, как и V, интеграл

$$\oint (E + P\,dz)$$

определен с точностью до подобия, и равенство (3.239) иногда будем записывать так:

$$V = \widehat{\oint} (E + P\,dz),$$

понимая под этим совпадение элементарных делителей у матриц, стоящих в левой и правой частях равенства.

Рассмотрим для примера систему с регулярной особой точкой

$$\frac{dX}{dz} = P(z)\,X,$$

где

$$P(z) = \frac{P_{-1}}{z-a} + \sum_{n=0}^{+\infty} P_n (z-a)^n.$$

Пусть

$$Q(z) = \frac{P_{-1}}{z-a}.$$

Пользуясь формулой VIII' предыдущего параграфа, дадим оценку модуля разности

$$D = \widehat{\oint}(E + P\,dz) - \widehat{\oint}(E + Q\,dz), \qquad (3.241)$$

выбрав в качестве пути интегрирования окружность радиуса r $(r < R)$ с положительным направлением обхода. Тогда при

$$\operatorname{mod} P_{-1} \leqslant p_{-1} I, \quad \operatorname*{mod}_{|z-a|=r} \sum_{n=0}^{\infty} P_n (z-a)^n \leqslant d(r) I, \quad I = \| 1 \|$$

мы можем в формуле VIII' положить:

$$q = \frac{p_{-1}}{r}, \quad d = d(r), \quad l = 2\pi r,$$

после чего получим:

$$\operatorname{mod} D \leqslant \frac{1}{n} e^{2\pi p_{-1}} \left(e^{2\pi n r d(r)} - 1\right) I.$$

Отсюда видно, что

$$\lim_{r \to 0} D = 0. \qquad (3.242)$$

При этом мы используем то, что при надлежащем выборе $d(r)$

$$\lim_{r \to 0} d(r) = d_0,$$

где d_0 — наибольший из модулей элементов матрицы P_0.

С другой стороны, система

$$\frac{dY}{dz} = QY$$

является системой Коши, и в этом случае при любом выборе начальной точки обхода z_0 и при любом $r < R$

$$\widehat{\oint_{z_0}}(E + Q\,dz) = e^{2\pi i P_{-1}}.$$

Поэтому из (3.241) и (3.242) следует:

$$\lim_{r \to 0} \oint_{z_0} (E + P\,dz) = e^{2\pi i P_{-1}}. \qquad (3.243)$$

Но элементарные делители интеграла

$$\overset{\frown}{\oint_{z_0}} (E + P\,dz)$$

не зависят от z_0 и r и совпадают с элементарными делителями интегральной подстановки V.

Отсюда Вольтерра делает вывод, что матрицы V и $e^{2\pi i P_{-1}}$ подобны, и потому интегральная подстановка V с точностью до подобия определяется матрицей «вычетов» P_{-1}.

Это утверждение Вольтерра ошибочно.

Из (3.239) и (3.243) можно лишь сделать вывод, что *характеристические числа интегральной подстановки V совпадают с характеристическими числами матрицы* $e^{2\pi i P_{-1}}$. Однако элементарные делители у этих матриц могут быть различными. Так, например, матрица

$$\begin{Vmatrix} \alpha & r \\ 0 & \alpha \end{Vmatrix}$$

при любом $r \neq 0$ имеет один элементарный делитель $(\lambda - \alpha)^2$, а предел этой матрицы при $r \to 0$, т. е. матрица $\begin{Vmatrix} \alpha & 0 \\ 0 & \alpha \end{Vmatrix}$, имеет два элементарных делителя

$$\lambda - \alpha,\ \lambda - \alpha.$$

Таким образом, утверждение Вольтерра не вытекает из формул (3.239) и (3.243). Но оно и вообще неверно, как показывает следующий пример.

Пусть

$$P(z) = \begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix} \frac{1}{z} + \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}.$$

Соответствующая система дифференциальных уравнений имеет вид:

$$\frac{dx_1}{dz} = x_2, \quad \frac{dx_2}{dz} = -\frac{x_2}{z}.$$

Интегрируя эту систему, находим:

$$x_1 = c \ln z + d, \quad x_2 = \frac{c}{z}.$$

Интегральная матрица

$$X(z) = \begin{Vmatrix} \ln z & 1 \\ z^{-1} & 0 \end{Vmatrix}$$

при однократном положительном обходе вокруг особой точки $z = 0$ умножается справа на матрицу

$$V = \begin{Vmatrix} 1 & 0 \\ 2\pi i & 1 \end{Vmatrix}.$$

Эта матрица имеет один элементарный делитель $(\lambda - 1)^2$. В то же время матрица

$$e^{2\pi i P_{-1}} = e^{2\pi i \begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix}} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = E$$

имеет два элементарных делителя $\lambda - 1, \lambda - 1$.

Рассмотрим теперь случай, когда матрица *P(z)* имеет конечное число отрицательных степеней *r — а (а* — регулярная или иррегулярная особая точка типа полюса):

$$P(z) = \frac{P_{-q}}{(z-a)^q} + \dots + \frac{P_{-1}}{z-a} + \sum_{n=0}^{\infty} P_n (z-a)^n \quad (q \geq 1; \ \ P_{-q} \neq 0).$$

Преобразуем данную систему

$$\frac{dX}{dz} = PX, \tag{3.244}$$

положив

$$X = A(z)\,Y, \tag{3.245}$$

где *А (z)* — матричная функция, регулярная в точке *z*=0 и принимающая в этой точке значение *Е:*

$$A(z) = E + A_1(z - a) + A_2(z - a)^2 + \dots;$$

степенной ряд в правой части сходится при $|z - a| < r_1$.

3.25. Регулярная особая точка

Исследуя поведение решения в окрестности особой точки, мы без нарушения общности рассуждения можем принять, что особой точкой является точка z= 0. Преобразованием $z' = z - a$ или $z' = \dfrac{1}{z}$ можно соответственно любую конечную точку $z = a$ или $z = \infty$ перевести в точку $z' = 0$.

1. Пусть дана система

$$\frac{dX}{dz} = P(z)\,X, \tag{3.246}$$

где

$$P(z) = \frac{P_{-1}}{z} + \sum_{m=0}^{\infty} P_m z^m \tag{3.247}$$

и ряд $\sum_{m=0}^{\infty} P_m z^m$ сходится внутри круга | z | < r.

Положим

$$X = A(z)\,Y, \tag{3.248}$$

где

$$A(z) = E + A_1 z + A_2 z^2 + \ldots \tag{3.249}$$

Оставляя пока в стороне вопрос о сходимости ряда (3.249), постараемся так определить матричные коэффициенты A_m этого ряда, чтобы преобразованная система

$$\frac{dY}{dz} = P^*(z)\,Y, \tag{3.250}$$

где

$$P^*(z) = \frac{P^*_{-1}}{z} + \sum_{m=0}^{\infty} P^*_m z^m, \tag{3.251}$$

имела возможно более простой («канонический») вид. Мы будем стремиться к тому, чтобы в ряду (3.254) было конечное (и притом возможно меньшее) число коэффициентов P_m, отличных от нуля.

Подставляя в (3.246) вместо X произведение AY и используя (3.250), мы получим:

$$A(z)\,P^*(z)\,Y + \frac{dA}{dz}\,Y = P(z)\,A(z)\,Y.$$

Умножая обе части этого равенства справа на Y^{-1}, найдем:

$$P(z)\,A(z) - A(z)\,P^*(z) = \frac{dA}{dz}.$$

Заменяя здесь $P(z)$, $A(z)$, $P^*(z)$ рядами (3.247), (3.249), (3.251) и приравнивая в левой и правой частях равенства коэффициенты при одинаковых степенях z, получим бесконечную систему матричных уравнений для искомых коэффициентов A_1, A_2, ... (во всех уравнениях, начиная со второго, мы заменим в силу первого уравнения матрицу P_{-1}^* на P_{-1}):

$$\left.\begin{aligned}
&1)\ P_{-1} = P^*_{-1},\\
&2)\ P_{-1}A_1 - A_1(P_{-1} + E) + P_0 = P^*_0,\\
&3)\ P_{-1}A_2 - A_2(P_{-1} + 2E) + P_0A_1 - A_1P^*_0 + P_1 = P^*_1,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&(m+2)\ P_{-1}A_{m+1} - A_{m+1}[P_{-1} + (m+1)E] +\\
&\quad + P_0A_m - A_mP^*_0 + P_1A_{m-1} - A_{m-1}P^*_1 + \dots + P_m = P^*_m.
\end{aligned}\right\} \quad (3.252)$$

2. Рассмотрим отдельно несколько случаев:

1° Матрица P_{-1} не имеет различных характеристических чисел, отличающихся друг от друга на целое число.

В этом случае при любом k = 1, 2, 3, ... матрицы P_{-1} и $P_{-1}+kE$ не имеют общих характеристических чисел, и потому матричное уравнение

$$P_{-1}U - U(P_{-1} + kE) = T$$

при любой правой части T имеет одно и только одно решение.

Это решение будем обозначать через

$$\Phi_k(P_{-1}, T).$$

Поэтому в уравнениях (3.252) можно положить все матрицы P^*_m $(m = 0, 1, 2, \dots)$ равными нулю и последовательно определить A_1, A_2, ... при помощи равенств

$$A_1 = \Phi_1(P_{-1}, -P_0), \quad A_2 = \Phi_2(P_{-1}, -P_1 - P_0A_1), \dots$$

Тогда преобразованная система является системой Коши

$$\frac{dY}{dz} = \frac{P_{-1}}{z}Y,$$

и потому решение X исходной системы (246) имеет вид

$$X = A(z)\,z^{P_{-1}}. \quad (3.253)$$

Формула (3.253) определяет одну интегральную матрицу системы (3.246). Произвольная интегральная матрица получается из (3.253) умножением справа на произвольную постоянную неособенную матрицу C.

2° Среди различных характеристических чисел матрицы P_{-1} имеются числа, разность между которыми является целым числом; при этом матрица P_{-1} имеет простую структуру.

Обозначим через $\lambda_1, \lambda_2, \dots, \lambda_n$ характеристические числа матрицы P_{-1}, расположенные так, чтобы имели место неравенства

$$\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \dots \geqslant \operatorname{Re}\lambda_n. \quad (3.254)$$

Не нарушая общности рассуждений, мы можем матрицу P_{-1} заменить любой, ей подобной. Это следует из того, что, умножая обе части уравнения (246) слева на неособенную матрицу T, а справа — на T^{-1}, мы фактически заменяем все P_m на TP_mT^{-1} $(m = -1, 0, 1, 2, \ldots)$ (при этом и X заменяется на TXT^{-1}). Поэтому мы будем считать, что в рассматриваемом случае P_{-1} — диагональная матрица:

$$P_{-1} = \| \lambda_i \delta_{ik} \|_1^n. \qquad (3.255)$$

Введем обозначения для элементов матриц P_m, P_m^* и A_m:

$$P_m = \| p_{ik}^{(m)} \|_1^n, \quad P_m^* = \| p_{ik}^{(m*)} \|_1^n, \quad A_m = \| x_{ik}^{(m)} \|_1^n. \qquad (3.256)$$

Для определения A_1 воспользуемся вторым из уравнений (3.252). Это матричное уравнение можно заменить скалярными уравнениями

$$(\lambda_i - \lambda_k - 1)\, x_{ik}^{(1)} + p_{ik}^{(0)} = p_{ik}^{(0*)} \quad (i, k = 1, 2, \ldots, n). \qquad (3.257)$$

Если ни одна из разностей $\lambda_i - \lambda_k$ не равна единице, то мы можем положить $P_0^* = 0$. Тогда из (3.252) $A_1 = \Phi_1(P_{-1}, -P_0)$. (Мы пользуемся обозначениями, введенными при разборе случая 1°).

В этом случае элементы матрицы A_1 однозначно определятся из уравнений (3.257):

$$x_{ik}^{(1)} = -\frac{p_{ik}^{(0)}}{\lambda_i - \lambda_k - 1} \quad (i, k = 1, 2, \ldots, n). \qquad (3.258)$$

Если же при некоторых i, k (в силу (89) это возможно лишь при $i < k$)

$$\lambda_i - \lambda_k = 1,$$

то соответствующее $p_{ik}^{(0*)}$ определяется из:

$$p_{ik}^{(0*)} = p_{ik}^{(0)},$$

а соответствующее $x_{ik}^{(1)}$ выбирается совершенно произвольно.

При тех же i и k, при которых $\lambda_i - \lambda_k \neq 1$, мы полагаем:

$$p_{ik}^{(0*)} = 0,$$

а соответствующее $x_{ik}^{(1)}$ находим по формуле (258).

Определив A_1 мы переходим к определению A_2 из третьего уравнения (252). Заменим это матричное уравнение системой n^2 скалярных уравнений:

$$(\lambda_i - \lambda_k - 2)\, x_{ik}^{(2)} = p_{ik}^{(1*)} - p_{ik}^{(1)} - (P_0 A_1 - A_1 P_0^*)_{ik} \qquad (3.259)$$

$$(i, k = 1, 2, \ldots, n).$$

Здесь мы поступаем так же, как и при определении A_1,

Если $\lambda_i - \lambda_k \neq 2$, то мы полагаем:

$$p_{ik}^{(1*)} = 0,$$

и тогда из (3.259) находим:

$$x_{ik}^{(2)} = -\frac{1}{\lambda_i - \lambda_k - 2}[p_{ik}^{(1)} - (P_0A_1 - A_1P_0^*)_{ik}].$$

Если же $\lambda_i - \lambda_k \neq 2$, то при этих i и k из (3.259) следует, что

$$p_{ik}^{(1*)} = p_{ik}^{(1)} + (P_0A_1 - A_1P_0^*)_{ik}.$$

В этом случае $x^{(2)}_{ik}$ выбирается произвольно.

Продолжая этот процесс далее, мы последовательно определим все матрицы $P_{-1}^*, P^*, P_1^*, \ldots$ и $A_1, A_2, \ldots$

При этом только конечное число из матриц P^*_m будет отлично от нуля, и, как нетрудно видеть, матрица $P^*(z)$ будет иметь вид

$$P^*(z) = \begin{Vmatrix} \frac{\lambda_1}{z} & a_{12}z^{\lambda_1-\lambda_2-1} & \ldots & a_{1n}z^{\lambda_1-\lambda_n-1} \\ 0 & \frac{\lambda_2}{z} & \ldots & a_{2n}z^{\lambda_2-\lambda_n-1} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & \frac{\lambda_n}{z} \end{Vmatrix}, \quad (3.260)$$

где $a_{ik} = 0$, если $\lambda_i - \lambda_k$ не есть целое положительное число, и

$$a_{ik} = p_{ik}^{(\lambda_i-\lambda_k-1*)},$$

если $\lambda_i - \lambda_k$ является целым положительным числом.

Следует отметить, что P_m $(m \geq 0)$ может быть отличным от нуля лишь тогда, когда существуют характеристические числа λ_i и λ_k матрицы P_{-1} такие, что $\lambda_i - \lambda_k - 1 = m$ (при этом в силу (3.254) $i < k$). При данном m каждому такому равенству соответствует элемент $p_{ik}^{(m*)} = a_{ik}$ матрицы P^*_m; этот элемент может быть отличен от нуля. Все остальные элементы матрицы P_m равны нулю.

Обозначим через m_i наибольшую целую часть числа $\operatorname{Re}\lambda_i$ (то есть m_i — наибольшее целое число, не превосходящее $\operatorname{Re}\lambda_i$ $(i = 1, 2, \ldots, n)$:

$$m_i = [\operatorname{Re}\lambda_i] \quad (i = 1, 2, \ldots, n). \quad (3.261)$$

Тогда в силу (3.254)

$$m_1 \geqslant m_2 \geqslant \ldots \geqslant m_n.$$

При этом если $\lambda_i - \lambda_k$ есть целое число, то

$$\lambda_i - \lambda_k = m_i - m_k.$$

Поэтому в выражении (3.260) канонической матрицы $P^*(z)$ мы можем все разности $\lambda_i - \lambda_k$ заменить на $m_i - m_k$. Кроме того, мы положим:

$$\tilde{\lambda}_i = \lambda_i - m_i \quad (i = 1, 2, \ldots, m), \quad (3.262)$$

$$M = \| m_i \delta_{ik} \|_1^n, \quad U = \begin{Vmatrix} \tilde{\lambda}_1 & a_{12} & \dots & a_{1n} \\ 0 & \tilde{\lambda}_2 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \tilde{\lambda}_n \end{Vmatrix}. \qquad (3.263)$$

Тогда из (3.260) следует:

$$P^*(z) = z^M \frac{U}{z} z^{-M} + \frac{M}{z} = D_z(z^M z^U).$$

Отсюда вытекает, что $Y = z^M z^U$ представляет собой решение уравнения (3.250), а

$$X = A(z) z^M z^U \qquad (3.264)$$

является решением уравнения (3.246) (специальный вид матриц (3.263) соответствует каноническому виду матрицы P_{-1}. Если матрица P_{-1}не имеет каноническую форму, то матрицы M и U в (3.264) подобны матрицам (3.263).

3° Переходим к общему случаю. Как было выяснено выше, мы можем, не нарушая общности, матрицу P_{-1} заменить любой матрицей, ей подобной. Мы примем, что матрица P_{-1} имеет нормальную жорданову форму

$$P_{-1} = \{\lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \dots, \lambda_u E_u + H_u\}, \qquad (3.265)$$

причем

$$\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \dots \geqslant \operatorname{Re}\lambda_u. \qquad (3.266)$$

Здесь E обозначает, единичную матрицу, а H — матрицу, у которой элементы первой «наддиагонали» равны единице, а остальные элементы равны нулю. Порядки матриц E_i и H_i в различных диагональных клетках будут, вообще говоря, различными; эти порядки совпадают со степенями соответствующих элементарных делителей матрицы P_{-1} (для сокращения обозначений мы не пишем при E_i и H_i индекса, указывающего порядок этих матриц.)

В соответствии с представлением (3.265) матрицы P_{-1} разобьем все матрицы P_m, P_m^*, A_m на блоки:

$$P_m = (P_{ik}^{(m)})_1^u, \quad P_m^* = (P_{ik}^{(m*)})_1^u, \quad A_m = (X_{ik}^{(m)})_1^u.$$

Тогда второе из уравнений (3.252) может быть заменено системой уравнений

$$(\lambda_i E_i + H_i) X_{ik}^{(1)} - X_{ik}^{(1)} [(\lambda_k + 1) E_k + H_k] + P_{ik}^{(0)} = P_{ik}^{(0*)} \qquad (3.267)$$

$$(i, k = 1, 2, \dots, u),$$

которые могут быть еще переписаны так:

$$(\lambda_i - \lambda_k - 1)\, X_{ik}^{(1)} + H_i X_{ik}^{(1)} - X_{ik}^{(1)} H_k + P_{ik}^{(0)} = P_{ik}^{(0*)} \quad (i, k = 1, 2, \ldots, u). \qquad (3.268)$$

Пусть (для сокращения обозначений мы опускаем индексы *i, k* при обозначении элементов матриц X_{ik}, $P_{ik}^{(0)}$, $P_{ik}^{(0*)}$.)

$$X_{ik}^{(1)} = \left\| \begin{matrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{matrix} \right\| = \| x_{st} \|, \qquad P_{ik}^{(0)} = \| p_{st}^{(0)} \|, \qquad P_{ik}^{(0*)} = \| p_{st}^{(0*)} \|.$$

Тогда матричное уравнение (3.268) (при фиксированных *i* и *k)* можно заменить системой скалярных уравнений вида

$$(\lambda_i - \lambda_k - 1)\, x_{st} + x_{s+1,\, t} - x_{s,\, t-1} + p_{st}^{(0)} = p_{st}^{(0*)} \qquad (3.269)$$

$$(x_{v+1,\, t} = x_{s,\, 0} = 0; \quad s = 1, 2, \ldots, v; \quad t = 1, 2, \ldots, w),$$

где *v* и *w* — порядки матриц $\lambda_i E_i + H_i$ и $\lambda_k E_k + H_k$ в (265).

Если $\lambda_i - \lambda_k \neq 1$, то в системе (3.269) можно положить все $p_{st}^{(0*)} = 0$ и однозначно определить все x_{st} из рекуррентных соотношений (3.269). Это означает, что в матричном уравнении (3.268) мы полагаем

$$P_{ik}^{(0*)} = 0$$

и однозначно определяем $X_{ik}^{(1)}$.

Если $\lambda_i - \lambda_k = 1$, то соотношения (269) принимают вид

$$x_{s+1,\, t} - x_{s,\, t-1} + p_{st}^{(0)} = p_{st}^{(0*)} \qquad (3.270)$$

$$(s = 1, 2, \ldots, v; \quad t = 1, 2, \ldots, w; \quad x_{v+1,\, t} = x_{s,\, 0} = 0).$$

Нетрудно показать, что из уравнений (3.270) можно так определить элементы x_{st} матрицы $X_{ik}^{(1)}$, чтобы матрица $P_{ik}^{(0*)}$ имела в соответствии со своими размерами *(v × w)* вид

$$\left\| \begin{matrix} a_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{v-1} & a_{v-2} & \cdots\; a_1 & a_0 \end{matrix} \right\|, \qquad \left\| \begin{matrix} a_0 & 0 & \cdots & 0 & 0 \ldots 0 \\ a_1 & a_0 & \cdots & 0 & 0 \ldots 0 \\ \vdots & \ddots & \ddots & & \vdots \\ a_{v-1} & & a_1 & a_0 & 0 \ldots 0 \end{matrix} \right\|.$$

$$(v = w) \qquad\qquad\qquad (v < w)$$

$$\begin{Vmatrix} 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot\cdot\cdot\cdot\cdot & \cdot \\ 0 & 0 & \dots & 0 \\ a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \cdot & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{w-1} & \dots & a_1 & a_0 \end{Vmatrix} . \qquad (3.271)$$
$$(v > w)$$

Про матрицы (3.271) будем говорить, что они имеют *правильную нижнюю треугольную форму*.

Аналогично определяются верхние правильные треугольные матрицы.

Из уравнений (2.270) все элементы матрицы $X^{(1)}_{ik}$ не определяются однозначно; имеется некоторый произвол в выборе элементов x_{st}. Это видно и непосредственно из уравнения (3.268): при $\lambda_i - \lambda_k = 1$ к матрице $X^{(1)}_{ik}$ можно прибавить произвольную матрицу, перестановочную с H, т. е. произвольную правильную верхнюю треугольную матрицу.

Из третьего уравнения (252) мы определяем матрицу A_2. Это уравнение можно заменить системой уравнений

$$(\lambda_i - \lambda_k - 2) X^{(2)}_{ik} + H_i X^{(2)}_{ik} - X^{(2)}_{ik} H_k + \{P_0 A_1 - A_1 P_0\}_{ik} + P^{(1)}_{ik} = P^{(1*)}_{ik} \quad (3.272)$$
$$(i, k = 1, 2, \dots, u).$$

Аналогично тому, как это было при определении A_1, если $\lambda_i - \lambda_k \neq 2$, то из соответствующего уравнения (3.272) матрица $X^{(2)}_{ik}$ определяется однозначно при $P^{(1*)}_{ik} = 0$. Если же $\lambda_i - \lambda_k = 2$, то можно так определить матрицу $X^{(3)}_{ik}$ чтобы матрица $P^{(1*)}_{ik}$ имела правильную нижнюю треугольную форму.

Продолжая этот процесс, мы определим последовательно все матричные коэффициенты $A_1, A_2, \dots$ и $P^*_{-1}, P^*_0, P^*_1, \dots$ При этом только конечное число из коэффициентов P^*_m будет отлично от нуля, и матрица P^* (z) будет иметь следующий блочный вид (размеры квадратных матриц E_i, H_i и прямоугольных матриц B_{ik} определяются размерами диагональных клеток в жордановой матрице P_{-1}, т. е. степенями элементарных делителей матрицы P_{-1}):

$$P^*(z)=\begin{pmatrix} \frac{\lambda_1 E_1+H_1}{z} & B_{12}z^{\lambda_1-\lambda_2-1} & \dots & B_{1u}z^{\lambda_1-\lambda_u-1} \\ 0 & \frac{\lambda_2 E_2+H_2}{z} & \dots & B_{2u}z^{\lambda_2-\lambda_u-1} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{\lambda_u E_u+H_u}{z} \end{pmatrix}, \qquad (3.273)$$

где

$$B_{ik}=\begin{cases} 0, & \text{если } \lambda_i-\lambda_k \text{ не есть целое положительное число;} \\ P_{ik}^{(\lambda_i-\lambda_k-1^*)}, & \text{если } \lambda_i-\lambda_k = \text{целому положительному числу} \\ & (i, k=1, 2, \dots, u). \end{cases}$$

Все матрицы $B_{ik}\ (i, k=1, 2, \dots, u;\ i<k)$ имеют правильную нижнюю треугольную форму.

Как и в предыдущем случае, обозначим через m_i целую часть $\operatorname{Re}\lambda_i$

$$m_i=[\operatorname{Re}\lambda_i] \quad (i=1, 2, \dots, u) \qquad (3.274)$$

и положим

$$\lambda_i=m_i+\tilde{\lambda}_i \quad (i=1, 2, \dots, u). \qquad (3.275)$$

Тогда снова в выражении (3.273) для P^*(z) мы можем всюду разность $\lambda_i-\lambda_k$ заменить разностью m_i — m_k. Вводя диагональную матрицу с целыми элементами M и верхнюю треугольную матрицу U при помощи равенств (здесь разбиение на блоки соответствует разбиению матриц P_{-1} и P^* (z)).

$$M=(m_i E_i \delta_{ik})_1^u, \quad U=\begin{pmatrix} \tilde{\lambda}_1 E_1+H_1 & B_{12} & \dots & B_{1u} \\ 0 & \tilde{\lambda}_2 E_2+H_2 & \dots & B_{2u} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \tilde{\lambda}_u E_u+H_u \end{pmatrix}, \qquad (3.276)$$

исходя из (3.273), легко получим следующее представление для матрицы P^*(z):

$$P^*(z)=z^M\frac{U}{z}\cdot z^{-M}+\frac{M}{z}=D_z\left(z^M z^U\right).$$

Отсюда следует, что решение уравнения (3.250) может быть задано в виде

$$Y=z^M z^U,$$

а решение уравнения (3.246) может быть представлено так:

$$X = A(z)\, z^{M} z^{U}, \tag{3.277}$$

Здесь A (z) — матричный ряд (3.249), M — диагональная матрица с постоянными целыми элементами, U — постоянная треугольная матрица. Матрицы M и U определяются равенствами (3.274), (3.275) и (3.276).

3. Переходим к доказательству сходимости ряда

$$A(z) = E + A_1 z + A_2 z^2 + \ldots$$

Воспользуемся леммой, которая представляет и самостоятельный интерес.

Лемма. *Если ряд*

$$x = a_0 + a_1 z + a_2 z^2 + \ldots \tag{3.278}$$

формально удовлетворяет системе (здесь $x = (x_1, x_2, ..., x_n)$ — столбец из неизвестных функций; a_0, a_1, a_2, ... — постоянные столбцы; P (z) — квадратная матрица коэффициентов.)

$$\frac{dx}{dz} = P(z)\, x, \tag{3.279}$$

для которой z = 0 *является регулярной особой точкой, то ряд* (3.278) *сходится в любой окрестности точки* z = 0, *в которой сходится разложение в ряд* (3.247) *для матрицы коэффициентов* P (z).

Доказательство. Пусть

$$P(z) = \frac{P_{-1}}{z} + \sum_{q=0}^{\infty} P_q z^q,$$

где ряд

$$\sum_{m=0}^{\infty} P_m z^m$$

сходится при $|z| < r$. Тогда существуют такие положительные постоянные p_{-1} и p, что

$$\operatorname{mod} P_{-1} \leq p_{-1} I,\ \operatorname{mod} P_m \leq \frac{p}{r^m} I,\ I = \|1\| \quad (m = 0, 1, 2, \ldots). \tag{3.280}$$

Подставляя в (3.279) вместо x ряд (3.278) и приравнивая между собой коэффициенты при одинаковых степенях в обеих частях равенства (3.279), получим бесконечную систему векторных (столбцевых) равенств

$$\left.\begin{aligned} P_{-1}a_0 &= 0, \\ (E - P_{-1})\,a_1 &= P_0a_0, \\ (2E - P_{-1})\,a_2 &= P_0a_1 + P_1a_0, \\ &\dots\dots\dots\dots\dots\dots \\ (mE - P_{-1})\,a_m &= P_0a_{m-1} + P_1a_{m-2} + \dots + P_{m-1}a_0 \\ &\dots\dots\dots\dots\dots\dots \end{aligned}\right\} \quad (3.281)$$

Нам достаточно доказать, что какой-либо остаток ряда (3.278)

$$x^{(k)} = a_kz^k + a_{k+1}z^{k+1} + \dots \qquad (3.282)$$

сходится в окрестности точки z = 0. Число k подчиним неравенству .

$$k > np_{-1}.$$

Тогда число k будет превосходить модули всех характеристических чисел матрицы P_{-1} и потому при $m \geq k$ будем иметь $|\, mE - P_{-1}\,| \neq 0$ и

$$(mE - P_1)^{-1} = \frac{1}{m}\left(E - \frac{P_{-1}}{m}\right)^{-1} = \frac{1}{m}E + \frac{1}{m^2}P_{-1} + \frac{1}{m^3}P_{-1}^2 + \dots \qquad (3.283)$$

$$(m = k, k + 1, \dots).$$

(Если λ_0 — характеристическое число матрицы

$$A = \|\, a_{ik} \,\|_1^n, \text{ то } |\,\lambda_0\,| \leqslant n \max_{1 \leqslant i, k \leqslant n} |\, a_{ik} \,|.$$

Действительно, пусть $Ax = \lambda_0 x$, где $x = \{x_1, x_2, ..., x_n\} \neq 0$. Тогда

$$\lambda_0 x_i = \sum_{k=1}^{n} a_{ik}x_k \qquad (i = 1, 2, \dots, n).$$

Пусть $|\, x_j \,| = \max\{|\, x_1 \,|, |\, x_2 \,|, \dots, |\, x_n \,|\}$. Тогда

$$|\,\lambda_0\,|\,|\,x_j\,| \leqslant \sum_{k=1}^{n} |\,a_{jk}\,|\,|\,x_k\,| \leqslant |\,x_j\,|\, n \max_{1 \leqslant i, k \leqslant n} |\,a_{ik}\,|.$$

Сокращая на $|\, x_j \,|$, получим нужное неравенство.)

В последней части этого равенства стоит сходящийся матричный ряд. Пользуясь этим рядом, мы из (3.281) можем все коэффициенты ряда (3.282) выразить однозначно через $a_0, a_1, \dots, a_{k-1}$ при помощи рекуррентных соотношений

$$a_m = \left(\frac{1}{m}E + \frac{1}{m^2}P_{-1} + \frac{1}{m^3}P_{-1}^2 + \dots\right)(f_{m-1} + P_0a_{m-1} + \dots + P_{m-k-1}a_k)$$

$$(m = k, k + 1, \dots), \qquad (3.284)$$

где

$$f_{m-1} = P_{m-k}a_{k-1} + \dots + P_{m-1}a_0 \qquad (m = k, k + 1, \dots). \quad (3.285)$$

Заметим, что ряд (3.282) формально удовлетворяет дифференциальному уравнению

$$\frac{dx^{(k)}}{dz} = P(z)\, x^{(k)} + f(z), \qquad (3.286)$$

где

$$f(z) = \sum_{m=k-1}^{\infty} f_m z^m = P(z)\,(a_0 + a_1 z + \ldots + a_{k-1} z^{k-1}) -$$
$$- a_1 - 2a_2 z - \ldots - (k-1)\, a_{k-1} z^{k-2}. \qquad (3.287)$$

Из (3.287) вытекает, что ряд

$$\sum_{m=k-1}^{\infty} f_m z^m$$

сходится при $|z| < r$, и потому существует такое число $N > 0$, что (здесь $\left\| \frac{N}{r^m} \right\|$ обозначает столбец, у которого все элементы равны одному и тому же числу $\frac{N}{r^m}$.)

$$\operatorname{mod} f_m \leq \left\| \frac{N}{r^m} \right\| \quad (m = k-1, k, \ldots). \qquad (3.288)$$

Из вида рекуррентных соотношений (3.284) следует, что, заменив в них матрицы P_{-1}, P_q, f_{m-1} мажорантными матрицами

$$p_{-1} I,\ \frac{p}{r^q} I,\ \left\| \frac{N}{r^{m-1}} \right\|,$$

а столбец a_m столбцом $\| \alpha_m \|$ (здесь $\| \alpha_m \|$ обозначает столбец $(\alpha_m, \alpha_m, \ldots, \alpha_m)$ (α_m — число; $m = k, k+1, \ldots$).

$(m = k, k+1, \ldots;\ q = 0, 1, 2, \ldots)$, мы получим соотношения, определяющие верхние границы $\| \alpha_m \|$ для mod a_m:

$$\operatorname{mod} a_m \leq \| \alpha_m \| \qquad (3.289)$$

Следовательно, ряд

$$\xi^{(k)} = \alpha_k z^k + \alpha_{k+1} z^{k+1} + \ldots \qquad (3.290)$$

после почленного умножения на столбец $\| 1 \|$ будет мажорантным рядом для ряда (3.282).

Заменив в (3.286) матричные коэффициенты P_{-1}, P_q, f_m рядов

$$P(z) = \frac{P_{-1}}{z} + \sum_{q=0}^{\infty} P_q z^q, \qquad f(z) = \sum_{m=k-1}^{\infty} f_m z^m$$

соответствующими мажорантными матрицами $p_{-1}I$, $\frac{p}{r^q}I$, $\left\| \frac{N}{r^m} \right\|$, а также заменив $x^{(k)}$ на $\| \xi^{(k)} \|$, мы получим дифференциальное уравнение для $\xi^{(k)}$:

$$\frac{d\xi^{(k)}}{dz} = n\left(\frac{p_{-1}}{z} + \frac{p}{1-\frac{z}{r}}\right)\xi^{(k)} + \frac{N\frac{z^{k-1}}{r^{k-1}}}{1-\frac{z}{r}}. \qquad (3.291)$$

Это линейное дифференциальное уравнение имеет частное решение

$$\xi^{(k)} = \frac{N}{r^{k-1}} \frac{z^{np_{-1}}}{\left(1-\frac{z}{r}\right)^{npr}} \int_0^z z^{k-np_{-1}-1}\left(1-\frac{z}{r}\right)^{npr-1} dz, \qquad (3.292)$$

которое регулярно в точке $z = 0$ и в окрестности этой точки разлагается в сходящийся при $|z| < r$ степенной ряд (3.290).

Из сходимости мажорантного ряда (3.290) следует сходимость ряда (3.282) при $|z| < r$. Лемма доказана.

Замечание 1. Приведенное доказательство позволяет определить все регулярные в особой точке решения дифференциальной системы (3.279), если таковые существуют.

Для существования регулярных решений (не равных тождественно нулю) необходимо и достаточно, чтобы матрица вычетов P_{-1} имела целое неотрицательное характеристическое число. Если s — наибольшее такое целое характеристическое число, то из первых $s+1$ уравнений (3.281) можно определить не обращающиеся одновременно в нуль столбцы $a_0, a_1, ..., a_s$, поскольку определитель соответствующей системы линейных однородных уравнений равен нулю:

$$\Delta = | P_{-1} | | E - P_{-1} | | \ldots | sE - P_{-1} | = 0.$$

Из остальных уравнений (3.281) столбцы a_{s+1}, a_{s+2}, ... однозначно выразятся через $a_0, a_1, ..., a_s$. Полученный ряд (3.278) сходится согласно лемме. Таким образом, линейно независимые решения первых $s + 1$ уравнений (3.281) определяют все линейно независимые регулярные в особой точке $z = 0$ решения системы (3.279).

Если $z = 0$ есть особая точка, то задание начального значения a_0 для регулярного в этой точке решения (3.279) (если таковое существует) не определяет однозначно этого решения. Однако решение, регулярное в регулярной особой точке, определяется однозначно, если заданы $a_0, a_1, ..., a_s$, т. е. если заданы начальные значения при $z = 0$ самого решения и его первых s производных (s — наибольшее неотрицательное целое характеристическое число матрицы вычетов P_{-1}).

Замечание 2. Доказанная лемма сохраняет свою силу и при $P_{-1} = 0$. В этом случае в доказательстве леммы в качестве p_{-1} можно взять любое положительное число. При $P_{-1} = 0$ лемма утверждает известное положение о существовании регулярного решения в окрестности регулярной точки системы. В этом случае решение однозначно определяется заданием a_0.

4. Пусть дана система

$$\frac{dX}{dz} = P(z)X, \qquad (3.293)$$

где

$$P(z) = \frac{P_{-1}}{z} + \sum_{m=0}^{\infty} P_m z^m$$

и ряд, стоящий в правой части, сходится при $|z| < r$.

Пусть, далее, полагая

$$X = A(z)Y \qquad (3.294)$$

и подставляя вместо A (z) ряд

$$A(z) = A_0 + A_1 z + A_2 z^2 + \ldots, \qquad (3.295)$$

мы после формальных преобразований получаем:

$$\frac{dY}{dz} = P^*(z)Y, \qquad (3.296)$$

где

$$P^*(z) = \frac{P^*_{-1}}{z} + \sum_{m=0}^{\infty} P^*_m z^m,$$

причем здесь, как и в выражении для P (z), ряд в правой части сходится при $|z| < r$.

Докажем что и ряд (3.295) сходится в окрестности $|z| < r$ точки $z = 0$.

Действительно, из (3.293), (3.294) и (3.296) следует, что ряд (3.295) формально удовлетворяет следующему дифференциальному матричному уравнению

$$\frac{dA}{dz} = P(z)A - AP^*(z). \qquad (3.297)$$

Мы будем рассматривать A как вектор (столбец) в пространстве всех матриц n-го порядка, т. е. в пространстве n^2 измерений. Если мы в этом пространстве определим линейный оператор $\widehat{P}(z)$ над матрицей A, аналитически зависящий от параметра z, при помощи равенства

$$\widehat{P}(z)[A] = P(z)A - AP^*(z), \tag{3.298}$$

то дифференциальное уравнение (3.297) можно будет записать в виде

$$\frac{dA}{dz} = \widehat{P}(z)[A]. \tag{3.299}$$

Правую часть этого уравнения можно рассматривать как произведение матрицы $\widehat{P}(z)$ порядка n^2 на столбец A из n^2 элементов. Из формулы (3.298) видно, что точка z = 0 является регулярной особой точкой для системы (3.299). Ряд (3.295) формально удовлетворяет этой системе. Поэтому, применяя лемму, заключаем, что ряд (3.295) сходится в окрестности $|z| < r$ точки z = 0.

В частности, сходится и ряд для A (z) в формуле (3.277).

Таким образом, нами доказана следующая

Теорема 2. *Всякая система*

$$\frac{dX}{dz} = P(z)X \tag{3.300}$$

с регулярной особой точкой z = 0:

$$P(z) = \frac{P_{-1}}{z} + \sum_{m=0}^{\infty} P_m z^m$$

имеет решение вида

$$X = A(z)\, z^M z^U, \tag{3.301}$$

где A (z) — матричная функция, регулярная при z = 0 *и обращающаяся в этой точке в единичную матрицу E, а M и U — постоянные матрицы, причем M имеет простую структуру и целые характеристические числа, и разность между любыми двумя различными характеристическими числами матрицы U не есть целое число.*

Если матрица P_{-1} приводится к жордановой форме при помощи неособенной матрицы T

$$P_{-1} = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2, \ldots, \lambda_s E_s + H_s\}\, T^{-1} \tag{3.302}$$

$$(\mathrm{Re}\,\lambda_1 \geqslant \mathrm{Re}\,\lambda_2 \geqslant \ldots \geqslant Re\,\lambda_s),$$

то можно взять M и U в виде

$$M = T\{m_1E_1, m_2E_2, \ldots, m_sE_s\}T^{-1},$$

$$U = T\begin{pmatrix} \tilde{\lambda}_1E_1 + H_1 & B_{12} & \ldots & B_{1s} \\ 0 & \tilde{\lambda}_2E_2 + H_2 & \ldots & B_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \tilde{\lambda}_sE_s + H_s \end{pmatrix}T^{-1},$$

(3.303-3.304)

где

$$m_i = [\lambda_i], \quad \tilde{\lambda}_i = \lambda_i - m_i, \quad (i=1,2,\ldots, s), \qquad (3.305)$$

B_{ik} *— правильные нижние треугольные матрицы* ($i,k=1,2,\ldots, s$)*, причем* $B_{ik} = 0$*, если* $\lambda_i - \lambda_k$ *не есть целое положителъное число* ($i,k = 1, 2, \ldots, s$)*. В частном случае, когда ни одна из разностей* $\lambda_i - \lambda_k$ ($i,k = 1, 2, \ldots, s$) *не равна целому положительному числу, в формуле* (3.301) *можно положить* $M = 0, U = P_{-1}$*, т. е. в этом случае решение представимо в виде*

$$X = A(z)z^{P_{-1}}. \qquad (3.306)$$

Замечание 1. Обращаем внимание на то, что в настоящем параграфе был установлен алгоритм для определения коэффициентов ряда

$$A(z) = \sum_{m=0}^{\infty} A_m z^m$$

($A_0=E$) через коэффициенты P_m ряда для $P(z)$. Кроме того, доказанная теорема определяет и интегральную подстановку V, на которую умножается решение (3.301) при однократном обходе особой точки $z = 0$ в положительном направлении:

$$V = e^{2\pi iU}.$$

Замечание 2. Из формулировки теоремы следует, что

$$B_{ik} = 0 \text{ при } \tilde{\lambda}_i \neq \tilde{\lambda}_k \quad (i, k = 1, 2, \ldots, s).$$

Поэтому матрицы

$$\tilde{\Lambda} = T\{\tilde{\lambda}_1E_1, \tilde{\lambda}_2E_2, \ldots, \tilde{\lambda}_sE_s\}T^{-1} \text{ и } \tilde{U} = T\begin{pmatrix} 0 & B_{12} & \ldots & B_{1s} \\ 0 & 0 & \ldots & B_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{pmatrix}T^{-1} \qquad (3.307)$$

перестановочны между собой:

$$\tilde{\Lambda}\tilde{U} = \tilde{U}\tilde{\Lambda}.$$

Отсюда

$$z^M z^U = z^M z^{\tilde{\Lambda}+\tilde{U}} = z^M z^{\tilde{\Lambda}} z^{\tilde{U}} = z^{\Lambda} z^{\tilde{U}}, \qquad (3.308)$$

где

$$\Lambda = M + \tilde{\Lambda} = T\{\lambda_1, \lambda_2, \ldots, \lambda_n\} T^{-1}, \qquad (3.309)$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$ — все характеристические числа матрицы P_{-1} расположенные в порядке $\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n$.
С другой стороны,

$$z^{\tilde{U}} = h(\tilde{U}),$$

где $h(\lambda)$ — интерполяционный многочлен Лагранжа-Сильвестра для функции

$$f(\lambda) = z^{\lambda}.$$

Поскольку все характеристические числа матрицы U равны нулю, то h (А) линейно зависит от

$$f(0), f'(0), \ldots, f^{(g-1)}(0), \text{ т. е. от } 1, \ln z, \ldots, (\ln z)^{g-1}$$

(g — наименьший показатель, при котором $\tilde{U}^g = 0$). Поэтому

$$h(\lambda) = \sum_{j=0}^{g-1} h_j(\lambda)(\ln z)^j$$

и потому

$$z^{\tilde{U}} = h(\tilde{U}) = \sum_{j=0}^{g-1} h_j(\tilde{U})(\ln z)^j = T\begin{pmatrix} 1 & q_{12} & \cdots & q_{1n} \\ 0 & 1 & \cdots & q_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} T^{-1}, \qquad (3.310)$$

где q_{ij} $(i,j = 1, 2, ..., n;\ i<j)$ — многочлены от $\ln z$ степени ниже g.

В силу (3.301), (3.308), (3.309) и (3.310) частное решение системы (3.293) можно взять в виде

$$X = A(z)\begin{Vmatrix} z^{\lambda_1} & 0 & \ldots & 0 \\ 0 & z^{\lambda_2} & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & z^{\lambda_n} \end{Vmatrix} \begin{Vmatrix} 1 & q_{12} & \cdots & q_{1n} \\ 0 & 1 & \ldots & q_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 \end{Vmatrix}, \qquad (3.311)$$

Здесь $\lambda_1, \lambda_2, \ldots, \lambda_n$— характеристические числа матрицы P_{-1}, расположенные в порядке

$$\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n, \text{ а } q_{ij}\ (i, j = 1, 2, \ldots, n;\ i < j)$$

— многочлены от $\ln z$ степени не выше g — 1, где g — максимальное количество характеристических чисел λ_i, отличающихся между собой на целое число; A (z) — матричная функция, регулярная в точке $z = 0$, причем $A(0) = T(|T| \neq 0)$. Если матрица P_{-1} имеет жорданову форму, то $T = E$.

3.26. Приводимые аналитические системы

В качестве приложения теоремы предыдущего параграфа выясним, в каких случаях система

$$\frac{dX}{dt} = Q(t)X, \tag{3.312}$$

где

$$Q(t) = \sum_{m=1}^{\infty} \frac{Q_m}{t^m} \tag{3.313}$$

— сходящийся ряд при $t>t_0$, является приводимой (по Ляпунову), т. е. в каких случаях существует решение системы вида

$$X = L(t)e^{Bt}, \tag{3.314}$$

где $L(t)$ — матрица Ляпунова (т. е. $L(t)$ удовлетворяет ранее приведенным условиям 1° — 3°, а B— постоянная матрица (если имеет место равенство (3.314), то преобразование Ляпунова $X=L(t)Y$ переводит систему (3.312) в систему $\frac{dY}{dt} = BY$). Здесь X, Q — матрицы с комплексными элементами, а t — вещественный аргумент.

Сделаем преобразование

$$z = \frac{1}{t}.$$

Тогда система (3.312) перепишется в виде

$$\frac{dX}{dz} = P(z)X, \tag{3.315}$$

где

$$P(z) = -z^{-2}Q\left(\frac{1}{z}\right) = -\frac{Q_1}{z} - \sum_{m=0}^{\infty} Q_{m+2}z^m. \tag{3.316}$$

Ряд, стоящий в правой части выражения для P (z), сходится при

$$|z| < \frac{1}{t_0}.$$

Могут представиться два случая:

1) $Q_1 = 0$. В этом случае точка $z = 0$ не является особой для системы (3.315). Эта система имеет решение, регулярное и нормированное в точке z = 0. Это решение задается сходящимся степенным рядом

$$X(z) = E + X_1 z + X_2 z^2 + \ldots \qquad \left(|z| < \frac{1}{t_0}\right).$$

Полагая

$$L(t) = X\left(\frac{1}{t}\right), \quad B = 0,$$

получим искомое представление (3.314). Система приводима.

2) $Q_1 \neq 0$. В этом случае система (3.315) имеет регулярную особую точку в точке $z = 0$.

Не нарушая общности рассуждений, можно считать матрицу вычетов $P_{-1} = — Q_1$ приведенной к жордановой форме, в которой диагональные элементы $\lambda_1, \lambda_2, \ldots, \lambda_n$ расположены в порядке $\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n$.

Тогда в формуле (3.311) $T = E$, и потому система (3.315) имеет решение

$$X = A(z) \begin{Vmatrix} z^{\lambda_1} & 0 & \ldots & 0 \\ 0 & z^{\lambda_2} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & z^{\lambda_n} \end{Vmatrix} \begin{Vmatrix} 1 & q_{12} & \ldots & q_{1n} \\ 0 & 1 & \ldots & q_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{Vmatrix},$$

где функция $A(z)$ регулярна при $z=0$ и принимает в этой точке значение E, а q_{ik} $(i, k = 1, 2, \ldots, n;\ i < k)$— многочлены от $\ln z$. Заменяя здесь z на $\frac{1}{t}$, будем иметь:

$$X = A\left(\frac{1}{t}\right) \begin{Vmatrix} \left(\frac{1}{t}\right)^{\lambda_1} & 0 & \ldots & 0 \\ 0 & \left(\frac{1}{t}\right)^{\lambda_2} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \left(\frac{1}{t}\right)^{\lambda_n} \end{Vmatrix} \begin{Vmatrix} 1 & q_{12}\left(\ln\frac{1}{t}\right) & \ldots & q_{1n}\left(\ln\frac{1}{t}\right) \\ 0 & 1 & \ldots & q_{2n}\left(\ln\frac{1}{t}\right) \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{Vmatrix}. \qquad (3.317)$$

Так как преобразование

$$X = A\left(\frac{1}{t}\right) Y$$

является преобразованием Ляпунова, то система (3.312) будет приводимой к некоторой системе с постоянными коэффициентами в том и только в том случае, когда произведение

$$L_1(t) = \begin{Vmatrix} t^{-\lambda_1} & 0 & \dots & 0 \\ 0 & t^{-\lambda_2} & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & t^{-\lambda_n} \end{Vmatrix} \cdot \begin{Vmatrix} 1 & q_{12}\left(\ln\frac{1}{t}\right) & \dots & q_{1n}\left(\ln\frac{1}{t}\right) \\ 0 & 1 & \dots & q_{2n}\left(\ln\frac{1}{t}\right) \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{Vmatrix} e^{-Bt}, \qquad (3.318)$$

где B — некоторая постоянная матрица, будет матрицей Ляпунова, т. е. когда матрицы $L_1(t)$, $\frac{dL_1}{dt}$ и $L_1^{-1}(t)$ будут ограничены. При этом, как следует из теоремы Еругина (п. 3.18), матрицу B можно считать матрицей с вещественными характеристическими числами.

Из ограниченности матриц $L_1(t)$ и $L_1^{-1}(t)$ при $t > t_0$ вытекает, что все характеристические числа матрицы B должны равняться нулю. Это следует из выражения для e^{Bt} и e^{-Bt}, получаемого из (3.318). Кроме того, все числа λ_1, λ_2, ..., λ_n должны быть чисто мнимыми, поскольку согласно (3.318) из ограниченности элементов последней строки в $L_1(t)$ и первого столбца в $L_1^{-1}(t)$ вытекает, что $\operatorname{Re}\lambda_n \geqslant 0$ и $\operatorname{Re}\lambda_1 \leqslant 0$.

Но если все характеристические числа матрицы P_{-1} чисто мнимы, то разность между любыми двумя различными характеристическими числами матрицы P_{-1} не равна целому числу. Поэтому имеет место формула (3.306):

$$X = A(z)\, z^{P_{-1}} = A\left(\frac{1}{t}\right) t^{Q_1},$$

и для приводимости системы необходимо и достаточно, чтобы матрица

$$L_2(t) = t^{Q_1} e^{-Bt} \qquad (3.319)$$

вместе со своей обратной была бы ограничена при $t > t_0$.

Поскольку все характеристические числа матрицы B должны равняться нулю, то минимальный многочлен для матрицы B имеет вид λ^d. Обозначим через

$$\psi(\lambda) = (\lambda - \mu_1)^{c_1} (\lambda - \mu_2)^{c_2} \dots (\lambda - \mu_u)^{c_u} \quad (\mu_i \neq \mu_k \quad \text{при} \quad i \neq k)$$

минимальный многочлен матрицы Q_1. Поскольку $Q_1 = -P_{-1}$, то числа μ_1, μ_2, ..., μ_n отличаются знаком от соответствующих чисел λ_i и потому все они — чисто мнимые числа. Тогда [см. формулы (3.169), (3.170)]

$$t^{Q_1} = \sum_{k=1}^{u} [U_{k0} + U_{k1}\ln t + \dots + U_{k,c_k-1}(\ln t)^{c_k-1}]\, t^{\mu_k},$$

$$e^{Bt} = V_0 + V_1 t + \dots + V_{d-1} t^{d-1}. \qquad (3.320\text{-}3.321)$$

Подставляя эти выражения в равенство

$$L_2(t)\, e^{Bt} = t^{Q_1},$$

получим:

$$[L_2(t)\, V_{d-1} + (*)]\, t^{d-1} = Z_0(t)\, (\ln t)^{c-1}, \qquad (3.322)$$

где c — наибольшее из чисел c_1, c_2, ..., c_n, (*) обозначает матрицу, стремящуюся к нулю при $t \to \infty$, а $Z_0(t)$ — ограниченная матрица при $t > t_0$.

Так как матрицы, стоящие в левой и правой частях равенства (3.322), должны иметь одинаковый порядок роста при $t \to \infty$, то

$$d = c = 1,$$

т. е.

$$B = 0,$$

и матрица Q_1 имеет простые элементарные делители.

Обратно, если матрица Q_1 имеет простые элементарные делители и чисто мнимые характеристические числа $\mu_1, \mu_2, \ldots, \mu_n$, то

$$X = A(z)\, z^{-Q_1} = A(z)\, \| z^{-\mu_i} \delta_{ik} \|_1^n$$

есть решение системы (316). Полагая здесь $s = \frac{1}{t}$, найдем:

$$X = A\left(\frac{1}{t}\right) \| t^{\mu_i} \delta_{ik} \|_1^n.$$

Функция $X(t)$ вместе с $\frac{dX(t)}{dt}$ и обратной матрицей $X^{-1}(t)$ ограничена при $t > t_0$. Поэтому система приводима ($B = 0$). Нами доказана

Теорема 3. *Система*

$$\frac{dX}{dt} = Q(t)\, X,$$

где матрица $Q(t)$ представима сходящимся при $t > t_0$ рядом

$$Q(t) = \frac{Q_1}{t} + \frac{Q_2}{t^2} + \ldots,$$

является приводимой в том и только в том случае, если у матрицы вычетов Q_1 все элементарные делители простые и все характеристические числа чисто мнимы.

3.27. Аналитические функции от многих матриц

Аналитическая функция от m матриц n-го порядка $X_1, X_2, ..., X_m$ может быть задана при помощи ряда

$$F(X_1, X_2, \ldots, X_m) = \alpha_0 + \sum_{\nu=1}^{\infty} \sum_{j_1, j_2, \ldots, j_\nu}^{(1 \ldots m)} \alpha_{j_1 j_2 \ldots j_\nu} X_{j_1} X_{j_2} \ldots X_{j_\nu}, \quad (3.323)$$

сходящегося для всех матриц n-го порядка X_j, удовлетворяющих неравенствам

$$\operatorname{mod} X_j < R_j \qquad (j = 1, 2, \ldots, m). \quad (3.324)$$

здесь коэффициенты

$$\alpha_0, s_{j_1 j_2 \ldots j_\nu} \qquad (j_1, j_2, \ldots, j_\nu = 1, 2, \ldots, m;\ \nu = 1, 2, 3, \ldots)$$

комплексные числа, Rj $(j = 1, 2, ..., m)$ — постоянные матрицы n-го порядка с положительными элементами и X_j $(j = 1,2, ..., m)$ — матрицы того же рядка, но переменные и с комплексными элементами.

Теория аналитических функций от нескольких матриц была развита И. А. Лаппо-Данилевским. На основе этой теории И. А. Лаппо-Данилевский провел фундаментальные исследования систем линейных дифференциальных уравнений с рациональными коэффициентами.

Система с рациональными коэффициентами путем надлежащего преобразования независимой переменной всегда может быть приведена к виду

$$\frac{dV}{dz} = \sum_{j=1}^{m} \left\{ \frac{U_{j0}}{(z-a_j)^{s_j}} + \frac{U_{j1}}{(z-a_j)^{s_j-1}} + \ldots + \frac{U_{j, s_j-1}}{z-a_j} \right\} X, \quad (3.325)$$

где U_{jk} — постоянные матрицы n-го порядка, a_j — комплексные числа, s_j — целые положительные числа

$$(k = 0, 1, \ldots, s_j - 1;\ j = 1, 2, \ldots, m)$$

В системе (3.325) все коэффициенты — правильные рациональные дроби относительно z. К такому виду приводятся любые рациональные коэффициенты, если при помощи дробно-линейного преобразования над переменной z перевести регулярную (для всех коэффициентов) конечную точку $z = c$ в точку $z = \infty$.

Некоторые результаты Лаппо-Данилевского мы проиллюстрируем на частном случае так называемых *регулярных* систем. Последние характеризуются условием $s_1 = s_2 = ... = s_m = 1$ и записываются в виде

$$\frac{dX}{dz} = \sum_{j=1}^{m} \frac{U_j}{z-a_j} X. \quad (3.326)$$

Следуя Лаппо-Данилевскому, введем в рассмотрение специальные аналитические функции — гиперлогарифмы, — определяемые следующими рекурентными соотношениями:

$$l_b(z; a_{j_1}) = \int_b^z \frac{dz}{z - a_{j_1}},$$

$$l_b(z; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) = \int_b^z \frac{l_b(z; a_{j_2}, a_{j_3}, \ldots, a_{j_\nu})}{z - a_{j_1}}\, dz.$$

Рассматривая точки a_1, a_2,…, a_m, ∞ как точки разветвления логарифмического типа, построим соответствующую риманову поверхность S $(a_1\ a_2,\ \ldots\ \ldots,\ a_m;\ \infty)$. Каждый гиперлогарифм будет однозначной функцией на этой поверхности. С другой стороны, матрицант системы (3.326) Ω_b^z (т. е. нормированное в точке $z=b$ решение), будучи аналитически продолжен, также может быть рассматриваем как однозначная функция на S $(a_1,\ a_2,\ \ldots,\ a_m;\ \infty)$; при этом в качестве b может быть выбрана любая конечная точка на S, отличная от $a_1, a_2, \ldots, a_m$.

Для нормированного решения Ω_b^z Лаппо-Данилевский дает явное выражение через определяющие матрицы U_1, U_2, … , U_m системы (3.326) в виде ряда

$$\Omega_b^z = E + \sum_{\nu=1}^{\infty} \sum_{j_1, \ldots, j_\nu}^{(1 \ldots m)} l_b(z; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu})\, U_{j_1} U_{j_2} \ldots U_{j_\nu}. \quad (3.327)$$

Это разложение сходится равномерно относительно z при любых U_1, U_2, ..., U_m и представляет Ω_b^z в любой конечной области на поверхности S $(a_1, a_2, \ldots, a_m; \infty)$, если только эта область не содержит внутри и на границе точек $a_1, a_2, \ldots, a_m$.

Если ряд (3.323) сходится при любых матрицах X_1, X_2,...,X_m, то соответствующая функция $F(X_1,\ X_2,\ \ldots,\ X_m)$ называется *целой.* Ω_b^z представляет собой целую функцию от матриц U_1, U_2, ..., U_m.

Заставляя в формуле (327) аргумент z обойти точку a_j в положительном направлении один раз так, чтобы контур обхода не захватывал других точек a_i (при $i \neq j$), мы получим выражение для *интегральной подстановки* V_j, соответствующей точке $z = a_j$:

$$V_j = E + \sum_{\nu=1}^{\infty} \sum_{j_1, \ldots, j_\nu}^{(1 \ldots m)} p_j(b; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu})\, U_{j_1} U_{j_2} \ldots U_{j_\nu} \quad (3.328)$$

$$(j = 1, 2, \ldots, m),$$

где в понятных обозначениях

$$p_j(b; a_{j_1}) = \int\limits_{(a_j)} \frac{\partial z}{z - a_{j_1}},$$

$$p_j(b; a_{j_1}, a_{j_2}, \ldots, a_{j_\nu}) = \int\limits_{(a_j)} \frac{L_b(z; a_{j_2}, a_{j_3}, \ldots, a_{j_\nu})}{z - a_{j_1}}\, dz$$

$$\begin{pmatrix} j_1, j_2, \ldots, j_\nu, \quad j = 1, 2, \ldots, m; \\ \nu = 1, 2, 3, \ldots \end{pmatrix}.$$

Ряд (3.328), как и ряд (3.327), представляет собой целую функцию от $U_1, U_2, .., U_m$.

Обобщив теорию аналитических функций на случай бесконечного, но счетного множества матриц-аргументов X_1, X_2, X_3, ..., Лаппо-Данилевский использовал эту теорию для исследования поведения решения системы в окрестности иррегулярной особой точки. Мы приведем основной результат.

Нормированное решение Ω_b^z системы

$$\frac{dX}{dz} = \sum_{j=-q}^{+\infty} P_j z^j X,$$

где степенной ряд в правой части сходится при $|z| < r$ $(r > 1)$, может быть представлено рядом

$$\Omega_b^z = E + \sum_{\nu=1}^{\infty} \sum_{j_1, j_2, \ldots, j_\nu = -q}^{\infty} P_{j_1} \ldots$$

$$\ldots P_{j_\nu} \sum_{\mu=0}^{\nu} b^{j_{\mu+1} + \ldots + j_\nu + \nu - \mu} z^{j_1 + \ldots + j_\mu + \mu} \sum_{\lambda=0}^{n-\mu} \alpha^{*(\lambda)}_{j_{\mu+1}, \ldots, j_\nu} \ln^\lambda b \sum_{\varkappa=0}^{\mu} \alpha^{(\varkappa)}_{j_1, \ldots, j_\mu} \ln^\varkappa z. \quad (3.329)$$

Здесь $\alpha^{*(\lambda)}_{j_{\mu+1}, \ldots, j_\nu}$ и $\alpha^{(\varkappa)}_{j_1, \ldots, j_\mu}$ — скалярные коэффициенты, определяемые по специальным формулам. Ряд (3.329) сходится при любых матрицах P_1, P_2, ... в кольце

$$\rho < |z| < r$$

(ρ — произвольное положительное число, меньшее r). Этому кольцу должна принадлежать и точка b $(\rho < b < r)$.

Микромодуль 8
Примеры решения типовых задач

1. Представление многочлена многочленной матрицей.

Матричный многочлен может быть представлен *многочленноq матрицей (λ-матрицей),* элементы которой являются многочленами относительно скалярной переменной x (или λ). К этому типу матриц относятся, в частности, характеристическая и присоединенная матрицы. Например:

$$A=\begin{bmatrix}4 & -8 & 1\\ 5 & -9 & 1\\ 4 & -6 & -1\end{bmatrix};\quad F(\lambda)=|\lambda E-A|=\begin{bmatrix}\lambda-4 & 8 & -1\\ -5 & \lambda+9 & -1\\ -4 & 6 & \lambda+1\end{bmatrix};$$

$$G(\lambda)=\mathrm{Adj}\,[\lambda E-A]=\begin{bmatrix}\lambda^2+10\lambda+15 & -8\lambda-14 & \lambda+1\\ 5\lambda+9 & \lambda^2-3\lambda-8 & \lambda+1\\ 4\lambda+6 & -6\lambda-8 & \lambda^2+5\lambda+4\end{bmatrix}=$$

$$=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\lambda^2+\begin{bmatrix}10 & -8 & 1\\ 5 & -3 & 1\\ 4 & -6 & 5\end{bmatrix}\lambda+\begin{bmatrix}15 & -14 & 1\\ 9 & -8 & 1\\ 6 & -8 & 4\end{bmatrix};$$

$$\Delta(\lambda)=\lambda^3+6\lambda^2+11\lambda+6=(\lambda+1)(\lambda+2)(\lambda+3).$$

Здесь $F(\lambda)$ и $G(\lambda)$ — многочленные матрицы, которые могут быть представлены в виде матричных многочленов; $\Delta(\lambda)$ — скалярный многочлен от λ. Экспоненциальная функция от A выражается многочленом от матрицы A:

$$e^A=1+A+\frac{A^2}{2!}+\frac{A^3}{3!}+\cdots=$$

$$=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}+\begin{bmatrix}4 & -8 & 1\\ 5 & -9 & 1\\ 4 & -6 & -1\end{bmatrix}+\frac{1}{2!}\begin{bmatrix}4 & -8 & 1\\ 5 & -9 & 1\\ 4 & -6 & -1\end{bmatrix}^2+\cdots$$

2. Пример применения теоремы Кэли-Гамильтона.

Проиллюстрируем теорему Кэли-Гамильтона на примере матрицы из примера 1:

$$\Delta(A) = A^3 + 6A^2 + 11A + 6E =$$

$$= \begin{bmatrix} 70 & -116 & 19 \\ 71 & -117 & 19 \\ 64 & -102 & 11 \end{bmatrix} + 6 \begin{bmatrix} -20 & 34 & -5 \\ -21 & 35 & -5 \\ -18 & 28 & -1 \end{bmatrix} + 11 \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix} +$$

$$+ 6 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0;$$

$$\Delta(A) = (A + 1E)(A + 2E)(A + 3E) =$$

$$= \begin{bmatrix} 5 & -8 & 1 \\ 5 & -8 & 1 \\ 4 & -6 & 0 \end{bmatrix} \begin{bmatrix} 6 & -8 & 1 \\ 5 & -7 & 1 \\ 4 & -6 & 1 \end{bmatrix} \begin{bmatrix} 7 & -8 & 1 \\ 5 & -6 & 1 \\ 4 & -6 & 2 \end{bmatrix} = 0.$$

Важным следствием полученной теоремы является возможность представления любого многочлена $f(A)$ от квадратной матрицы A n-го порядка многочленом $r(A)$ степени n — 1, т.е. $f(A) = r(A)$. Пусть, например,

$$f(A) = A^4 + 4A^3 + 2A^2 - 12A - 10E.$$

Разделив соответствующий скалярный многочлен

$$f(\lambda) = \lambda^4 + 4\lambda^3 + 2\lambda^{2-12}\lambda - 10$$

на характеристический многочлен

$$\Delta(\lambda) = \lambda^3 + 6\lambda^2 + 11\lambda + 6,$$

получим остаток $r(\lambda) = 3\lambda^2 + 4\lambda + 2$.

Следовательно, $f(A) = 3A^2 + 4A + 2E$.

Теорему Кэли-Гамильтона можно также использовать для вычисления степеней матрицы и определение обратной матрицы. Так как $\Delta(A) = 0$, то и $A^k\Delta(A) = 0$, где k — любое целое число. Поэтому любая степень матрицы линейно выражается через ее первые $(n - 1)$ степеней. Так, для нашего примера:

$$A^3 = -6A^2 - 11A - 6E;$$

$$A^4 = -6A^3 - 11A^2 - 6A = 25A^2 + 60A + 36E \text{ и т.д.}$$

Для обратной матрицы необходимое соотношение получаем умножением $\Delta(A) = 0$ на A^{-1} т.е. $A^{-1}(A^3 + 6A^2 + 11A + 6E) = 0$,

откуда имеем: $A^{-1} = (-1/6)(A^2 + 6A + 11E)$.

3. Пример на применение минимального многочлена.

Естественно стремиться свести функцию $f(A)$ к многочлену $r(A)$ возможно меньшей степени. Поскольку степень $r(A)$ всегда на единицу ниже степени аннулирующего многочлена, то эта задача означает поиск аннулирующего многочлена $\psi(\lambda)$ наименьшей степени (со старшим коэффициентом, равным единицы), называемого *минимальным многочленом*.

Если все собственные значения матрицы A различны, то характеристический многочлен $\Delta(\lambda)$ является одновременно и минимальным. В общем же случае может быть несколько аннулирующих многочленов, степень которых не превышает n, и среди них только один минимальный многочлен степени $m \leq n$.

Рассмотрим пример:

$$A = \begin{bmatrix} -3 & 3 & -2 \\ 1 & -5 & 2 \\ 1 & -3 & 0 \end{bmatrix}; \quad F(\lambda) = \begin{bmatrix} \lambda+3 & -3 & 2 \\ -1 & \lambda+5 & -2 \\ -1 & 3 & \lambda \end{bmatrix};$$

$$\Delta(\lambda) = \lambda^3 + 8\lambda^2 + 20\lambda + 16 = (\lambda + 2)^2(\lambda + 4).$$

Матрица имеет двукратное собственное значение $\lambda_1 = -2$ и простое $\lambda_2 = -4$. Присоединенная матрица:

$$G(\lambda) = \begin{bmatrix} (\lambda+2)(\lambda+3) & 3(\lambda+2) & -2(\lambda+2) \\ \lambda+2 & (\lambda+2)(\lambda+1) & 2(\lambda+2) \\ \lambda+2 & -3(\lambda+2) & (\lambda+2)(\lambda+6) \end{bmatrix} = (\lambda+2)C(\lambda).$$

Общий наибольший делитель d=(λ+2), следовательно:

$$\psi(\lambda) = \frac{\Delta(\lambda)}{d(\lambda)} = (\lambda+2)(\lambda+4);$$

$$C(\lambda) = \begin{bmatrix} \lambda+3 & 3 & -2 \\ 1 & \lambda+1 & 2 \\ 1 & -3 & \lambda+6 \end{bmatrix}.$$

4. Интерполяционный многочлен.

Определим экспоненциальную функцию от матрицы с помощью интерполяционного многочлена:

$$A = \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix}.$$

Так как все собственные значения матрицы различны (λ_1=-1; λ_2=-2; λ_3=-3), то $\psi(\lambda)=\Delta(\lambda)$. Тогда:

$$V = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & 4 \\ 1 & -3 & 9 \end{bmatrix}; \quad V^{-1} = \frac{1}{2}\begin{bmatrix} 6 & -6 & 2 \\ 5 & -8 & 3 \\ 1 & -2 & 1 \end{bmatrix};$$

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 6 & -6 & 2 \\ 5 & -8 & 3 \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} e^{-t} \\ e^{-2t} \\ e^{-3t} \end{bmatrix} = \begin{bmatrix} 3e^{-t} - 3e^{-2t} + e^{-3t} \\ \frac{5}{2}e^{-t} - 4e^{-2t} + \frac{3}{2}e^{-3t} \\ \frac{1}{2}e^{-t} - e^{-2t} + \frac{1}{2}e^{-3t} \end{bmatrix};$$

$$\exp(At) = \alpha_0 E + \alpha_1 A + \alpha_2 A^2 =$$

$$= \alpha_0 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} -20 & 34 & -5 \\ -21 & 35 & -5 \\ -18 & 28 & -1 \end{bmatrix}.$$

5. Интерполяционный многочлен Лагранжа-Сильвестра.

Найти интерполяционный многочлен Лагранжа-Сильвестра для матрицы *A* из примера 4.

Для матрицы *A,* все собственные значения которой различны (λ_1=-1; λ_2=-2; λ_3=-3) и $\psi(\lambda)=\Delta(\lambda)=(\lambda+1)(\lambda+2)(\lambda+3)$, имеем:

$$r(\lambda) = \frac{(\lambda-\lambda_2)(\lambda-\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)} f(\lambda_1) + \frac{(\lambda-\lambda_1)(\lambda-\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)} f(\lambda_2) +$$
$$+ \frac{(\lambda-\lambda_1)(\lambda-\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)} f(\lambda_3) = \frac{1}{2}(\lambda+2)(\lambda+3) f(-1) -$$
$$-(\lambda+1)(\lambda+3) f(-2) + \frac{1}{2}(\lambda+1)(\lambda+2) f(-3).$$

Для экспоненциальной функции $f(\lambda)=\exp(\lambda t)$ находим

$$f(-1)=\exp(-t),\ f(-2)=\exp(-2t),\ f(-3)=\exp(-3t).$$

Подставляя вместо λ матрицу *A*, получаем

$$e^{At} = \frac{1}{2}(A+2E)(A+3E)e^{-t} - (A+E)(A+3E)e^{-2t} +$$
$$+ \frac{1}{2}(A+E)(A+2E)e^{-3t},$$

Проиллюстрируем применение полученной формулы на примере матрицы с кратными собственными значениями:

$$A = \begin{bmatrix} -2 & 2 & -3 \\ -10 & 4 & -5 \\ -5 & 4 & -6 \end{bmatrix}; \quad F(\lambda) = \begin{bmatrix} \lambda+2 & -2 & 3 \\ 10 & \lambda-4 & 5 \\ 5 & -4 & \lambda+6 \end{bmatrix};$$

$$\Delta(\lambda) = \lambda^3 + 4\lambda^2 + 5\lambda + 2 = (\lambda+1)^2(\lambda+2);$$
$$\lambda_1 = -1\ (m_1 = 2); \quad \lambda_2 = -2\ (m_2 = 1).$$

Так как $F(\lambda_1)$ — просто вырожденная матрица (ее дефект равен единице), то $\psi(\lambda) = \Delta(\lambda)$, а также $\psi_1(\lambda)=\lambda+2$ и $\psi_2(\lambda)=(\lambda + 1)^2$. Поэтому можно записать:

$$r(\lambda) = [\mu_{11} + \mu_{12}(\lambda - \lambda_1)]\, \psi_1(\lambda) + \mu_{21}\psi_2(\lambda),$$

где

$$\mu_{11} = \frac{f(\lambda_1)}{\psi_1(\lambda_1)} = f(\lambda_1);$$

$$\mu_{12} = \left[\frac{f(\lambda)}{\psi_1(\lambda)}\right]'_{\lambda=\lambda_1} = \frac{f'(\lambda_1)\,\psi_1(\lambda_1) - f(\lambda_1)\,\psi_1(\lambda_1)}{\psi_1(\lambda_1)^2} = f'(\lambda_1) - f(\lambda_1);$$

$$\mu_{21} = \frac{f(\lambda_2)}{\psi_2(\lambda_2)} = f(\lambda_2).$$

Подставляя значение коэффициентов, получаем

$$r(\lambda) = \{f(\lambda_1) + [f'(\lambda_1) - f(\lambda_1)](\lambda + 1)\}(\lambda + 2) + f(\lambda_2)(\lambda + 1)^2 =$$
$$= -f(\lambda_1)(\lambda + 2)\lambda + f'(\lambda_1)(\lambda + 1)(\lambda + 2) + f(\lambda_2)(\lambda + 1)^2.$$

Обозначим многочлены от λ, входящие в полученные выражения,

$$\varphi_{11} = (\lambda + 2)\,\lambda;\quad \varphi_{12} = (\lambda + 1)(\lambda + 2);\quad \varphi_{21} = (\lambda + 1)^2;$$

тогда:

$$r(\lambda) = -f(\lambda_1)\,\varphi_{11}(\lambda) + f'(\lambda_1)\,\varphi_{12}(\lambda) + f(\lambda_2)\,\varphi_{21}(\lambda).$$

Заменив λ на матрицу *A,* с учетом *r(A) = f(A),* получим:

$$f(A) = -f(\lambda_1)(A + 2E)A + f'(\lambda_1)(A + E)(A + 2E) +$$
$$+ f(\lambda_2)(A + E)^2 = -f(\lambda_1)\begin{bmatrix} 0 & 2 & -3 \\ -10 & 6 & -5 \\ -5 & 4 & -4 \end{bmatrix}\begin{bmatrix} -2 & 2 & -3 \\ -10 & 4 & -5 \\ -5 & 4 & -6 \end{bmatrix} +$$
$$+ f'(\lambda_1)\begin{bmatrix} -1 & 2 & -3 \\ -10 & 5 & -5 \\ -5 & 4 & -5 \end{bmatrix}\begin{bmatrix} 0 & 2 & -3 \\ -10 & 6 & -5 \\ -5 & 4 & -4 \end{bmatrix} + f(\lambda_2)\begin{bmatrix} -1 & 2 & -3 \\ -10 & 5 & -5 \\ -5 & 4 & -5 \end{bmatrix}^2 =$$
$$= f(\lambda_1)\begin{bmatrix} 5 & 4 & -8 \\ 15 & 16 & -30 \\ 10 & 10 & -19 \end{bmatrix} + f'(\lambda_1)\begin{bmatrix} -5 & -2 & 5 \\ -25 & -10 & 25 \\ -15 & -6 & 15 \end{bmatrix} +$$
$$+ f(\lambda_2)\begin{bmatrix} -4 & -4 & 8 \\ -15 & -15 & 30 \\ -10 & -10 & 20 \end{bmatrix}.$$

Определив $f(\lambda_1)$, $f'(\lambda_1)$, $f(\lambda_2)$ для данной скалярной функции $f(\lambda)$ и подставив в это выражение, получим функцию от матрицы *f(A),* определенную на ее спектре.

6. Алгоритм Фаддєєва.

Алгоритм Фаддеева используется для определения присоединенной матрицы.

Приведем пример использования алгоритма Фаддеева на матрице из примера 1.

Полагая $G_1=E$, имеем:

$$B_1 = G_1A = A = \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix}; \quad a_1 = -\operatorname{tr} B_1 = -(4-9-1) = 6;$$

$$G_2 = B_1 + a_1E = \begin{bmatrix} 10 & -8 & 1 \\ 5 & -3 & 1 \\ 4 & -6 & 5 \end{bmatrix}; \quad B_2 = G_2A = \begin{bmatrix} 4 & -14 & 1 \\ 9 & -19 & 1 \\ 6 & -8 & -7 \end{bmatrix};$$

$$a_2 = -\frac{1}{2}\operatorname{tr} B_2 = -\frac{1}{2}(4-19-7) = 11;$$

$$G_3 = B_2 + a_2E = \begin{bmatrix} 15 & -14 & 1 \\ 9 & -8 & 1 \\ 6 & -8 & 4 \end{bmatrix}; \quad B_3 = G_3A = \begin{bmatrix} -6 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & -6 \end{bmatrix};$$

$$a_3 = -\frac{1}{3}\operatorname{tr} B_3 = -\frac{1}{3}(-6-6-6) = 6.$$

Соотношение $G_nA=-a_nE$ можно использовать для проверки правильности вычислений. Действительно, в нашем примере $G_3A=B_3=-a_3E$. Таким образом, получаем

$$\Delta(\lambda)=\lambda^3+6\lambda^2+11\lambda+6;$$

$$G(\lambda) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\lambda^2 + \begin{bmatrix} 10 & -8 & 1 \\ 5 & -3 & 1 \\ 4 & -6 & 5 \end{bmatrix}\lambda + \begin{bmatrix} 15 & -14 & 1 \\ 9 & -8 & 1 \\ 6 & -8 & 4 \end{bmatrix}.$$

Этот алгоритм требует выполнения $n - 1$ операций умножения матриц G_iA, каждая из которых сводится к n^3 операций умножения, т.е. всего $(n - 1)n^3 \approx n^4$. Кроме того, при каждом умножении матриц необходимо выполнить $(n-1)n^2$ операций сложения (или вычитания) чисел, т.е. всего $(n - 1)^2n^2 \approx n^4$. При этом вычитание близких по величине чисел может привести к существенному снижению точности.

7. Значение присоединенной матрицы.

Определим значение присоединенной матрицы и ее производной для матрицы третьего порядка из примера 5, собственные значения которой — двукратное $\lambda_1=-1$ и простое $\lambda_2=-2$.

Согласно полученным формулам

$$G(\lambda_1) = (A - \lambda_1 E)(A - \lambda_2 E);\ \ G'(\lambda_1) = A - \lambda_2 E;$$
$$G(\lambda_2) = (A - \lambda_1 E)^2;$$
$$G(\lambda_1) = \begin{bmatrix} -1 & 2 & -3 \\ -10 & 5 & -5 \\ -5 & 4 & -5 \end{bmatrix} \begin{bmatrix} 0 & 2 & -3 \\ -10 & 6 & -5 \\ -5 & 4 & -4 \end{bmatrix} = \begin{bmatrix} -5 & -2 & 5 \\ -25 & -10 & 25 \\ -15 & -6 & 15 \end{bmatrix};$$
$$G'(\lambda_1) = \begin{bmatrix} 0 & 2 & -3 \\ -10 & 6 & -5 \\ -5 & 4 & -4 \end{bmatrix};$$
$$G(\lambda_2) = \begin{bmatrix} -1 & 2 & -3 \\ -10 & 5 & -5 \\ -5 & 4 & -5 \end{bmatrix}^2 = \begin{bmatrix} -4 & -4 & 8 \\ -15 & -15 & 30 \\ -10 & -10 & 20 \end{bmatrix}.$$

Полученные результаты можно проверить подстановкой значений λ_1 и λ_2 в общее выражение для $G(\lambda)$.

8. Применение теоремы Сильвестра.

Использование теоремы Сильвестра в общем случае кратных характеристических чисел (нулей минимального многочлена) иллюстрируется на примере матрицы из примера 5, для которой $\lambda_1 = -1$ $(m_1 = 2)$ и $\lambda_2 = -2$ $(m_2 = 1)$:

$$f(A) = f(\lambda_1) Z_{11} + f'(\lambda_1) Z_{12} + f(\lambda_2) Z_{21};$$
$$Z_{11} = \frac{1}{1!\,0!} \left[\frac{C(\lambda)}{\psi_1(\lambda)}\right]'_{\lambda=\lambda_1} = \frac{C'(\lambda_1)\psi_1(\lambda_1) - C(\lambda_1)\psi'(\lambda_1)}{\psi_1^2(\lambda_1)};$$
$$Z_{12} = \frac{C(\lambda_1)}{\psi_1(\lambda_1)};\quad Z_{21} = \frac{C(\lambda_2)}{\psi_2(\lambda_2)}.$$

Так как $\psi_1(\lambda) = \lambda + 2$ и $\psi_2(\lambda) = (\lambda + 1)^2$, то $\psi_1(\lambda_1) = 1$, $\psi'_1(\lambda_1) = 1$; $\psi_2(\lambda_2) = 1$ и с учетом полученных в примере 7 значений $G(\lambda)$, которые в данном случае совпадают с соответствующими значениями $C(\lambda)$, имеем:

$$Z_{11} = \begin{bmatrix} 5 & 4 & -8 \\ 15 & 16 & -30 \\ 10 & 10 & -19 \end{bmatrix};\qquad Z_{12} = \begin{bmatrix} -5 & -2 & 5 \\ -25 & -10 & 25 \\ -15 & -6 & 15 \end{bmatrix};$$
$$Z_{21} = \begin{bmatrix} -4 & -4 & 8 \\ -15 & -15 & 30 \\ -10 & -10 & 20 \end{bmatrix}.$$

В результате по формуле Сильвестра получаем выражение для функции от матрицы $f(A)$, которое совпадает с полученным ранее в примере 5.

Для экспоненциальной функции от матрицы формула Сильвестра запишется следующим образом:

$$e^{At}=\sum_{k=1}^{q}\sum_{j=1}^{m_k} t^{j-1}e^{\lambda_k t}Z_{kj}.$$

Так, для нашего примера имеем:

$$e^{At}=e^{-t}\begin{bmatrix}5 & 4 & -8\\ 15 & 16 & -30\\ 10 & 10 & -19\end{bmatrix}+te^{-t}\begin{bmatrix}-5 & -2 & 5\\ -25 & -10 & 25\\ -15 & -6 & 15\end{bmatrix}+$$

$$+e^{-2t}\begin{bmatrix}-4 & -4 & 8\\ -15 & -15 & 30\\ -10 & -10 & 20\end{bmatrix}=$$

$$=\begin{bmatrix}5(1-t)e^{-t}-4e^{-2t} & 2(2-t)e^{-t}-4e^{-2t} & (-8+5t)e^{-t}+8e^{-2t}\\ 5(3-5t)e^{-t}-15e^{-2t} & 2(8-5t)e^{-t}-15e^{-2t} & 5(-6+5t)e^{-t}+30e^{-2t}\\ 5(2-3t)e^{-t}-10e^{-2t} & 2(5-3t)e^{-t}-10e^{-2t} & (-19+15t)e^{-t}+20e^{-2t}\end{bmatrix}.$$

Полученное выражение можно рассматривать как фундаментальную матрицу $\Phi(t)$ линейных дифференциальных уравнений с постоянными коэффициентами в нормальной форме $dx/dt=Ax$, в которой матрица A совпадает с заданной.

9. Компоненты матрицы. Так как компоненты матрицы Z_{kj} не зависят от вида функции $f(\lambda)$, то их можно определить из системы m уравнений, которые получаем подстановкой в основное уравнение теоремы Сильвестра m независимых аналитических функций. В качестве таких функций удобно использовать многочлены, получаемые последовательным делением минимального многочлена $\psi(\lambda)$ на простейшие сомножители. В результате всегда получим m функций, включая и $f(\lambda) = 1$.

Определим этим способом компоненты матрицы из примера 5. Так как $\psi(\lambda)=(\lambda+1)^2(\lambda+2)$, то в качестве простейших множителей принимаем $f_1(\lambda) = (\lambda + 1)^2$, $f_2(\lambda) = \lambda + 1$, $f_3(\lambda) = 1$. По основной формуле

$$f(A)= f(-1)Z_{11}+ f'(-1)Z_{12} + f(-2)Z_{21}.$$

Подставляя сюда простейшие многочлены, получаем систему уравнений:

$$(A + E)^2 = Z_{21}; \quad (A + E) = Z_{12}-Z_{21}; \quad E = Z_{11} + Z_{21},$$

откуда последовательной подстановкой находим:

$$Z_{21} =(A + E)^2; \quad Z_{12} = (A + E) - Z_{21};\ Z_{11} = E- Z_{21}.$$

Для данной матрицы *A* после выполнения соответствующих операций получаем компоненты, которые совпадают с найденными в примере 7.

Компоненты матрицы связаны общими соотношениями, которые можно использовать для проверки правильности вычислений:

$$\sum_{k=1}^{q} (\lambda_k Z_{k1} + Z_{k2}) = A; \quad \sum_{k=1}^{q} Z_{k1} = E; \quad Z_{k1}^2 = Z_{k1} \quad (k = 1,\ 2,\ \ldots,\ q);$$
$$Z_{kj} Z_{li} = 0; \quad (k \neq l; \quad 1 \leqslant j \leqslant m_k; \quad 1 \leqslant i \leqslant m_l).$$

Первая пара соотношений получается подстановкой в основную формулу функций $f(\lambda) = \lambda$ и $f(\lambda) = 1$.

Микромодуль 8
Индивидуальные тестовые задачи

1. Дана матрица

$$A = \begin{bmatrix} 7 & 4 & 3 & 2 \\ -3 & -2 & -5 & 2 \\ -6 & -4 & 0 & -4 \\ -6 & -4 & 1 & -5 \end{bmatrix},$$

собственные значения которой $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 2$, $\lambda_4 = -2$. Найдите общее выражение для функции от матрицы *A* с помощью:

а) интерполяционного многочлена

$$f(A) = \sum_{i=0}^{m-1} \alpha_i A^i;$$

б) интерполяционного многочлена Лагранжа;

в) теоремы Сильвестра;

г) решения уравнений для компонент матрицы *A*,

2. По результатам задачи 1 запишите функции e^A, $\sin A$, $\ln A$.

3. Для каждой из приведенных ниже матриц:

$$X = \begin{bmatrix} 0 & 1 & 3 \\ 6 & 0 & 2 \\ -5 & 2 & 4 \end{bmatrix}; \quad Y = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}; \quad Z = \begin{bmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{bmatrix}$$

а) найдите с помощью алгоритма Фаддеева выражение для $G(\lambda)$ и $\Delta(\lambda)$;

б) определите минимальный многочлен $\psi(\lambda)$ и приведенную присоединенную матрицу $C(\lambda)$;

в) получите значения $G(\lambda_k)$ и $\Delta(\lambda_k)$, а также их производных с помощью формул, приведенных в микромодуле 8 «Примеры решения типовых задач, пример 9», и проверьте результат непосредственной подстановкой значений $\lambda = \lambda_k$, в $G(\lambda)$ и $\Delta(\lambda)$, вычисленные по алгоритму Фаддеева.

4. Найдите общие выражения функций от матриц из предыдущей задачи с помощью:

а) интерполяционного многочлена Лагранжа-Сильвестра;

б) теоремы Сильвестра;

в) решения уравнений для компонент матрицы.

5. На основе общих выражений, полученных в предыдущей задаче, найдите e^X, e^Y, e^Z, а также $\sin X$, $\sin Y$, $\sin Z$.

6. Исходя из алгоритма Фаддеева, выведите рекуррентные соотношения для определения коэффициентов характеристического многочлена $\Delta(\lambda)$:

$$a_1 = -\operatorname{tr} A; \quad a_i = -\frac{1}{i}\sum_{j=1}^{i} a_{i-j} \operatorname{tr} A^j \quad (i = 2,\ 3,\ \ldots,\ n),$$

известные под названием формул Ньютона. Получите на основании этих формул характеристический многочлен матрицы из задачи 2.

7. Дана система дифференциальных уравнений:

$$\left.\begin{aligned}
\frac{dx_1}{dt} &= x_2 \\
\frac{dx_2}{dt} &= -4x_1 + 4x_2 \\
\frac{dx_3}{dt} &= -2x_1 + x_2 + 2x_3 + e^{3t}
\end{aligned}\right\}.$$

а) Воспользовавшись каким-либо способом определения функции от матрицы, найдите фундаментальную матрицу системы $\Phi(t) = e^{At}$.

б) Найдите решение системы дифференциальных уравнений при начальных значениях $x_{10} = 2$, $x_{20} = -1$, $x_{30} = 5$.

Микромодуль 9
Матричные преобразования

Основные типы матричных преобразований. При решении систем линейных алгебраических и дифференциальных уравнений уже встречались преобразования вещественных матриц (LU-разложение и приведение матрицы к диагональной форме), которые являются частными случаями общих типов матричных преобразований (рис. 3.1).

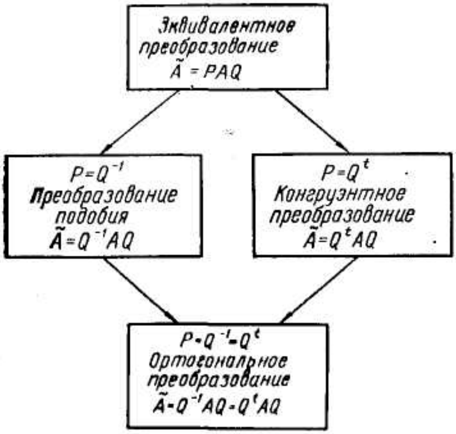


Рис. 3.1. Матричные преобразования.

Наиболее общим является *эквивалентное преобразование* $\tilde{A}= PAQ$, где P и Q — неособые квадратные матрицы. Матрицы A и $\tilde{A}$ называются *эквивалентными.* Треугольное разложение является частным случаем эквивалентных преобразований, когда $\tilde{A}$ — единичная матрица E. Из $PAQ = E$ следует $A=P^{-1}Q^{-1} = LU$, т.е. $L =P^{-1}$ и $U = Q^{-1}$ — соответственно нижняя и верхняя треугольные матрицы.

При $P = Q^{-1}$ имеем *преобразование подобия* $\tilde{A} = Q^{-1}AQ$, где $\tilde{A}$ и A называют *подобными матрицами.* Приведение матрицы A к диагональной форме представляет собой случай этого преобразования $A=H^{-1}AH,$ причем $Q=H$ является модальной матрицей.

При $P=Q^{t}$ имеем *конгруэнтное преобразование* $\tilde{A}= Q^{t}AQ$, что, как и преобразование подобия, является частным случаем эквивалентного преобразования. Если преобразование удовлетворяет одновременно свойствам конгруэнтности и подобия, то оно называется *ортогональным*. При этом $\tilde{A} = Q^{-1}AQ = Q^{t}AQ$, где $Q=(Q^{-1})^{t}$, называется *ортогональной матрицей.*

Матричные преобразования тесно связаны с различными типами линейных преобразований в n-мерном евклидовом пространстве, которые рассматриваются ниже.

3.28. Эквивалентные преобразования многочленной матрицы.

Рассмотрим линейное преобразование $y=Ax$ в n-мерном пространстве. Умножив это уравнение слева на неособенную матрицу n-го порядка P, получим $Py=PAx$ или $y'=PAx$, что соответствует преобразованию вектора y в вектор $y' = Py$. Если вместо вектора x рассматривать вектор x', связанный с x неособенным преобразованием $x = Qx'$, то исходное уравнение приводится к виду $y' = PAQx'$ или $y' = \tilde{A}x'$. Матрица $\tilde{A}$ связана с A *эквивалентным преобразованием*

$$\tilde{A} = PAQ.$$

Строки произведения PA можно представить как сумму произведений строк матрицы A на элементы матрицы P, т.е.

$$PA = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{21} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix} \begin{bmatrix} a_{(1)} \\ a_{(2)} \\ \dots \\ a_{(n)} \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^{n} a_{(i)} p_{1i} \\ \sum\limits_{i=1}^{n} a_{(i)} p_{2i} \\ \dots \\ \sum\limits_{i=1}^{n} a_{(i)} p_{ni} \end{bmatrix}.$$

Аналогично, столбцы произведения AQ можно представить как сумму произведений столбцов матрицы A на элементы матрицы Q, т.е.

$$AQ = [a^{(1)},\ a^{(2)},\ \dots,\ a^{(n)}] \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \dots & \dots & \dots & \dots \\ q_{n1} & q_{n2} & \dots & q_{nn} \end{bmatrix} =$$

$$= \left[\sum_{l=1}^{n} a^{(l)} q_{l1},\ \sum_{l=1}^{n} a^{(l)} q_{l2},\ \dots,\ \sum_{l=1}^{n} a^{(l)} q_{ln} \right].$$

Определение 1. *Многочленной матрицей* или *λ-матрицей* называется прямоугольная матрица A (λ), элементы которой суть многочлены от λ:

$$A(\lambda) = \| a_{ik}(\lambda) \| = \| a_{ik}^{(0)} \lambda^l + a_{ik}^{(1)} \lambda^{l-1} + \dots + a_{ik}^{(l)} \| \quad \begin{pmatrix} i = 1, 2, \dots, m; \\ k = 1, 2, \dots, n \end{pmatrix};$$

здесь l — наибольший из степеней многочленов a_{ik} (λ).
Полагая

$$A_j = \| a_{ik}^{(j)} \| \quad (i = 1, 2, \dots, m;\ k = 1, 2, \dots, n;\ j = 0, 1, \dots, l),$$

мы можем представить многочленную матрицу A(λ) в виде матричного многочлена относительно λ, т.е. в виде многочлена с матричными коэффициентами:

$$A(\lambda) = A_0 \lambda^l + A_1 \lambda^{l-1} + \dots + A_{l-1} \lambda + A_l.$$

Введем в рассмотрение следующие *элементарные операции* над многочленной матрицей A (λ):

1° Умножение какой-либо, например i-й, строки на число $c \neq 0$.

2° Прибавление к какой-либо, например i-й, строке другой, например j-й, строки, предварительно умноженного на произвольный многочлен b (λ).

3° Перестановка местами любых двух строк, например i-й и j-й строк.

Предлагаем читателю проверить, что операции 1°, 2°, 3° равносильны умножению многочленной матрицы A (λ) *слева* соответственно на следующие квадратные матрицы порядка m:

$$
\left.
\begin{array}{l}
S' = \begin{array}{c} \quad (i) \\ \begin{Vmatrix} 1 & \cdots & \vdots & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ \vdots & & c & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{Vmatrix} \end{array}, \quad
S'' = \begin{array}{cc} \quad (j) & \\ \begin{Vmatrix} 1 & \cdots & \cdots & \vdots & \cdots & 0 \\ \vdots & \ddots & & \vdots & & \vdots \\ \vdots & & 1 & \cdots\, b(\lambda) & \cdots & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{Vmatrix} & \begin{array}{c} \\ \\ \cdots (i) \\ \\ \\ \\ \end{array} \end{array} \\
\\
S''' = \begin{array}{c} \quad (i) \quad\quad (j) \\ \begin{Vmatrix} 1 & & \vdots & & \vdots & & 0 \\ & \ddots & \vdots & & \vdots & & \\ \cdots & \cdots & 0 & \cdots & 1 & \cdots & \cdots \\ & & \vdots & \ddots & \vdots & & \\ \cdots & \cdots & 1 & \cdots & 0 & \cdots & \cdots \\ & & \vdots & & \vdots & \ddots & \\ 0 & & \vdots & & \vdots & & 1 \end{Vmatrix} \end{array},
\end{array}
\right\}
\qquad (3.330)
$$

(в матрицах (3.330) все неотмеченные элементы на главной диагонали равны единицы, а в остальнiх местах — нулю), т.е. в результате применения операций 1°, 2°, 3° матрица $A(\lambda)$ преобразуется соответственно в матрицi $S'\bullet A(\lambda)$, $S''\bullet A(\lambda)$, $S'''\bullet A(\lambda)$. Поэтому операции типа 1°, 2°, 3° называются *левыми элементарными операциями.*

Совершенно аналогично определяются *правые элементарные операции* над многочленной матрицей (эти операции выполняются не над строками, а над столбцами многочленной матрицы) и соответствующие им матрицы (порядка *n)*:

$$T' = \begin{array}{c} \begin{Vmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & c & \cdots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{Vmatrix} \end{array} \begin{array}{l} \\ \\ \ldots (i), \\ \\ \\ \end{array} \quad T'' = \begin{Vmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & 1 & \cdots & \cdots & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & b(\lambda) & \cdots & \cdots & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{Vmatrix} \begin{array}{l} \\ \\ \ldots (i) \\ \\ \ldots (j) \\ \\ \\ \end{array}$$

$$T''' = \begin{Vmatrix} 1 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & \vdots \\ \vdots & 0 & \cdots & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \vdots & 1 & \cdots & 0 & \cdots & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{Vmatrix} \begin{array}{l} \\ \\ \ldots (i) \\ \\ \ldots (j) \\ \\ \\ \end{array}$$

В результате применения правой элементарной операции матрица $A(\lambda)$ умножается *справа* на соответствующую матрицу T.

Матрицы типа S', S'', S''' (или, что то же самое, типа T', T'', T''') мы будем называть *элементарными матрицами.*

Определитель любой элементарной матрицы не зависит от λ и отличен от нуля. Поэтому для каждой левой (правой) элементарной операции существует обратная операция, которая также является левой (соответственно правой) элементарной операцией. (Отсюда следует, что если матрица $B(\lambda)$ получается из $A(\lambda)$ путем применения левых (правых; левых и правых) элементарных операций, то и обратная матрица $A(\lambda)$ может быть получена из $B(\lambda)$ путем применения элементарных операций такого же типа. Как левые элементарные операции, так и правые образуют группу).

Определение 2. Две многочленные матрицы $A(\lambda)$ и $B(\lambda)$ называются 1) *левоэквивалентными,* 2) *правоэквивалентными,* 3) *эквивалентными,* если одна из них получается из другой путем применения соответственно 1) левых элементарных операций, 2) правых элементарных операций, 3) левых и правых элементарных операций (из определения следует, что левоэквивалентными, правоэквивалентными или просто эквивалентными могут быть только прямоугольные матрицы одинаковых размеров).

Пусть матрица $B(\lambda)$ получается из $A(\lambda)$ при помощи левых элементарных операций, которые соответствую матрицам $S_1, S_2, ..., S_p$. Тогда

$$B(\lambda) = S_pS_{p-1}...S_1A(\lambda). \qquad (3.331)$$

Обозначая через $P(\lambda)$ произведение $S_pS_{p-1} ... S_1$, мы равенство (3.331) запишем в виде

$$B(\lambda)=P(\lambda)A(\lambda), \qquad (3.332)$$

где $P(\lambda)$, как и каждая из матриц $S_1, S_2, ..., S_p$, имеет отличный от нуля *постоянный* (т.е. не зависимый от λ) определитель.

В следующем разделе будет доказано, что каждая квадратная λ-матрица $P(\lambda)$ с постоянным отличным от нуля определителем может быть представлена в виде произведения элементарных матриц. Поэтому равенство (3.332) эквивалентно равенству (3.331) и потому означает левую эквивалентность матриц $A(\lambda)$ и $B(\lambda)$.

В случае правой эквивалентности многочленных матриц $A(\lambda)$ и $B(\lambda)$ вместо равенства (3.332) будем иметь равенство

$$B(\lambda)=A(\lambda)Q(\lambda), \qquad (3.333)$$

а в случае (двусторонней) эквивалентности - равенство

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda); \qquad (3.334)$$

здесь опять $P(\lambda)$ и $Q(\lambda)$ — матрицы с отличными от нуля и не зависимыми от λ определителями.

Таким образом, определение 2 можно заменить равносильным определениям.

Определение 3. Две прямоугольные λ-матрицы $A(\lambda)$ и $B(\lambda)$ называются 1) *левоэквивалентными,* 2) *правоэквивалентными,* 3) *эквивалентными,* если соответственно

1) $B(\lambda) = P(\lambda) A(\lambda)$, 2) $B(\lambda) = A(\lambda) Q(\lambda)$, 3) $B(\lambda) = P(\lambda) A(\lambda) Q(\lambda)$,

где $P(\lambda)$ и $Q(\lambda)$ - многочленные квадратные матрицы с постоянными и отличными от нуля определителями.

Все введенные высшее понятия проиллюстрируем на следующем примере.

Рассмотрим систему m линейных однородных дифференциальных уравнений l-го порядка с n неизвестными функциями $x_1, x_2, ..., x_n$ аргумента t с постоянными коэффициентами:

$$\left.\begin{array}{l} a_{11}(D)\,x_1 + a_{12}(D)\,x_2 + \ldots + a_{1n}(D)\,x_n = 0, \\ a_{21}(D)\,x_1 + a_{22}(D)\,x_2 + \ldots + a_{2n}(D)\,x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}(D)\,x_1 + a_{m2}(D)\,x_2 + \ldots + a_{mn}(D)\,x_n = 0; \end{array}\right\}$$

(3.335)

здесь

$$a_{ik}(D) = a_{ik}^{(0)}D^l + a_{ik}^{(1)}D^{l-1} + \ldots + a_{ik}^{(l)} \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

- многочлен относительно D с постоянными коэффициентами, $D=d/dt$ - оператор дифференцирования.

Матрица операторних коэффициентов

$$A(D) = \|\, a_{ik}(D) \,\| \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

является многочленной матрицей или D -матрицей.

Очевидно, что левая элементарная операция 1° над матрицей $A(D)$ означает почленное умножение i-го дифференциального уравнения системы на число $c\neq 0$. Левая элементарная операция 2° означает почлене прибавление к i-му уравнение j-го уравнения, предварительно подвергнутого дифференциальной операции $b(D)$. Левая элементарная операция 3° означает перестановку местами i-го и j-го уравнений.

Таким образом, если в системе уравнений (3.335) матрицу операторних коэффициентов $A(p)$ заменить левоэквивалентной ей матрицей $B(D)$, то мы получим новую систему уравнений. Поскольку, обратно, исходная система может быть получена из этой системы при помощи аналогичных операций, то обе системы уравнений равносильны. (При этом принимается, что искомые функции $x_1, x_2, \ldots, x_n$ таковы, что производные от этих функций всех порядков, которые встречаются при преобразованиях, существуют. Тогда две системы уравнений с левоэквивалентными матрицами $A(D)$ и $B(D)$ имеют одни и те же решения).

Нетрудно проинтерпретировать на данном примере и правые элементарные операции. Первая из этих операций означает введение вместо одной из неизвестных функций x_i - новой неизвестной функции $x'_i=(1/c)x_i$; вторая элементарная операция означает введение новой неизвестной функции $x'_j = x_j — b(D)x_s$ (вместо x_j); третья операция означает перемену местами в уравнениях членов, которые содержат x_i и x_j (т.е. $x_i = x'$, $x_j = x'_i$).

3.29. Канонический вид λ-матрицы

1. Выясним сначала, к какому сравнительно простому виду можно привести прямоугольную многочленную матрицу $A(\lambda)$ путем применения одних только левых элементарных операций.

Допустим, что в первом столбце матрицы $A(\lambda)$ имеются элементы, не равные тождественно нулю. Возьмем среди них многочлен наименьшей степени и путем перестановки строк сделаем его элементом $a_{11}(\lambda)$. После этого разделим многочлен $a_{i1}(\lambda)$ на $a_{11}(\lambda)$; частное и остаток обозначим через $q_{i1}(\lambda)$ и $r_{i1}(\lambda)$

$$a_{i1}(\lambda) = a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda) \quad (i = 2, ..., m).$$

Вычтем теперь из i-й строки первую строку, предварительно умноженную на $q_{i1}(\lambda)$ ($i = 2, ..., m$). Если при этом не все остатки $r_{i1}(\lambda)$ равны тождественно нулю, то тот из них, который не равен нулю и имеет наименьшую степень, может быть перестановкой строк поставлен на место $a_{11}(\lambda)$. В результате всех этих операций степень многочлена $a_{11}(\lambda)$ понизится.

Теперь мы снова повторим этот процесс и т.д. Так как степень многочлена $a_{11}(\lambda)$ конечна, то на некотором этапе этот процесс уже нельзя будет продолжить, т.е. на этом этапе все элементы $a_{21}(\lambda), a_{31}(\lambda), ..., a_{m1}(\lambda)$ окажутся равными тождественно нулю.

После этого возьмем элемент $a_{22}(\lambda)$ и применим ту же процедуру к строкам с номерами 2, 3, ..., m. Тогда добьемся того, что и $a_{32}(\lambda) = ... = a_{m2}(\lambda) = 0$. Продолжая так далее, мы в конце концов приведем матрицу $A(\lambda)$ к следующему виду:

$$\left\|\begin{matrix} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1m}(\lambda) & \dots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2m}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{mm}(\lambda) & \dots & b_{mn}(\lambda) \\ & & (m \leqslant n) & & & \end{matrix}\right\|, \quad \left\|\begin{matrix} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn}(\lambda) \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \\ & (m \geqslant n) & & \end{matrix}\right\|. \tag{3.336}$$

Если многочлен $b_{22}(\lambda)$ не равен тождественно нулю, то, применяя левую элементарную операцию второго типа, мы сделаем степень элемента $b_{12}(\lambda)$ меньшей, чем степень $b_{22}(\lambda)$ (если $b_{22}(\lambda)$ имеет нулевую степень, то $b_{12}(\lambda)$ станет тождественно равен нулю). Точно так же, если

$a_{33}(\lambda) \not\equiv 0$, то с помощью левых элементарных операций второго типа мы сделаем степени элементов $b_{13}(\lambda)$, $b_{23}(\lambda)$ меньшими, чем степень $b_{33}(\lambda)$, не изменив при этом элемента $b_{12}(\lambda)$, и т.д.

Сформулируем следующую теорему:

Теорема 1. *Произвольная прямоугольная многочленная матрица с размерами $m \times n$ при помощи левых элементарных операций всегда может быть приведена к виду* (3.336), *где многочлени $b_{1k}(\lambda)$, $b_{2k}(\lambda)$, ..., $b_{k-1,k}(\lambda)$ имеют меньшую степень, чем $b_{kk}(\lambda)$, если только $b_{kk}(\lambda) \not\equiv 0$, и все равны тождественно нулю, если $b_{kk}(\lambda) = \text{const} \neq 0$ ($k = 2, 3, ..., \min(m, n)$).*

Совсем аналогично доказывается следующая теорема.

Теорема 2. *Произвольная прямоугольная многочленная матрица с размерами $m \times n$ при помощи правых элементарных операций всегда может быть приведена к виду*

$$\left\| \begin{matrix} c_{11}(\lambda) & 0 & \dots & 0 & 0 \dots 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \dots & 0 & 0 \dots 0 \\ \dots & \dots & \dots & \dots & \dots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \dots & c_{mm}(\lambda) & 0 \dots 0 \\ & & (m \leqslant n) & & \end{matrix} \right\|, \quad \left\| \begin{matrix} c_{11}(\lambda) & 0 & \dots & 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ c_{n1}(\lambda) & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \\ \dots & \dots & \dots & \dots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \dots & c_{mn}(\lambda) \\ & (m \geqslant n) & & \end{matrix} \right\|, \tag{3.337}$$

где многочлены $c_{k1}(\lambda)$, $c_{k2}(\lambda)$, ..., $c_{k-1,k}(\lambda)$ имеют меньшую степень, чем $c_{kk}(\lambda)$, если только $c_{kk}(k) \not\equiv 0$, и все равны тождественно нулю, если $c_{kk}(\lambda) = \text{const} \neq 0$ ($k = 2, 3, ..., \min(m, n)$).

2. Из теорем 1 и 2 вытекает следующее

Следствие. *Если определитель квадратной многочленной матрицы $P(\lambda)$ не зависит от λ и отличен от нуля, то эту матрицу можно представить в виде произведения конечного числа элементарных матриц.*

Действительно, согласно теореме 1 матрицу $P(\lambda)$ при помощи левых элементарных операций можно привести к виду

$$\left\| \begin{matrix} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn}(\lambda) \end{matrix} \right\|, \tag{3.338}$$

где n — порядок матрицы $P(\lambda)$. Так как при применении элементарных операций к квадратной многочленной матрице определитель этой матрицы умножается лишь на постоянный отличный от нуля множитель, то определитель матрицы (3.338), как и определитель $P(\lambda)$, не зависит от λ и отличен от нуля, т.е.

$$b_{11}(\lambda)\, b_{22}(\lambda) \dots b_{nn}(\lambda) = \text{const} \neq 0.$$

Отсюда

$$b_{kk}(\lambda) = \text{const} \neq 0 \quad (k = 1, 2, \dots, n).$$

Но тогда в силу той же теоремы 1 матрица (3.338) имеет диагональный вид $\| \delta_{kk}\ \delta_{ik} \|_1^n$ и потому может быть приведена с помощью левых элементарных операций типа 1° к единичной матрице E. Тогда и обратно, единичную матрицу E можно привести к $P(\lambda)$ при помощи левых элементарных операций с матрицами $S_1, S_2, \dots, S_p$. Следовательно,

$$P(\lambda) = S_p S_{p-1} \dots S_1 E = S_p S_{p-1} \dots S_1.$$

Из доказанного следствия получаем равносильность двух определений 2 и 3 эквивалентности многочленных матриц.

3. Вернемся к нашему примеру системы дифференциальных уравнений (3.335). Применим теорему 1 к матрице операторных коэффициентов $\| a_{ik}(D) \|$. Тогда, как было указано раньше, система (3.335) заменится равносильной системой

$$\left.\begin{aligned}
b_{11}(D)\,x_1 + b_{12}(D)\,x_2 + \dots + b_{1s}(D)\,x_s &= \\
&= -b_{1,s+1}(D)\,x_{s+1} - \dots - b_{1n}(D)\,x_n, \\
b_{22}(D)\,x_2 + \dots + b_{2s}(D)\,x_s &= \\
&= -b_{2,s+1}(D)\,x_{s+1} - \dots - b_{2n}(D)\,x_n, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots& \\
b_{ss}(D)\,x_s = -b_{s,s+1}(D)\,x_{s+1} &- \dots - b_{sn}(D)\,x_n,
\end{aligned}\right\} \tag{3.339}$$

где $s = \min(m, n)$. В этой системе мы функции $x_{s+1}, \dots, x_n$ можем выбрать произвольно, после чего последовательно определятся функции $x_s, x_{s-1}, \dots, x_1$, причем на каждом этапе этого определения приходится интегрировать одно дифференциальное уравнение с одной неизвестной функцией.

4. Перейдем теперь к установлению «канонического» вида, к которому можно привести прямоугольную многочленную матрицу $A(\lambda)$, применяя к ней как левые, так и правые элементарные операции.

Среди всех не равных тождественно нулю элементов $a_{ik}(\lambda)$ матрицы $A(\lambda)$ возьмем тот элемент, который имеет наименьшая степень относительно λ, и путем соответствующей перестановки строк и столбцов сделаем его элементом $a_{11}(\lambda)$. После этого найдем частные и остатки от деления многочленов $a_{i1}(\lambda)$ и $a_{1k}(\lambda)$ на $a_{11}(\lambda)$:

$$a_{i1}(\lambda) = a_{11}(\lambda)\, q_{i1}(\lambda) + r_{i1}(\lambda), \qquad a_{1k}(\lambda) = a_{11}(\lambda)\, q_{1k}(\lambda) + r_{1k}(\lambda)$$
$$(i = 2, 3, \ldots, m; \qquad k = 2, 3, \ldots, n).$$

Если хотя бы один из остатков $r_{i1}(\lambda)$, $r_{1k}(\lambda)$ (r = 2, ..., m; k = 2, ..., n), например $r_{1k}(\lambda)$, не равен тождественно нулю, то, вычитая из k-го столбца первый столбец, предварительно умноженный на $q_{1k}(\lambda)$, мы заменим элемент $a_{1k}(\lambda)$ остатком $r_{1k}(\lambda)$, который имеет меньшую степень, чем $a_{11}(\lambda)$. Тогда мы имеем возможность снова уменьшить степень элемента, который стоит в левом верхнем углу матрицы, поместив на это место элемент с наименьшей степенью относительно λ.

Если же все остатки $r_{21}(\lambda)$, ..., $r_{m1}(\lambda)$; $r_{12}(\lambda)$,…,$r_{1г}(\lambda)$ равны тождественно нулю, то, вычитая из i-й строки первый, умноженный предварительно на $q_{ik}(\lambda)$ (i= 2, ..., m), а из k -го столбца — первый, предварительно умноженный на $q_{ik}(\lambda)$ (k = 2, ..., n), мы приведем нашу многочленную матрицу к виду

$$\begin{Vmatrix} a_{11}(\lambda) & 0 & \ldots & 0 \\ 0 & a_{22}(\lambda) & \ldots & a_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{m2}(\lambda) & \ldots & a_{mn}(\lambda) \end{Vmatrix}.$$

Если при этом хотя бы один из элементов $a_{ik}(\lambda)$ (i = 2, ..., m; k = 2,...,n) не делится без остатка на $a_{11}(\lambda)$, то, прибавляя к первому столбцу тот столбец, который содержит этот элемент, мы придем к предыдущему случаю и, следовательно, снова сможем заменить элемент $a_{11}(\lambda)$ многочленом меньшей степени.

Поскольку первоначальный элемент $a_{11}(\lambda)$ имел определенную степень и процесс уменьшения этой степени не может неограниченно продолжаться, то после конечного числа элементарных операций мы должны получить матрицу вида

$$\begin{Vmatrix} a_{1}(\lambda) & 0 & \ldots & 0 \\ 0 & b_{22}(\lambda) & \ldots & b_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & b_{m2}(\lambda) & \ldots & b_{mn}(\lambda) \end{Vmatrix}, \tag{3.340}$$

в которой все элементы b_{ik} (λ) делятся без остатка на a_1 (λ). Если среди этих элементов b_{ik}(λ) имеются не равные тождественно нулю, то, продолжая тот же процесс приведения для строк с номерами 2, ..., m и столбцов с номерами 2, ..., n, мы матрицу (3.340) приведем к виду

$$\begin{Vmatrix} a_1(\lambda) & 0 & 0 & \dots & 0 \\ 0 & a_2(\lambda) & 0 & \dots & 0 \\ 0 & 0 & c_{33}(\lambda) & \dots & c_{3n}(\lambda) \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & c_{m3}(\lambda) & \dots & c_{mn}(\lambda) \end{Vmatrix},$$

где a_2 (λ) делится без остатка на a_1(λ), а все многочлены c_{ik}(λ) делятся без остатка на a_2(λ). Продолжая этот процесс далее, мы в конце концов придем к матрице вида

$$\begin{Vmatrix} a_1(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_2(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_s(\lambda) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{Vmatrix}, \tag{3.341}$$

где многочлены a_1(λ), a_2 (λ), ..., a_s (λ) ($s \leq m, n$) не равны тождественно нулю и каждый из них делится без остатка на предыдущий.

Умножая первые s строк на соответствующие отличные от нуля числовые множители, мы сможем добиться того, чтобы старшие коэффициенты многочленов a_1(λ), a_2(λ), ..., a_s(λ) были равны единице.

Определение 4. Многочленная прямоугольная матрица называется *канонической диагональной,* если она имеет вид (3.341), где 1) многочлены a_1(λ), a_2 (λ), ..., a_s(λ) не равны тождественно нулю и 2) каждый из многочленов a_2(λ), ..., a_s(λ) делится без остатка на предыдущий. При этом предполагается, что старшие коэффициенты всех многочленов a_1(λ), a_2 (λ), ..., a_s (λ) равны единице.

Таким образом, доказано, что *произвольная прямоугольная многочленная матрица A(λ) эквивалентна некоторой канонической диагональной*. Далее будет показано, что *многочлены a_1(λ),a_2(λ),...,a_s (λ) однозначно определяются заданием матрицы A(λ)*, и установим формулы, которые связывают эти многочлены с элементами матрицы A(λ).

3.30. Преобразование подобия.

Эквивалентное преобразование можно рассматривать как результат перехода к новым (вообще различным) координатным базисам для векторов x и y, т.е. $x'=Q^{-1}x$ и $y'=Py$. Иначе говоря, преобразование $\tilde{A}=PAQ$ соответствует независимым преобразованиям координат, определяемым матрицами Q^{-1} и P.

Если векторы x и y преобразуются к одному и тому же координатному базиса, то приняв $P = Q^{-1}$, приходим к преобразованию подобия:

$$\tilde{A}= Q^{-1}AQ.$$

Она означает, что уравнение $y=Ax$ при переходе к новому координатному базису, определяемому матрицей Q^{-1} ($x' = Q^{-1}x$ и $y'=Q^{-1}y$), преобразуется к $y'=\tilde{A}x'$. Важнейшее свойство преобразования подобия состоит в том, что определитель матрицы инвариантен относительно этого преобразования:

$$\det \tilde{A} = \det Q^{-1} \det A \det Q = \det A,$$

так как определитель обратной матрицы $\det Q^{-1}$ равен обратному значению определителя $\det Q$. Ясно, что и собственные значения матрицы не изменяются при преобразовании подобия. Действительно,

$$\lambda E - \tilde{A} = \lambda E - Q^{-1}AQ = Q^{-1}[\lambda Q\, Q^{-1} - A]Q = Q^{-1}\,[\lambda E - A]Q.$$

Откуда следует

$$\det[\lambda E - \tilde{A}] = \det[\lambda E - A].$$

При определенных условиях преобразование подобия приводит матрицу A к диагональной

$$\Lambda = H^{-1}AH = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения матрицы A, а роль преобразующей матрицы Q играет модальная матрица H.

Можно указать, по крайней мере, три случая, когда матрица приводится к диагональной форме:

1) все собственные значения матрицы A различны;
2) дефекты матриц $[\lambda_k E - A]$ равны кратностям m_k (ранги равны

$n - m_k$) соответствующих собственных значений λ_k *(кратная вырожденность);*

3) симметричные матрицы.

Первый из них рассмотрен раньше, второй будет исследован ниже, а третий связан с ортогональным преобразованием. В общем случае преобразование подобия приводит к *квазидиагональной канонической форме.*

Матрица простой структуры. В случае кратной вырожденности говорят, что матрица имеет *простую структуру.* При этом каждое из уравнений $(\lambda_k E - A)h=0$ дает m_k независимых решений для собственных векторов h, соответствующих m_k-кратному корню λ_k. Таким образом, получим всего $m_1 + m_2 + \ldots + m_q=n$ независимых векторов, которые и составят модальную матрицу H. Если при определении модальных столбцов исходить из присоединенной матрицы $G(\lambda) = \mathrm{Adj}[\lambda E - A]$, то следует иметь в виду, что ее значение и значение всех производных до $(m_k - 2)$-й включительно при $\lambda=\lambda_k$ — нулевые матрицы. Поэтому m_k независимых модальных столбцов для λ_k выбираем из $(m_k - 1)$-х производных

$$\left[\frac{d^{m_k-1}}{d\lambda^{m_k-1}} G(\lambda)\right]_{\lambda=\lambda_k} \qquad (k = 1,\ 2,\ \ldots,\ q).$$

В результате получаем $m_1+m_2+\ldots+ m_q=n$ модальных столбцов матрицы H, которая преобразует матрицу A к диагональной форме, причем каждое из собственных значений кратности m_k повторяется на главной диагонали m_k раз. Например:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}; \qquad [\lambda E - A] = \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -1 & \lambda - 2 & -1 \\ 0 & 0 & \lambda - 1 \end{bmatrix} = F(\lambda).$$

Определив $\Delta(\lambda)$ и $G(\lambda)$ с помощью алгоритма Фаддеева, получим:

$$\Delta(\lambda) = \lambda^3 - 5\lambda^2 + 7\lambda - 3 = (\lambda - 1)^2(\lambda - 3);$$

$$G(\lambda) = \begin{bmatrix} (\lambda - 2)(\lambda - 1) & \lambda - 1 & \lambda - 1 \\ \lambda - 1 & (\lambda - 2)(\lambda - 1) & \lambda - 1 \\ 0 & 0 & (\lambda - 1)(\lambda - 3) \end{bmatrix}.$$

Так как дефект $F(\lambda_1)$ равен двум и дефект $F(\lambda_2)$ — единице, то матрица A имеет простую структуру и приводится к диагональной форме. Модальные столбцы, соответствующие двукратному корню $\lambda_1=1$, выбираем из независимых столбцов матрицы $G'(\lambda_1)$:

$$G'(\lambda) = \begin{bmatrix} 2\lambda - 3 & 1 & 1 \\ 1 & 2\lambda - 3 & 1 \\ 0 & 0 & 2\lambda - 4 \end{bmatrix}; \quad G'(1) = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix}.$$

Для простого корня λ_2=3 в качестве модального столбца выбираем пропорциональный одному из столбцов матрицы:

$$G(3) = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}; \qquad h^{(3)} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Таким образом имеем:

$$H = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & -2 & 0 \end{bmatrix}; \quad H^{-1} = \frac{1}{2}\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix};$$

$$\Lambda = \frac{1}{2}\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & -2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Каноническая форма Жордана. В общем случае матрица A может не иметь простой структуры. Это значит, что хотя бы для одного собственного значения λ дефект матрицы $\lambda_k E - A$ отличается от m_k — кратности λ_k. Тогда преобразование подобия приводит к *канонической матрице Жордана* $J = H^{-1}AH$, для которой характерна квазидиагональная структура.

Главную диагональ матрицы J, как и в случае матрицы простой структуры, занимают собственные значения λ_i $(i = 1, 2, ..., q)$, причем каждое из них представлено m_i раз. Над главной диагональю (по первой наддиагонали) располагаются единичные элементы, но они не обязательно занимают все клетки наддиагонали.

Расположение единичных элементов зависит от структуры матрицы A. Остальные элементы матрицы Жордана равны нулю.

С каждым собственным значением λ_k связанны одна или несколько *клеток Жордана r-го порядка* $(r \le m_k)$, имеющих следующую структуру:

$$J_r(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & & & \\ & \lambda_k & 1 & & & \\ & & \cdots & \cdots & & \\ & & & & \lambda_k & 1 \\ & & & & & \lambda_k \end{bmatrix}.$$

Все такие клетки являются блоками квазидиагональной матрицы J и сумма их порядков равна порядку n исходной матрицы A. Количество клеток Жордана, соответствующих кратному собственному значению λ_k, равно дефекту матрицы $\lambda_k E$ — A, а сумма их порядков равна кратности m_k. Это позволяет выяснить структуру матрицы Жордана лишь в следующих случаях.

1) При кратной вырожденности, когда дефект матрицы $\lambda_k E$ — A для m_k -кратного λ_k равен m_k, имеем m_k клеток первого порядка, т.е. справа от диагональных элементов λ_k единицы отсутствуют. Этот случай был рассмотрен раньше.

2) При простой вырожденности, когда для m_k -кратного λ_k дефект матрицы $\lambda_k E$ — A равен единице, собственному значению λ_k соответствует только одна клетка m_k -го порядка. Это значит, что везде справа от λ_k в клетке Жордана стоят единицы. При этом матрицы

$$G^{(j-1)}(\lambda_k) = \left[\frac{d^{j-1}}{d\lambda^{j-1}} \operatorname{Adj}(\lambda E - A)\right]_{\lambda=\lambda_k} \qquad (j = 1,\ 2,\ \dots\ ,\ m_k)$$

содержат в совокупности m_k (и только m_k) независимых столбцов, которые (или пропорциональные им) можно принять в качестве m_k собственных векторов.

Рассмотрим, например, матрицу:

$$A = \begin{bmatrix} -2 & 2 & -3 \\ -10 & 4 & -5 \\ -5 & 4 & -6 \end{bmatrix}; \quad \lambda E - A = \begin{bmatrix} \lambda + 2 & -2 & 3 \\ 10 & \lambda - 4 & 5 \\ 5 & -4 & \lambda + 6 \end{bmatrix} = F(\lambda);$$

$$\Delta(\lambda) = (\lambda + 1)^2 (\lambda + 2); \quad \lambda_1 = -1 \quad (m_1 = 2); \quad \lambda_2 = -2 \quad (m_2 = 1).$$

Здесь есть простая вырожденность для двукратного корня $\lambda_1 = -1$, так как дефект матрицы $F(-1)$ равен единице, а ранг двум. Поэтому с точностью до порядка расположения клеток матрица Жордана имеет вид:

$$J=\left[\begin{array}{cc|c}-1 & 1 & 0\\ 0 & -1 & 0\\ \hline 0 & 0 & -2\end{array}\right]=\begin{bmatrix}J_2(\lambda_1) & \\ & J_1(\lambda_2)\end{bmatrix}.$$

Клетка $J_2(\lambda_2)$ второго порядка соответствует $\lambda_1=-1$, а $J_1(\lambda_2)$ — клетка первого порядка, совпадающая с элементом $\lambda_2=-2$.

Для вычисления модальной матрицы воспользуемся значениями присоединенных матриц и их производных, полученных раньше:

$$G(\lambda_1)=\begin{bmatrix}-5 & -2 & 5\\ -25 & -10 & 25\\ -15 & -6 & 15\end{bmatrix};\quad G'(\lambda_1)=\begin{bmatrix}0 & 2 & -3\\ -10 & 6 & -5\\ -5 & 4 & -4\end{bmatrix};$$

$$G(\lambda_2)=\begin{bmatrix}-4 & -4 & 8\\ -15 & -15 & 30\\ -10 & -10 & 20\end{bmatrix}.$$

Хотя ранг $G(\lambda_1)$ равен единице, а $G'(\lambda_1)$ — двум, но в совокупности они имеют только два независимых столбца, пропорциональные которым принимаем как собственные векторы для двукратного собственного значения λ_1. Третий вектор принимаем пропорциональным столбцу матрицы $G(\lambda_2)$, ранг которой равен единице. В результате получаем модальную матрицу:

$$H=\begin{bmatrix}1 & 0 & 4\\ 5 & 2 & 15\\ 3 & 1 & 10\end{bmatrix};\qquad H^{-1}=\begin{bmatrix}5 & 4 & -8\\ -5 & -2 & 5\\ -1 & -1 & 2\end{bmatrix}.$$

Проверка по формуле $J=H^{-1}AH$ приводит к приведенной выше матрице Жордана.

3) Если для λ_k кратности m_k дефект d матрицы λ_kE-A равен m_k-1, то этому собственному значению соответствует m_k-1 клеток Жордана, общий порядок которых равен m_k. Поэтому среди них может быть только одна клетка второго порядка, а остальные m_k-2 клеток будут первого порядка. С точностью до расположения этих клеток блок для λ_k имеет при $d=m_k-1$ следующую диагональную структуру:

$$J(\lambda_k)=\begin{bmatrix} J_2(\lambda_k) & & & \\ & J_1(\lambda_k) & & \\ & & \ddots & \\ & & & (J_1\lambda_k) \end{bmatrix}=\begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & & \\ & & \ddots & \\ & & & \lambda_k \end{bmatrix}.$$

$$\underbrace{\qquad\qquad}_{m_k-2} \qquad\qquad \underbrace{\qquad\qquad}_{m_k-2}$$

Например:

$$A=\begin{bmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{bmatrix}; \quad F(\lambda)=\begin{bmatrix} \lambda-3 & -2 & 3 \\ -4 & \lambda-10 & 12 \\ -3 & -6 & \lambda+7 \end{bmatrix};$$

$$\Delta(\lambda)=\lambda^3-6\lambda^2+12\lambda-8=(\lambda-2)^3.$$

Так как при $\lambda = 2$ ранг матрицы $F(2)$ равен единице, т.е. ее дефект равен двум, то каноническая матрица Жордана имеет вид:

$$J=\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Матрица Жордана в общем случае. Если дефект d_k матрицы $\lambda_k E$ — A для m_k-кратного λ_k больше единицы и меньше m_k, то определение соответствующих клеток Жордана встречает существенные трудности. Априори известно лишь количество d_k и суммарный порядок m_k клеток, соответствующих собственному значению λ_k, но выяснение порядка каждой из них требует дополнительного и довольно сложного исследования в каждом конкретном случае.

Если структура матрицы Жордана известна, то для вычисления модальной матрицы H можно воспользоваться уравнением $AH=HJ$, которое получается из преобразование подобия $J=H^{-1}AH$ умножением слева обеих частей равенства на H. Для клетки Жордана r-го порядка, соответствующей λ_k, из уравнения $AH=HJ$ имеем уравнение относительно r линейно-независимых собственных векторов $h^{(1)}, h^{(2)},....., h^{(r)}$:

$$A[h^{(1)}, h^{(2)}, \dots, h^{(r)}]=[h^{(1)}, h^{(2)}, \dots, h^{(r)}]\,J_r(\lambda_k)=$$

$$=[h^{(1)}, h^{(2)}, \dots, h^{(r)}]\begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & 1 & \\ & & \dots & \dots \\ & & & \lambda_k \end{bmatrix}.$$

Отсюда получаем систему r уравнений

$$Ah^{(1)} = \lambda_k h^{(1)}; \quad Ah^{(2)} = h^{(1)} + \lambda_k h^{(2)}; \ \dots \ ; \ Ah^{(r)} = h^{(r-1)} + \lambda_k h^{(r)},$$

или

$$(\lambda_k E - A)\, h^{(1)} = 0; \quad (\lambda_k E - A) h^{(2)} = -h^{(1)}; \ \dots \ ;$$
$$(\lambda_k E - A) h^{(r)} = -h^{(r-1)}.$$

Решая эти уравнения для каждой клетки Жордана, определяем соответствующие столбцы матрицы H. Вектор $h^{(1)}$ можно выбирать пропорциональным любому столбцу матрицы $G(\lambda_k) = \mathrm{Adj}\,[\lambda_k E - A]$.

При неизвестной структуре матрицы Жордана приведенные соотношения можно использовать для всевозможных вариантов расположения ее клеток. Если система уравнений для данного варианта совместна, то это свидетельствует о его соответствии истинной структуре матрицы Жордана, подобной данной матрице A.

Совокупность линейно-независимых собственных векторов $h^{(1)}$, $h^{(2)}$,....., $h^{(r)}$ соответствующих клетке $J_r(\lambda_k)$, образует *жорданову цепочка векторов* в r-мерном векторном пространстве, которое является подпространством n-мерного пространства. Совокупность всех жордановых цепочек представляет *жордановый базис* в этом n-мерном пространстве.

Рассмотрим, например, матрицу:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 & -3 \\ 0 & 0 & -1 & 2 & -2 \\ 1 & -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 & 2 \end{bmatrix}.$$

Для нее

$$\lambda E - A = \begin{bmatrix} \lambda-1 & 0 & 0 & -1 & 1 \\ 0 & \lambda-1 & 2 & -3 & 3 \\ 0 & 0 & \lambda+1 & -2 & 2 \\ -1 & 1 & -1 & \lambda & -1 \\ -1 & 1 & -1 & 1 & \lambda-2 \end{bmatrix};$$

$$\Delta(\lambda) = (\lambda-1)^4(\lambda+1).$$

Подставляя значение четырехкратного корня $\lambda=1$ в матрицу $\lambda E - A$, убеждаемся, что дефект этой матрицы равен двум. Следовательно, возможна одна из следующих двух форм матрицы Жордана:

$$J' = \begin{bmatrix} 1 & 1 & 0 & & \\ 0 & 1 & 1 & & \\ 0 & 0 & 1 & & \\ & & & 1 & \\ & & & & -1 \end{bmatrix}; \qquad J'' = \begin{bmatrix} 1 & 1 & & & \\ 0 & 1 & & & \\ & & 1 & 1 & \\ & & 0 & 1 & \\ & & & & -1 \end{bmatrix}.$$

Проверка с помощью уравнений для собственных векторов показывает, что правильной является вторая форма.

Функции от матрицы Жордана. Так как преобразование подобия $\tilde{A}=Q^{-1}AQ$ не изменяет собственных значений матрицы A, а матрица Q вещественная, то

$$f(\tilde{A}) = Q^{-1}f(A)\,Q; \quad f(A) = Qf(\tilde{A})\,Q^{-1}.$$

Это означает, что определение функции от произвольной матрицы A можно свести к определению функции от подобной ей $\tilde{A}$ и, определив матрицу преобразования Q, найти затем $f(A)$ по приведенной выше формуле. Ясно, что наибольший для практики интерес представляет случай, когда $\tilde{A}$ имеет стандартную простейшую форму. Одной из таких форм и есть каноническая матрица Жордана, которая для матрицы A простой структуры принимает диагональную форму. Последняя уже использовалась при определении экспоненциальной функции от матрицы с различными собственными значениями.

Итак, задача состоит в том, чтобы найти общее выражение для функции от матрицы Жордана $f(J)$. В силу квазидиагональности J можно получить $f(J)$ заменой каждой клетки J_r Жордана на матрицу $f(J_r)$ и свести задачу к определению функции от этой клетки.

Воспользуемся интерполяционным методом. Так как характеристический многочлен для $J_r(k_k)$ равен $(\lambda—\lambda_k)^r$, то интерполяционный многочлен имеет вид:

$$r(\lambda) = f(\lambda_k) + \frac{f'(\lambda_k)}{1!}(\lambda - \lambda_k) + \cdots + \frac{f^{(r-1)}(\lambda_k)}{(r-1)!}(\lambda - \lambda_k)^{r-1}.$$

Заменяя λ на матрицу J_r и учитывая, что $f(J_r)=r(J_r)$, получаем

$$f(J_r) = f(\lambda_k)\,E + \frac{f'(\lambda_k)}{1!}(J_r - \lambda_k E) + \cdots + \frac{f^{(r-1)}(\lambda_k)}{(r-1)!}(J_r - \lambda_k E)^{r-1}.$$

Учитывая структуру клетки J_r, легко убедиться, что $J_r —\lambda_k E$ представляет собой матрицу r-го порядка с единицами на первой наддиагонали, а остальные ее элементы равны нулю. При каждом умножении этой матрицы на себя единицы смещаются на следующую наддиагональ, так что матрица $(J_r— \lambda_k E)^j$ будет содержать единицы только на j-й наддиагонали. Очевидно, $(J_r - \lambda_k E)^{r-1}$ содержит единственный ненулевой элемент, равный единице, в правом верхнем углу. Таким образом, приходим к следующему виду функции от клетки J_r для собственного значения λ_k:

$$f(J_k) = \begin{bmatrix} f(\lambda_k) & \frac{f'(\lambda_k)}{1!} & \dots & \frac{f^{(r-2)}(\lambda_k)}{(r-2)!} & \frac{f^{(r-1)}(\lambda_k)}{(r-1)!} \\ & f(\lambda_k) & \dots & \frac{f^{(r-3)}(\lambda_k)}{(r-3)!} & \frac{f^{(r-2)}(\lambda_k)}{(r-2)!} \\ & & \dots & \dots & \dots \\ & & & f(\lambda_k) & \frac{f'(\lambda_k)}{1!} \\ & & & & f(\lambda_k) \end{bmatrix}.$$

Располагая такие блоки для всех собственных значений по диагонали, получаем квазидиагональную форму для функции $f(J)$ от матрицы Жордана. В частности, для диагональной матрицы

$$f(\Lambda) = \begin{bmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \dots & \\ & & & f(\lambda_n) \end{bmatrix},$$

где функции для кратных собственных значений повторяются соответственно их кратности.

Зная $f(J)$ и модальную матрицу H, можно по формуле $f(A)=Hf(J)H^{-1}$ определить функцию от матрицы A, которой соответствует матрица J.

Экспоненциальную функцию от клетки Жордана $\exp(J_k t)$ запишем с учетом соотношений $f(\lambda_k) = \exp(\lambda_k t)$;

$$f'(\lambda_k) = \frac{1}{1!}\exp(\lambda_k t), \ \dots \ , \ f^{(r-1)}(\lambda_k) = \frac{1}{(r-1)!}\exp(\lambda_k t)$$

в следующем виде

$$\exp(J_k t) = \begin{bmatrix} 1 & \frac{1}{1!}t & \frac{1}{2!}t^2 & \dots & \frac{1}{(r-2)!}t^{r-2} & \frac{1}{(r-1)!}t^{r-1} \\ & 1 & \frac{1}{1!}t & \dots & \frac{1}{(r-3)!}t^{r-3} & \frac{1}{(r-2)!}t^{r-2} \\ & & 1 & \dots & \frac{1}{(r-4)!}t^{r-4} & \frac{1}{(r-3)!}t^{r-3} \\ & & & \dots & \dots & \dots \\ & & & & 1 & \frac{1}{1!}t \\ & & & & & 1 \end{bmatrix} e^{\lambda_k t}.$$

Например, для рассмотренной ранее матрицы

$$A = \begin{bmatrix} -2 & 2 & -3 \\ -10 & 4 & -5 \\ -5 & 4 & -6 \end{bmatrix}$$

имеем

$$\exp(Jt) = \begin{bmatrix} e^{-t} & te^{-t} & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-2t} \end{bmatrix}; \quad \exp(At) = H\exp(Jt)H^{-1};$$

$$\exp(At) = \begin{bmatrix} 1 & 0 & 4 \\ 5 & 2 & 15 \\ 3 & 1 & 10 \end{bmatrix} \begin{bmatrix} e^{-t} & te^{-t} & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-2t} \end{bmatrix} \begin{bmatrix} 5 & 4 & -8 \\ -5 & -2 & 5 \\ -1 & -1 & 2 \end{bmatrix},$$

что приводит к результату, полученному для россматриваемого примера другим методом.

3.31. Конгруэнтное преобразование

Пусть, как и при эквивалентном преобразовании, выполняется переход к различным координатным базисам для векторов x и y в линейном преобразовании $y = Ax$. Потребовав, чтобы при этом сохранялось неизменным скалярное произведение этих векторов, получим другой частный случай эквивалентного преобразования $\tilde{A} = PAQ$ — *конгруэнтное преобразование*. Так как $x'=Q^{-1}x$ и $y' = Py$, и условие $\langle x, y \rangle = \langle x', y' \rangle$ означает $x'y = x'^{t}y'$, что соответствует $x^t y = x^t (Q^{-1})Py$. Это равенство имеет место только при $(Q^{-1})^t P = E$, откуда получаем $P=Q^t$ и, следовательно, $y'=Q^t y$. Таким образом, конгруэнтное преобразование выражается соотношением

$$\tilde{A} = Q^t AQ.$$

Наряду с инвариантностью скалярного произведения это преобразование характерно также тем, что сохраняет свойство симметричности матрицы A. Действительно, если A симметрична ($A = A^t$), то

$$\tilde{A}^t = Q^t A^t Q^{tt} = Q^{(t)} At = Q^t AQ = \tilde{A}$$

т.е. преобразованная матрица $\tilde{A}$ также симметрична. Это обстоятельство широко используется в теории электрических цепей и других физических систем для преобразования системы координат.

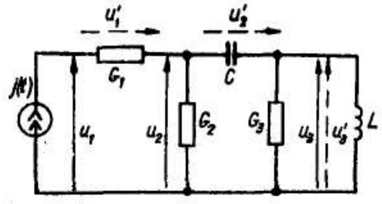


Рис. 3.2 Схема электрической цепи, на примере которой выполнены преобразования системы координат.

Например, схема (рис. 3.2) относительно напряжений u_1, u_2, u_3 описывается системой дифференциальных уравнений (в операторной записи):

$$\begin{bmatrix} G_1 & -G_1 & 0 \\ -G_1 & G_1 + G_2 + pC & -pC \\ 0 & -pC & pC + \frac{1}{pL} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} j(t) \\ 0 \\ 0 \end{bmatrix},$$

где p соответствует оператору дифференцирования d/dt; $1/p$ оператору интегрирования $\int dt$. Пусть требуется перейти к другой системе независимых напряжений u_1', u_2', u_3'. Тогда матрица преобразования Q определяется из зависимостей между напряжениями, которые записываются по второму закону Кирхгофа:

$$\left.\begin{aligned} u_1 &= -u_1' - u_2' + u_3' \\ u_2 &= -u_2' + u_3' \\ u_3 &= u_3' \end{aligned}\right\}; \quad Q = \begin{bmatrix} -1 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Конгруэнтное преобразование переводит матричное уравнение схемы $Yu=j$ к виду $(Q^tYQ)u' = Q^tj$ или $\tilde{Y}u'=j'$, где матрично-векторные параметры в новой системе координат выражаются как $\tilde{Y}=Q^tYQ$ и $j'=Q^tj$. В нашем примере:

$$Y' = Q^t Y Q = \begin{bmatrix} -1 & 0 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} G_1 & -G_1 & 0 \\ -G_1 & G_1 + G_2 + pC & -pC \\ 0 & -pC & pC + \frac{1}{pL} \end{bmatrix} \times$$

$$\times \begin{bmatrix} -1 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} G_1 & 0 & 0 \\ 0 & G_2 + pC & -G_2 \\ 0 & -G_2 & G_2 + \frac{1}{pL} \end{bmatrix};$$

$$J' = \begin{bmatrix} -1 & 0 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} j(t) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -j(t) \\ -j(t) \\ j(t) \end{bmatrix}.$$

Таким образом, уравнение схемы в новой системе координат запишем следующим образом:

$$\begin{bmatrix} G_1 & 0 & 0 \\ 0 & G_2 + pC & -G_2 \\ 0 & -G_2 & G_2 + \frac{1}{pL} \end{bmatrix} \begin{bmatrix} u'_1 \\ u'_2 \\ u'_3 \end{bmatrix} = \begin{bmatrix} -i(t) \\ -i(t) \\ i(t) \end{bmatrix}.$$

Канонической формой, к которой можно привести симметричную матрицу ранга r с помощью конгруэнтного преобразования, является диагональная матрица вида:

$$\Lambda = Q^t A Q = \begin{bmatrix} E_p & & \\ & -E_{r-p} & \\ & & 0 \end{bmatrix},$$

где E_p и E_{r-p} — единичные матрицы порядков p и $r — p$. Число p называется *индексом матрицы,* а $s=p—(r-p)=2p-r$ — *сигнатурой* (предполагается в общем случае, что $r\leq n$). Две матрицы n-го порядка считаются конгруэнтными, если они одинакового ранга и равны их индексы (или сигнатуры).

Биортонормированные базисы. При конгруэнтном преобразовании осуществляется переход к новым различным базисам для двух векторов x и $y,$ участвующих в линейном преобразовании $y=Ax$. При этом $x'=Q^{-1}x=Bx$ и $y' = Q^t y = Cy,$ что соответствует $y'= \tilde{A}x'$, где $\tilde{A}=Q^t AQ$. Столбцы матриц $B = Q^{-1}$ и $C = Q'$ образуют два новых базиса для векторов $x = (x_1, x_2.....x_n)$ и $y = (y_1, y_2.....y_n)$:

$$x' = b^{(1)}x_1 + b^{(2)}x_2 + \cdots + b^{(n)}x_n;$$
$$y' = c^{(1)}y_1 + c^{(2)}y_2 + \cdots + c^{(n)}y_n.$$

Из условия инвариантности скалярного произведения $\langle x, y \rangle = \langle x', y' \rangle$ имеем:

$$\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle b^{(i)}, c^{(j)} \rangle x_i y_j,$$

Откуда

$$\langle b^{(i)}, c^{(j)} \rangle = \delta_{ij} \qquad (i, j = 1, 2, \dots, n),$$

где δ_{ij} — символ Кронекера: $\delta_{ij} = 1$ при $i=j$ и $\delta_{ij}=0$ при $i \neq j$. Две совокупности векторов $\{b^{(1)}, b^{(2)}, \dots, b^{(n)}\}$ и $\{c^{(1)}, c^{(2)}, \dots, c^{(n)}\}$, обладающие этим свойством, называют *биортонормированными*.

Таким образом, конгруэнтное преобразование соответствует переходу к *биортонормированным базисам*.

Ортогональное преобразование. Если потребовать, чтобы преобразование векторов x и y осуществлялось к одному базису (как и преобразование подобия) с сохранением неизменным их скалярного произведения (как в конгруэнтном преобразовании), то следует положить $B=C$ или $Q^{-1} = Q^t$. При этом два биортонормированных базиса сливаются в общий ортонормированный. Матрица A линейного преобразования $x=Ay$ в старых координатах преобразуется в $\tilde{A} = Q^{-1}AQ = Q^tAQ$, причем $Q=(Q^{-1})^t$.

Столбцы матрицы Q^t или строки Q образуют ортонормированную систему векторов в n-мерном евклидовом пространстве, так как они отвечают соотношению

$$< q_{(i)}, q_{(j)} > = \delta_{ij}.$$

В связи с этим матрица Q называется *ортогональной*, а преобразование $A=Q^tAQ$ с такой матрицей — *ортогональным преобразованием*. Ортогональными являются и соответствующие преобразования векторов $x' = Q^tx$ и $y' = Q^ty$.

Ортогональные матрицы обладают следующими свойствами:

1. Если Q ортогональна, то Q^t также ортогональна, так как из условия ортогональности $QQ^t=Q^tQ=E$. Следовательно, совокупность столбцов ортогональной матрицы также образует ортогональный базис.

2. Из соотношения $QQ^t = E$ следует $\det Q \det Q^t = 1$, и так как $\det Q=\det Q^t \neq 0$, то $(\det Q)^2=1$ и $\det Q=\pm 1$, т.е. определитель ортогональной матрицы равен ± 1.

3. Произведение двух ортогональных матриц Q и R есть также — ортогональная матрица. В самом деле, если $QR = QR$, то

$$(V^{-1})^t = (R^tQ^t)^{-1} = (Q^{-1})^t (R^{-1})^t = QR = QR$$

4. Если ортогональная матрица симметрична *($Q=Q^t$)*, то она инволютивна, т.е. равна своей обратной ($Q = Q^{-1}$). Действительно, из $QQ^t = E$ и $Q = Q^t$ следует $QQ = E$ или $Q = Q^{-1}$

Ортогональное преобразование обладает всеми свойствами преобразования подобия и конгруэнтного преобразования. Его дополнительным свойством является инвариантность длины (нормы) векторов. Действительно, так как $x' = Q^t x$, то

$$\langle x', x' \rangle = x'^t x' = x^t (Q^t)^t Q^t x = x^t Q Q^t x = x^t x = \langle x, x \rangle ,$$

поскольку $QQ^t = E$. Ясно, что угол γ между векторами, который определяется соотношением

$$cos\gamma = \frac{\langle x, y\rangle}{\| x \| \cdot \| y \|}$$

где нормы векторов

$$\| x \| = \sqrt{\langle x, x\rangle} \text{ и } \| y \| = \sqrt{\langle y, y\rangle}$$

также не изменяется при ортогональном преобразовании.

Так как ортогональное преобразование (как и конгруэнтное) сохраняет симметрию матрицы A и в тот же время обладает свойствами преобразования подобия, то любую симметричную матрицу можно привести к симметричной канонической форме Жордана. Но симметричность матрицы Жордана означает, что она должна быть диагональной. Таким образом, для симметричной матрицы, независимо от кратности собственных значений всегда можно найти такую матрицу Q, которая приводит ее преобразованием $A = Q^{-1}AQ$ к диагональной форме, на что указывалось раньше.

Можно показать, что собственные значения вещественной симметричной матрицы (а также эрмитовой матрицы) вещественны. Ее собственные векторы образуют ортогональную систему, причем с собственным значением λ_k кратности m_k связанные m_k собственных векторов.

Конъюнктивное и унитарное преобразования. До сих пор предполагалось, что матрица A вещественна и вещественны также преобразующие матрицы, т.е. рассмотренные преобразования относятся к евклидовым пространствам. Они легко обобщаются и для *унитарных пространств,* когда элементами векторов и матриц являются комплексные числа.

Скалярное произведение $< x, y>$ в унитарном пространстве определяется как $x^t y^*$, где y^* — комплексно-сопряженный вектор, а преобразованию $y' = Py$ соответствует $y^{*\prime} = \overline{P} y^*$, где $\overline{P}$ — матрица, комплексно-сопряженная с P. Тогда для конгруэнтного преобразования из условия инвариантности скалярного произведения получим $\overline{P}=Q^t$ или $P=\overline{Q}^t=Q^*$, где Q^* — сопряженная с Q. Конгруэнтное преобразование в унитарном пространстве называют *конъюнктивным преобразованием:*

$$\tilde{A} = Q^*AQ.$$

Если Q — вещественная матрица, то $Q^*=Q^t$ и формула преобразования имеет тот же вид, что и в евклидовом пространстве.

Аналогичное обобщение ортогонального преобразования на унитарное пространство характеризуется соотношением $Q\overline{Q}^t=E$ или $QQ^* = E$, откуда $Q = (Q^*)^{-1}$. Ортогональное преобразование в унитарном пространстве называется *унитарным* и характеризуется соотношением

$$\tilde{A} == Q^{-1}AQ = Q^*AQ,$$

где $Q = (Q^*)^{-1}$. Матрица Q, обратная к своей сопряженной Q^*, называется *унитарной.* Ее определитель выражается комплексным числом, модуль которого равен ± 1. Свойства рассмотренных видов матричных преобразований сведены в табл. 3.1

Таблица 3.1

Матричные преобразования

Преобразование	$y=Ax\to y'=\tilde{A}x'$		
	x'	y'	$\tilde{A}$
Эквивалентное	$Q^{-1}x$	Py	PAQ
Подобия	$Q^{-1}x$	$Q^{-1}y$	$Q^{-1}AQ$
Конгруэнтное	$Q^{-1}x$	Q^ty	Q^tAQ
Ортогональное	$Q^{-1}x=Q^tx$	$Q^{-1}y=Q^ty$	$Q^{-1}AQ= =Q^tAQ$
Конъюнктивное	$Q^{-1}x$	Q^*y	Q^*AQ
Унитарное	$Q^{-1}x= =Q^*x$	$Q^{-1}y= =Q^*y$	$Q^{-1}AQ= =Q^*AQ$

Новые координатные базисы для x и y	Инварианты матричного преобразования
Различные произвольные в евклидовом пространстве	Ранг матрицы
Общий произвольный в евклидовом пространстве	Ранг, определитель и собственные значения матрицы
Различные биортонормированные в евклидовом пространстве	Ранг и симметрия матрицы, скалярное произведение векторов
Общий ортонормированный в евклидовом пространстве $Q^{-1} = Q^t$	Ранг, определитель, собственные значения и симметрия матрицы; скалярное произведение, углы и нормы векторов
Различные биортонормированные в унитарном пространстве	Ранг и сопряженность матрицы; скалярное произведение векторов
Общий ортонормированный в унитарном пространстве $Q^{-1} = Q^*$	Ранг, определитель, собственные значения и сопряженность матрицы; скалярное произведение, углы и нормы векторов

Преобразование квадратичных форм. Квадратичная форма от n переменных $x_1, x_2 ..., x_n$ - это сумма, каждый член которой является квадратом одной из них или произведением двух различных переменных:

$$\psi(x) = (a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n) + (a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n) + \cdots + (a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j.$$

Это выражение можно записать в матричной форме:

$$\psi(x) = [x_1, x_2, \ldots, x_n] \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} =$$
$$= x^t A x = \langle x, Ax \rangle.$$

Матрица A коэффициентов формы называется *ассоциированной* с квадратичной формой $\psi(x)$ или просто *матрицей формы,* а ранг матрицы A является и рангом определяемой ею формы $\psi(x)$. В двойной сумме каждый член, содержащий x_i и x_j, встречается дважды как $a_{ij}x_ix_j$ и $a_{ji}x_jx_i$. Целесообразно положить a_{ij} и a_{ji} равными 1/2 $(a_{ij}+a_{ji})$, от чего сумма не изменится, но зато матрицу A можно считать симметричной, что существенным образом упрощает исследование и преобразование квадратичных форм.

Обозначив $y=Ax$, можно рассматривать квадратичную форму как скалярное произведение $<x, y>$. Очевидно, для преобразования квадратичной формы подходит всякое преобразование матрицы A, не изменяющеее скалярного произведения. Так, при конгруэнтном преобразовании $\tilde{A} = Q^tAQ$ вектор x преобразуется к $x' = Q^{-1}x$, где Q — произвольная неособенная матрица.

Особый интерес представляет преобразование квадратичной формы *к канонической,* в которой она содержит только сумму членов с квадратами переменных x^2_1 , x^2_2, ..., x^2_n Ясно, что для этого необходимо преобразовать симметричную матрицу A к диагональной, что всегда осуществимо с помощью ортогонального преобразования. При этом вектор x преобразуется к $x'=Q^tx$. Разработано несколько практических способов такого преобразования. Один из них иллюстрируется следующим примером.

Пусть данная квадратичная форма

$$\psi(x) = 7x^2_1 + 6x^2_2 + 5x^2_3 — 4x_1x_2— 4\, x_2x_3.$$

Тогда

$$A = \begin{bmatrix} 7 & -2 & -0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{bmatrix}; \quad \lambda E - A = \begin{bmatrix} \lambda - 7 & 2 & 0 \\ 2 & \lambda - 6 & 2 \\ 0 & 2 & \lambda - 5 \end{bmatrix};$$

$$\Delta(\lambda) = \lambda^3 - 18\lambda^2 + 99\lambda - 162; \quad \lambda_1 = 3; \quad \lambda_2 = 6; \quad \lambda_3 = 9.$$

Найдем собственные векторы, соответствующие трем различным собственным значениям. Для этого воспользуемся, например, значениями присоединенной матрицы:

$$G(\lambda_1) = (\lambda_2 E - A)(\lambda_3 E - A); \quad G(\lambda_2) = (\lambda_1 E - A)(\lambda_3 E - A);$$
$$G(\lambda_3) = (\lambda_1 E - A)(\lambda_2 E - A).$$

Так как все столбцы этих матриц пропорциональны, то достаточно вычислить лишь по одном столбцу и принять их или пропорциональные им в качестве столбцов модальной матрицы. После соответствующих вычислений получим:

$$h^{(1)} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}; \qquad h^{(2)} = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}; \qquad h^{( )} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Для того, чтобы применить ортогональное преобразование, пронормируем эти векторы, разделив их составляющие на норму, которая для всех векторов равна:

$$\| h \| = \sqrt{1^2 + 2^2 + 2^2} = 3.$$

Таким образом получаем ортогональную матрицу, которая является и нормированной модальной матрицей преобразования подобия:

$$Q = \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}.$$

Непосредственной проверкой можно убедиться, что $Q^t = Q^{-1}$. По формуле $\tilde{A} = Q^t A Q$ приходим к диагональной матрице:

$$\tilde{A} = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix} =$$
$$= \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

Таким образом, в новому ортонормированном базисе, который образуют столбцы матрицы *Q*, квадратичная форма принимает канонический вид

$$\psi(x') = 3(x_1')^2 + 6(x_2')^2 + 9(x_3')^2,$$

причем

$$x' = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = Q^t x = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$
$$= \frac{1}{3}\begin{bmatrix} x_1 + 2x_2 + 2x_3 \\ -2x_1 - x_2 + 2x_3 \\ 2x_1 - 2x_2 + x_3 \end{bmatrix}.$$

Положительная определенность. Квадратичная форма называется *положительно определенной,* если она положительна для всех значений *x,* кроме *x*=0. Ясно, что это возможно только при положительности виде всех собственных значений неособенной матрицы *A* данной формы.

Квадратичная форма <*x, Ax*> называется *положительно полуопределенной,* если она неотрицательна для всех значений *x* и существуют значения $x \neq 0$, для которых она равна нулю. Этот случай обусловлен требованием неотрицательности всех собственных значений матрицы *A,* среди которых хотя бы одно равняется нулю, т.е. матрица *A* должна быть особенной.

Матрица положительно определенной (полуопределенной) формы также называется *положительно определенной (полуопределенной) матрицей.* Аналогично определяются отрицательно определенная и полуопределенная формы. Если матрица *A* обладает как положительными, так и отрицательными собственными значениями, то квадратичная форма является *неопределенной.*

Характер определенности квадратичной формы можно выяснить, если известные собственные значения ее матрицы или если эта матрица преобразована к канонической форме. Однако такой сложный путь можно обойти, исследуя главные миноры матрицы квадратичной формы.

Можно показать, что необходимым и достаточным условием подложительной определенности квадратичной формы (или симметричной матрицы) является подложительность главных миноров ее матрицы:

$$\Delta_1 = a_{11} > 0; \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0; \quad \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0; \; \ldots$$
$$\ldots\;;\; \Delta_n = |A| > 0.$$

Миноры Δ_k $(k = 1, 2, \ldots, n)$, образованные первыми k строками и k столбцами матрицы, называются ее *дискриминантами,* а приведенное выше условие — *дискриминантным критерием.*

Например,

$$\psi(x) = 2x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1x_2 - 2x_2x_3$$

является положительно определенной, так как

$$\Delta_3 = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 4 \end{vmatrix} > 0; \quad \Delta_2 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} > 0; \quad \Delta_1 = 2 > 0.$$

Микромодуль 9
Индивидуальные тестовые задачи

1. Укажите операции над строками (столбцами) матрицы четвертого порядка, которым соответствует умножение ее слева (справа) на матрицы:

а) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$; б) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; в) $\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$;

г) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; д) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{bmatrix}$; е) $\begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 3 & 1 \end{bmatrix}$.

Какие из этих матриц являются элементарными? Выразите неэлементарные матрицы через произведения матриц, которые осуществляют элементарные преобразования.

2. Найдите матрицу, при умножении которой на квадратную матрицу четвертого порядка (слева или справа) осуществляются следующие эквивалентные преобразования:

а) перестановка второй и четвертой строк, а также прибавление к первой строке второй строки, умноженной на 2, и вычитание из первой строки третьей строки (матрица *P);*

б) перестановка первого и второго столбцов, умножение четвертого столбца на 3 и вычитание из третьего столбца удвоенного первого столбца (матрица *Q).*

3. Дана матрица

$$A = \begin{bmatrix} 2 & 7 & 1 & 3 \\ -1 & 3 & 0 & -5 \\ 4 & 2 & -1 & 1 \\ 0 & 1 & 2 & 4 \end{bmatrix}.$$

а) Используя результаты предыдущей задачи, выполните умножение PA и AQ и убедитесь в выполнении соответствующих преобразований над строками и столбцами маирицы A.

б) Определите эквивалентную матрицу по формуле $\tilde{A}= PAQ$.

в) С помощью элементарных преобразований приведите матрицы A и $\tilde{A}$ к канонической форме и покажите, что эти матрицы имеют одинаковые ранги.

4. Какие из приведенных ниже матриц являются эквивалентными:

$$\text{а) } \begin{bmatrix} 4 & 3 & 6 \\ 1 & 5 & 7 \\ 2 & 9 & 1 \end{bmatrix}; \quad \text{б) } \begin{bmatrix} 0 & 4 & 1 \\ 4 & 8 & 7 \\ 1 & 7 & 3 \end{bmatrix}; \quad \text{в) } \begin{bmatrix} 4 & 1 & 2 \\ 0 & 1 & 0 \\ 8 & 4 & 5 \end{bmatrix}?$$

5. Покажите, что эквивалентность матриц обладает всеми свойствами отношения эквивалентности (рефлексивность, симметричность, транзитивность).

6. Найдите матрицу, подобную A, с помощью преобразующей матрицы Q, если

$$A = \begin{bmatrix} 2 & 5 & 0 & 1 \\ 0 & 3 & 4 & 2 \\ -1 & 1 & 5 & 0 \\ 3 & 2 & 0 & -3 \end{bmatrix}; \quad Q = \begin{bmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{bmatrix}.$$

7. С помощью преобразования подобия приведите к диагональной форме следующие матрицы:

$$\text{а) } \begin{bmatrix} 1 & -2 & 2 \\ 1 & 4 & -2 \\ 1 & 5 & -3 \end{bmatrix}; \quad \text{б) } \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

8. Дана матрица

$$A = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix}.$$

а) Найдите подобную ей диагональную матрицу и объясните, почему такая матрица существует?

б) Покажите, что преобразующая матрица H в преобразовании подобия $A = H^{-1}AH$ является ортогональной, т.е. $H^{-1} = H^{t}$.

9. Найдите каноническую форму Жордана для следующих матриц:

$$\text{а)}\ \begin{vmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{vmatrix}; \qquad \text{б)}\ \begin{vmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{vmatrix}; \qquad \text{в)}\ \begin{vmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{vmatrix}.$$

10. Для каждой из матриц, приведенных в предыдущей задаче, найдите модальную матрицу H, преобразующую к канонической форме Жордана, и проверьте результат по формуле $J = H^{-1}AH$.

11. Запишите экспоненциальные функции от матриц, приведенных в задаче 9.

12. Жорданaнова форма некоторой матрицы A имеет вид:

$$J = \begin{bmatrix} 2 & 1 & & & & & & \\ 0 & 2 & & & & & & \\ & & 2 & & & & & \\ & & & -1 & 1 & 0 & & \\ & & & 0 & -1 & 1 & & \\ & & & 0 & 0 & -1 & & \\ & & & & & & 5 & 1 \\ & & & & & & 0 & 5 \end{bmatrix}.$$

Что можно сказать о характеристических числах λ_k и дефектах матриц $F(\lambda_k) = \lambda_k E - A$?

13. Покажите, что любая числовая квадратная матрица подобна треугольной (верхний или нижний) матрицы над полем комплексных чисел. В каких случаях элементы треугольной матрицы комплексные?

14. Покажите, что в преобразовании подобия $\tilde{A} = Q^{-1}AQ$ роль преобразующей матрицы, наряду с Q, может играть также любая матрица $Q' = CQ$, где C — произвольная неособенная матрица, перестановочная с A.

15. Покажите, что матрица

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{bmatrix}$$

приводится к канонической форме Жордана J с помощью преобразующей матрицы H, где

$$J=\begin{bmatrix}-1 & 1 & 0 & 0\\ 0 & -1 & 1 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix};\quad H=\begin{bmatrix}0 & 1 & 0 & 1\\ 1 & -5 & 0 & 0\\ 0 & 4 & 1 & 5\\ -1 & 11 & 0 & 7\end{bmatrix}.$$

Проверьте, что $J = H^{-1}AH$ (или $HJ = AH$).

16. Матрицы являются подобными, если преобразованием подобия они приводятся к одинаковой канонической форме. Являются ли подобными матрицы

$$A=\begin{bmatrix}2 & 2 & 1\\ 1 & 3 & 1\\ 1 & 2 & 2\end{bmatrix};\quad B=\begin{bmatrix}2 & 1 & -1\\ 0 & 2 & -1\\ -3 & -2 & 3\end{bmatrix}?$$

17. Найдите модальную матрицу, каноническую форму и экспоненциальную функцию для каждой из следующих матриц:

$$\text{а)}\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 1 & 3 & 2\\ 0 & 1 & -1 & 2\end{bmatrix};\quad \text{б)}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & -1 & -1 & -1\end{bmatrix};\quad \text{в)}\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\end{bmatrix}.$$

18. С помощью преобразования подобия найдите фундаментальную матрицу Ф(t) системы дифференциальных уравнений:

$$\left.\begin{aligned}\frac{dx_1}{dt}&=3x_1-x_2-4x_3+2x_4\\ \frac{dx_2}{dt}&=2x_1+3x_2-2x_3-4x_4\\ \frac{dx_3}{dt}&=2x_1-x_2-3x_3+2x_4\\ \frac{dx_4}{dt}&=x_1+2x_2-x_3-3x_4\end{aligned}\right\}.$$

19. Дана квадратичная форма

$$\psi(x)=x_1^2-2x_2^2+x_3^2+4x_1x_2-8x_1x_3-4x_2x_3.$$

а) Запишите матрицу A формы и представьте ее в виде $\psi(x) = x^tAx$.

б) Приведите данную форму к каноническому виду.

в) Выразите векторы, образующие новый ортонормированный базис, через компоненты вектора x.

г) Установите, является ли данная форма положительно определенной.

20. Какие из приведенных ниже матриц являются положительно определенными:

$$\text{а)}\begin{bmatrix}1 & 1 & 1\\ 1 & 3 & 1\\ 1 & 1 & 1\end{bmatrix};\quad \text{б)}\begin{bmatrix}2 & 1 & 1\\ 1 & 3 & 0\\ 1 & 0 & 1\end{bmatrix};\quad \text{в)}\begin{bmatrix}2 & 1 & 1\\ 1 & 3 & 2\\ 1 & 2 & 1\end{bmatrix}?$$

Модуль 4
Инвариантные многочлены и матричные уравнения

Микромодуль 10
Эквивалентные преобразования многочленных матриц

4.1. Инвариантные многочлены и элементарные делители многочленной матрицы

1. Введем понятие об инвариантных многочленах λ-матрицы $A(\lambda)$.

Пусть многочленная матрица $A(\lambda)$ имеет ранг r, т.е. в этой матрице имеются не равные тождественно нулю миноры r-го порядка, в то время как все миноры порядка $>r$ равны нулю тождественно относительно λ. Обозначим через $D_j(\lambda)$ наибольший общий делитель всех миноров λ-го порядка матрицы $A(\lambda)$ ($j=1,2,\ldots,r$) (в каждом $D_j(\lambda)$ берем старший коэффициент равным единицы). Тогда, как нетрудно видеть, в ряду

$$D_r(\lambda),\quad D_{r-1}(\lambda),\ \ldots,\ D_1(\lambda),\ \ D_0(\lambda)\equiv 1$$

каждый многочлен делится без остатка на следующий (если к какому-нибудь минору j-го порядка применить разложение Безу по элементам какой-либо строки, то каждое слагаемое в этом разложении будет делиться на $D_{j-1}(\lambda)$; следовательно, любой минор j-го порядка, а значит и $D_j(\lambda)$, делится на $D_{j-1}(\lambda)$ ($j=1,2,\ldots,r$). Соответствующие частные обозначим через $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$:

$$i_1(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)},\quad i_2(\lambda)=\frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)},\ \ldots,\ i_r(\lambda)=\frac{D_1(\lambda)}{D_0(\lambda)}=D_1(\lambda). \tag{4.1}$$

Определение 1. Многочлены $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$, которые определяются формулами (4.1), называются *инвариантными многочленами* прямоугольной матрицы $A(\lambda)$.

Термин «инвариантные многочлены» связан со следующими соображениями. Пусть $A(\lambda)$ и $B(\lambda)$ — две эквивалентные многочленные матрицы. Тогда они получаются друг из друга при помощи элементарных операций. Но нетрудно непосредственно проверить, что элементарные операции не изменяют ни ранга r матрицы $A(\lambda)$, ни самих многочленов $D_1(\lambda)$, $D_2(\lambda)$, ..., $D_r(\lambda)$. Действительно, применяя к тождеству (3.334) формулу, которая выражает минор произведения матриц через миноры сомножителей, мы для произвольного минора матрицы $B(\lambda)$ получим выражение

$$B\begin{pmatrix} j_1 \, j_2 \, \dots \, j_p \\ k_1 \, k_2 \, \dots \, k_p \end{pmatrix} ; \lambda \Bigg) =$$

$$= \sum_{\substack{1 \leq \alpha_1 < \alpha_2 < \dots < \alpha_p \leq m \\ 1 \leq \beta_1 < \beta_2 < \dots < \beta_p \leq m}} P\begin{pmatrix} j_1 & j_2 & \dots & j_p \\ \alpha_1 & \alpha_2 & \dots & \alpha_p \end{pmatrix} A\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ \beta_1 & \beta_2 & \dots & \beta_p \end{pmatrix}; \lambda\Bigg) Q\begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}$$

$$(p = 1, 2, \dots, \min(m, n)).$$

Отсюда следует, что все миноры порядка $\geq r$ матрицы $B(\lambda)$ равны нулю и, следовательно, для ранга r^* матрицы $B(\lambda)$ имеем:

$$r^* \leq r.$$

Кроме того, из этой же формулы следует, что $D_p^*(\lambda)$ — наибольший общий делитель всех миноров p-го порядка матрицы $B(\lambda)$ — делится на $D_p(\lambda)$ нацело $(p = 1, 2, ..., \min (m, n))$. Но матрицы $A(\lambda)$ и $B(\lambda)$ можно поменять ролями. Поэтому $r \leq r^*$, и $D_p(\lambda)$ делится без остатка на $D_p^*(\lambda)$ $(p = 1, 2, ..., \min (m, n))$. Отсюда (старшие коэффициенты в $D_p(\lambda)$ и $D^*(\lambda)$ $(p= 1,2, ..., r)$ равны единице).

$$r = r^*, \quad D_1^*(\lambda) = D_1(\lambda), \quad D_2^*(\lambda) = D_2(\lambda), \quad \dots, \quad D_r^*(\lambda) = D_r(\lambda).$$

Поскольку элементарные операции не меняют многочленов $D_1(\lambda)$, $D_2(\lambda)$, ..., $D_r(\lambda)$, то они не изменяют и многочленов $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$, определяемых формулами (4.1).

Таким образом, многочлены $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$ остаются неизменными, инвариантными при переходе от одной матрицы к другой, ей эквивалентной.

Если многочленная матрица имеет канонический диагональный вид (3.341), то, как нетрудно видеть, для этой матрицы

$$D_1(\lambda) = a_1(\lambda), \quad D_2(\lambda) = a_1(\lambda)\, a_2(\lambda), \dots, D_r(\lambda) = a_1(\lambda)\, a_2(\lambda) \dots a_r(\lambda).$$

Но тогда в силу соотношений (4.1) диагональные многочлены $a_1(\lambda)$, $a_2(\lambda)$, ..., $a_r(\lambda)$ в (3.341) совпадают с инвариантными многочленами

$$i_1(\lambda) = a_r(\lambda), \quad i_2(\lambda) = a_{r-1}(\lambda), \dots, i_r(\lambda) = a_1(\lambda). \tag{4.2}$$

Здесь $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$ являются одновременно и инвариантными многочленами исходной матрицы $A(\lambda)$, поскольку эта матрица эквивалентна матрице (3.341).

Полученные результаты мы можем сформулировать в виде следующей теоремы.

Теорема 1. *Многочленная прямоугольная матрица* $A(\lambda)$ *всегда эквивалентна канонической диагональной матрице*

$$\left\| \begin{matrix} i_r(\lambda) & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & i_{r-1}(\lambda) & \ldots & 0 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & i_1(\lambda) & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{matrix} \right\| .$$

(4.3)

При этом здесь обязательно r — ранг, а $i_1(\lambda)$, i_2 (λ), ..., $i_r(\lambda)$ — *инвариантные многочлены матрицы* $A(\lambda)$, *которые определяются формулами* (4.1).

Следствие 1. *Для того чтобы две прямоугольные матрицы одинаковых размеров* A (λ) *и* B (λ) *были эквивалентны, необходимо и довольно, чтобы они имели одни и те же инвариантные многочлены.*

Действительно, необходимость этого условия была выяснена выше. Достаточность следует из того, что две многочленные матрицы, имеющие одни и те же инвариантные многочлены, эквивалентны одной и той же канонической диагональной матрице и, следовательно, эквивалентны между собой.

Таким образом, *инвариантные многочлены образуют полную систему инвариантов* λ *-матрицы.*

Следствие 2. В ряду инвариантных многочленов

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \ldots, i_r(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \qquad (D_0(\lambda) \equiv 1)$$

(4.3)

каждый многочлен, начиная со второго, является делителем предыдущего.

Это утверждение не вытекает непосредственно из формул (3.4). Оно вытекает из того, что многочлены $i_1(\lambda)$, i_2 (λ), ..., $i_r(\lambda)$ совпадают с многочленами $a_1(\lambda)$, a_2 (λ), ..., $a_r(\lambda)$ канонической диагональной матрицы (3.341).

2. Укажем методы вычисления инвариантных многочленов для квазидиагональных λ-матриц, если известны инвариантные многочлены матриц, которые стоят в диагональных клетках.

Теорема 2. *Если в квазвдиагональной прямоугольной матрице*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

любой инвариантный многочлен матрицы $A(\lambda)$ *является делителем любого инвариантного многочлена матрицы* $B(\lambda)$, *то совокупность инвариантных многочленов матрицы* $C(\lambda)$ *получается объединением инвариантных многочленов матриц* $A(\lambda)$ *и* $B(\lambda)$.

Доказательство. Обозначим через $i'_1(\lambda)$, $i'_2(\lambda)$, ..., $i'_r(\lambda)$ и $i''_1(\lambda)$, $i''_2(\lambda)$, ..., $i''_q(\lambda)$ соответственно инвариантные многочлены λ-матриц $A(\lambda)$ и $B(\lambda)$. Тогда (знаком ~ мы здесь обозначаем эквивалентность матриц, а фигурными скобками { } — диагональную прямоугольную матрицу вида (4.3))

$$A(\lambda) \sim \{i'_r(\lambda), \ldots, i'_1(\lambda), 0, \ldots, 0\}, \quad B(\lambda) \sim \{i''_q(\lambda), \ldots, i''_1(\lambda), 0, \ldots, 0\}$$

и, следовательно,

$$C(\lambda) \sim \{i'_r(\lambda), \ldots, i'_1(\lambda), \quad i''_q(\lambda), \ldots, i''_1(\lambda), 0, \ldots, 0\}. \tag{4.5}$$

λ -матрица, которая стоит в правой части этого соотношения, имеет каноническую диагональную форму. Тогда согласно теореме 1 не равные тождественно нулю диагональные элементы этой матрицы образуют полную систему инвариантных многочленов матрицы $C(\lambda)$. Теорема доказана.

Для того чтобы в общем случае при произвольных инвариантных многочленах матриц $A(\lambda)$ и $B(\lambda)$ определить инвариантные многочлены $C(\lambda)$, мы воспользуемся важным понятием об элементарных делителях.

Разложим инвариантные многочлены $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$ на неприводимые в данном числовом поле K множители:

$$\begin{aligned} i_1(\lambda) &= [\varphi_1(\lambda)]^{c_1} [\varphi_2(\lambda)]^{c_2} \ldots [\varphi_s(\lambda)]^{c_s}, \\ i_2(\lambda) &= [\varphi_1(\lambda)]^{d_1} [\varphi_2(\lambda)]^{d_2} \ldots [\varphi_s(\lambda)]^{d_s}, \qquad \begin{pmatrix} c_k \geqslant d_k \geqslant \ldots \geqslant l_k \geqslant 0; \\ k = 1, 2, \ldots, s \end{pmatrix} \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ i_r(\lambda) &= [\varphi_1(\lambda)]^{l_1} [\varphi_2(\lambda)]^{l_2} \ldots [\varphi_s(\lambda)]^{l_s} \end{aligned} \tag{4.6}$$

(Некоторые из показателей c_k, d_k, ...,l_k (k= 1,2, ..., s) могут равняться нулю).

Здесь $\varphi_1(\lambda)$, $\varphi_2(\lambda)$, ..., $\varphi_s(\lambda)$ — все различные неприводимые в поле K многочлены (со старшими коэффициентами, равными единицы), входящие в состав $i_1(\lambda)$, $i_2(\lambda)$, ..., $i_r(\lambda)$.

Определение 2. Все отличные от единицы степени среди

$$[\varphi_1(\lambda)]^{c_1},...,[\varphi_s(\lambda)]^{l_s}$$

в разложении (4.6) называются *элементарными делителями* матрицы $A(\lambda)$ в поле K (формулы (4.6) дают возможность не только по инвариантным многочленах определить элементарные делители $A(\lambda)$ в поле K, но и наоборот — по элементарным делителям определить инвариантные многочлены).

Теорема 3. *Совокупность элементарных делителей прямоугольной квазидиагональноц матрицы*

$$C(\lambda)=\begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

всегда получается объединением элементарных делителей матрицы $A(\lambda)$ *с элементарными делителями матрицы* $B(\lambda)$.

Доказательство. Разложим инвариантные многочлены матриц $A(\lambda)$ и $B(\lambda)$ на неприводимые в поле K множители (если какой-либо неприводимый многочлен, $\varphi_k(\lambda)$ входит множителем в одни инвариантные многочлены и не входит в другие, то в эти последние многочлены мы вписываем $\varphi_k(\lambda)$ с нулевым показателем):

$$i_1'(\lambda)=[\varphi_1(\lambda)]^{c_1'}[\varphi_2(\lambda)]^{c_2'}\ldots[\varphi_s(\lambda)]^{c_s'},\quad i_1''(\lambda)=[\varphi_1(\lambda)]^{c_1''}[\varphi_2(\lambda)]^{c_2''}\ldots[\varphi_s(\lambda)]^{c_s''},$$

$$i_2'(\lambda)=[\varphi_1(\lambda)]^{d_1'}[\varphi_2(\lambda)]^{d_2'}\ldots[\varphi_s(\lambda)]^{d_s'},\quad i_2''(\lambda)=[\varphi_1(\lambda)]^{d_1''}[\varphi_2(\lambda)]^{d_2''}\ldots[\varphi_s(\lambda)]^{d_s''},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$i_r'(\lambda)=[\varphi_1(\lambda)]^{h_1'}[\varphi_2(\lambda)]^{h_2'}\ldots[\varphi_s(\lambda)]^{h_s'},\quad i_q''(\lambda)=[\varphi_1(\lambda)]^{g_1''}[\varphi_2(\lambda)]^{g_2''}\ldots[\varphi_s(\lambda)]^{g_s''}.$$

Обозначим через

$$c_1\geq d_1\geq ...\geq=l_1>0 \qquad (4.7)$$

все отличные от нуля числа среди

$$c_1',\ d_1',\ \ldots,\ h_1',\ c_1'',\ d_1'',\ \ldots,\ g_1''.$$

Тогда матрица $C(\lambda)$ эквивалентна матрице (4.5), а последняя перестановкой строк и столбцов может быть приведена к «диагональному» виду

$$\{[\varphi_1(\lambda)]^{c_1}\cdot(*),\ [\varphi_1(\lambda)]^{d_1}\cdot(*),\ldots,[\varphi_1(\lambda)]^{l_1}\cdot(*),\ (**),\ \ldots,(**)\}, \qquad (4.8)$$

где через ($*$) мы обозначили многочлены, взаимно простые с $\varphi_1(\lambda)$, а через ($*$ $*$) — многочлены, взаимно простые с $\varphi_1(\lambda)$ либо тождественно уровные нулю. Из вида матрицы (4.8) непосредственно вытекают следующие разложения многочленов $D_r(\lambda)$, $D_{r-1}(\lambda)$, ... и $i_1(\lambda)$, $i_2(\lambda)$, ... для матрицы $C(\lambda)$:

$$D_r(\lambda) = [\varphi_1(\lambda)]^{c_1+d_1+\dots+l_1} \cdot (*), \qquad D_{r-1}(\lambda) = [\varphi_1(\lambda)]^{d_1+\dots+l_1} \cdot (*) \text{ и т. д.,}$$

$$i_1(\lambda) = [\varphi_1(\lambda)]^{c_1} \cdot (*), \quad i_2(\lambda) = [\varphi_1(\lambda)]^{d_1} \cdot (*) \text{ и т. д.}$$

Отсюда следует, что

$$[\varphi_1(\lambda)]^{c_1}, [\varphi_1(\lambda)]^{d_1}, \dots, [\varphi_1(\lambda)]^{l_1},$$

т.е. все не равные единицы из степеней

$$[\varphi_1(\lambda)]^{c_1'}, \dots, [\varphi_1(\lambda)]^{h_1'}, \quad [\varphi_1(\lambda)]^{c_1''}, \dots, [\varphi_1(\lambda)]^{g_1''},$$

являются элементарными делителями матрицы $C(\lambda)$.

Аналогично определяются элементарные делители матрицы $C(\lambda)$, являющиеся степенями $\varphi_2(\lambda)$ и т.д. Теорема доказана.

Примечание. Совсем аналогично предыдущему можно построить теорию эквивалентности для целочисловых матриц (т.е. матриц, в которых элементы — целые числа). При этом в 1°, 2° $c=\pm 1$, $b(\lambda)$ заменяется целым числом, а в формулах (3.332), (3.333), (3.334) вместо $P(\lambda)$ и $Q(\lambda)$ стоят целочисловые матрицы с определителями, равными ± 1.

3. Пусть дана теперь матрица $A=\|a_{ik}\|^n_1$ с элементами из поля K. Составим для нее характеристическую матрицу

$$\lambda E - A = \left\| \begin{matrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{matrix} \right\| . \tag{4.9}$$

Характеристическая матрица является λ-матрицей ранга n. Ее инвариантные многочлены

$$i_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \dots, i_n(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1) \tag{4.10}$$

называются *инвариантными многочленами матрицы А,* а соответствующие элементарные делители в поле K — *элементарными делителями в поле К матрицы А*. Первый инвариантный многочлен $i_1(\lambda)$ совпадает с минимальным многочленом $\psi(\lambda)$ матрицы A (см. формулу (3.56); там $\Delta(\lambda)\equiv D_n(\lambda)$). Знание инвариантных многочленов (а следовательно, и элементарных делителей) матрицы A позволяет исследовать ее структуру. Поэтому представляют интерес практические способы вычисления инвариантных многочленов матрицы. Сами формулы (4.10) дают алгоритм для вычисления этих многочленов, но этот алгоритм при больших n очень громоздкий.

Теорема 1 дает другой способ вычисления инвариантных многочленов, основанный на приведении характеристической матрицы (4.9) при помощи элементарных операций к каноническому диагональному виду.

Пример.

$$A = \begin{Vmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{Vmatrix}, \quad \lambda E - A = \begin{Vmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ -6 & -1 & \lambda-2 & -1 \\ 14 & 5 & 1 & \lambda \end{Vmatrix}.$$

В характеристической матрице λЕ— *А* прибавим к четвертой строке третью, предварительно помноженную на λ:

$$\begin{Vmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ -6 & -1 & \lambda-2 & -1 \\ 14-6\lambda & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

Теперь, прибавляя к первым трем столбцам четвертый, предварительно помноженный соответственно на —6, —1, λ — 2, получим:

$$\begin{Vmatrix} \lambda-3 & -1 & 0 & 0 \\ 4 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 14-6\lambda & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

К первому столбцу прибавляем второй, помноженный на λ - 3:

$$\begin{Vmatrix} 0 & -1 & 0 & 0 \\ \lambda^2-2\lambda+1 & \lambda+1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2+2\lambda-1 & 5-\lambda & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

Прибавим к второц и четвертоц строкам первую, помноженную соответственно на λ +1 и 5— λ*;* найдем:

$$\begin{Vmatrix} 0 & -1 & 0 & 0 \\ \lambda^2-2\lambda+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2+2\lambda-1 & 0 & \lambda^2-2\lambda+1 & 0 \end{Vmatrix}.$$

Прибавим к четвертой строке вторую, затем умножим первую и третью строки на -1. После перестановки строк и столбцов получим:

$$\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (\lambda-1)^2 & 0 \\ 0 & 0 & 0 & (\lambda-1)^2 \end{Vmatrix}.$$

Матрица A имеет два элементарных делителя: $(\lambda - 1)^2$ и $(\lambda - 1)^2$.

4.2. Эквивалентность линейных двучленов

В предыдущих разделах мы рассматривали прямоугольные λ-матрицы. В этом же разделе мы рассмотрим две квадратные λ-матрицы $A(\lambda)$ и $B(\lambda)$ n-го порядка, в которых все элементы имеют степень не выше единицы относительно λ. Эти многочленные матрицы могут быть представлены в виде матричных двучленов:

$$A(\lambda) = A_0 \lambda + A_1, \quad B(\lambda) = B_0 \lambda + B_1.$$

Будем предполагать, что эти двучлены имеют первую степень и регулярны, т.е. что $|A_0| \neq 0$, $|B_0| \neq 0$.

Следующая теорема устанавливает критерий эквивалентности таких двучленов:

Теорема 4. *Если два регулярных двучлена первой степени $A_0\lambda+A_1$ и $B_0\lambda + B_1$ эквивалентны, то эти двучлены строго эквивалентны, т.е. в тождестве*

$$B_0\lambda + B_1 = P(\lambda)(A_0\lambda+A_1)Q(\lambda) \qquad (4.11)$$

*можно $P(\lambda)$ и $Q(\lambda)$ — матрицы с постоянными и отличными от нуля определителями — заменить постоянными неособенными матрицами P і Q (*тождество (4.12) равносильно двум матричным равенствам: $B_0 = PA_0Q$ и $B_1 = PA_1Q$):

$$B_0\lambda + B_1 = P(A_0\lambda+A_1)Q. \qquad (4.12)$$

Доказательство. Так как определитель матрицы $P(\lambda)$ не зависит от λ и отличен от нуля, то обратная матрица $M(\lambda)=P^{-1}(\lambda)$ также будет многочленной. Пользуясь этой матрицей, мы тождество (4.11) перепишем в виде

$$M(\lambda)(B_0\lambda + B_1) = (A_0\lambda+A_1)Q(\lambda). \qquad (4.13)$$

(Эквивалентность двучленов $A_0\lambda+A_1$ и $B_0\lambda+B_1$ означает наличие тождества (4.11), в котором $|P(\lambda)| = \text{const} \neq 0$ и $|Q(\lambda)| = \text{const} \neq 0$. Однако последние соотношения в этом случае вытекают из самой тождества (4.11). Действительно, определители регулярных двучленов первой степени имеют степень n:

$$|A_0\lambda+A_1| = |A_0|\lambda^n + \ldots, \ |B_0\lambda+B_1| = B_0|\lambda^n + \ldots; \ |A_0| \neq 0, \ |B_0| \neq 0.$$

Поэтому из

$$| B_0\lambda+B_1| = | P(\lambda) | A_0\lambda+A_1| | Q (\lambda) |$$

следует:

$$|P(\lambda)| = \text{const} \neq 0 \ , | Q (\lambda) | = \text{const} \neq 0).$$

Рассматривая M (λ) и Q (λ) как матричные многочлены, разделим M (λ) слева на $A_0\lambda+A_1$, а Q (λ) — справа на $B_0\lambda+B_1$:

$$M (\lambda)=(A_0\lambda+A_1) S(\lambda) +M \qquad (4.14)$$

$$Q (\lambda)= T(\lambda)(B_0\lambda+B_1)+ Q \qquad (4.15)$$

здесь M и Q — постоянные (не зависящие от λ) квадратные матрицы n-го порядка. Полученные выражения для M(λ) и Q(λ) подставим в (4.13). После небольших преобразований получим:

$$(A_0\lambda+A_1)[T(\lambda)- S(\lambda)](B_0\lambda+B_1)=M (B_0\lambda+B_1)- (A_0\lambda+A_1) Q. \qquad (4.16)$$

Разность, которая стоит в квадратных скобках, должна тождественно равняться нулю, так как в противном случае произведение, которое стоит в левой части равенства (4.16), имело бы степень ≥ 2, в то время как в правой части этого же равенства стоит многочлен не выше первой степени. Поэтому

$$S(\lambda)= T(\lambda); \qquad (4.17)$$

но тогда из (4.16) получим:

$$M (B_0\lambda+B_1)= (A_0\lambda+A_1) Q. \qquad (4.18)$$

Покажем теперь, что M — неособенная матрица. Для этого разделим P (λ) слева на $B_0\lambda+B_1$:

$$P(\lambda) = (B_0 \lambda + B_1)U(\lambda) + P. \qquad (4.19)$$

Из (4.13), (4.14) и (4.19) следует:

$$E= M(\lambda) P(\lambda)= M(\lambda) (B_0 \lambda + B_1)U(\lambda)+ M(\lambda)P=$$
$$=(A_0\lambda+A_1) Q(\lambda) U(\lambda)+ (A_0\lambda+A_1) S(\lambda)P+ MP=$$
$$= (A_0\lambda+A_1)[Q(\lambda) U(\lambda)+ S(\lambda)P]+ MP \qquad (4.20)$$

Так как последняя часть этой цепочки равенств должна иметь нулевую степень относительно λ (поскольку она равна E), то выражение в квадратных скобках должно тождественно равняться нулю. Но тогда из (4.20)

$$MP = E. \qquad (4.21)$$

Отсюда вытекает: $| M | \neq 0$ и $M^{-1} = P$.

Умножая обе части равенства (4.18) слева на P, получим:

$$B_0 \lambda + B_1 = P(A_0 \lambda + A_1)Q.$$

Неособенность матрицы P следует из (4.21). Но неособенность матриц P и Q следует и из самого тождества (4.12), так как из этого тождества следует равенство

$$B_0 = PA_0Q$$

и потому

$$|P| \, | A_0| \, | Q | = |B_0| \neq 0.$$

Теорема доказана.

Примечание. Из доказательства следует [см. (4.15) и (4.19)], что в качестве постоянных матриц P и Q, которыми мы заменяем λ-матрицы $P(\lambda)$ и $Q\ (\lambda)$ в тождестве (4.11), можно взять соответственно левый и правый остатки от деления $P\ (\lambda)$ и $Q\ (\lambda)$ на $B_0\lambda + B_1$.

4.3. Критерий подобия матриц

Пусть дана матрица $A = \|\ a_{ik}\ \|^n_1$ с числовыми элементами из поля K. Ее характеристическая матрица $\lambda E — A$ является λ-матрицей ранга n и потому имеет n инвариантных многочленов

$$i_1\ (\lambda), i_2\ (\lambda),\ ..., i_n\ (\lambda).$$

Следующая теорема показывает, что эти инвариантные многочлены определяют исходную матрицу A с точностью до преобразования подобия.

Теорема 5. *Для того чтобы две матрицы* $A = \|\ a_{ik}\ \|^n_1$ *и* $B = \|\ b_{ik}\ \|^n_1$ *были подобны* ($B=T^{-1}AT$), *необходимо и достаточно, чтобы они имели одни и те же инвариантные многочлены, или, что то же, одни и те же элементарные делители в поле К.*

Доказательство. ***Условие необходимости***. Действительно, если матрицы A и B подобны, то существует такая неособенная матрица T, что

$$B = T^{-1}AT.$$

Отсюда

$$\lambda E — B = T^{-1}\ (\lambda E - A)\ T.$$

Это равенство показывает, что характеристические матрицы $\lambda E — A$ и $\lambda E — B$ эквивалентны и поэтому имеют одни и те же инвариантные многочлены.

Условие достаточности. Пусть характеристические матрицы $\lambda E — A$ і $\lambda\ E — B$ одни и те же инвариантные многочлены. Тогда эти λ-матрицы эквивалентны (см. следствие 1 из теоремы 1), и, следовательно, существуют две многочленные матрицы $P\ (\lambda)$ и $Q\ (\lambda)$ такие, что

$$\lambda\ E - B = P\ (\lambda)\ (\lambda\ E - A)\ Q\ (\lambda). \qquad (4.22)$$

Применяя к матричным двучленам $\lambda E — A$ і $\lambda E — B$ теорему 2, мы можем в тождестве (4.22) заменить λ-матрицы $P(\lambda)$ и $Q(\lambda)$ постоянными матрицами P и Q:

$$\lambda\ E- B = P(\lambda\ E-A)Q, \qquad (4.23)$$

причем в качестве P и Q можно взять соответственно левый и правый остатки от деления $P(\lambda)$ и $Q(\lambda)$ на $\lambda E — B$, т.е. на основании обобщенной теоремы Безу можно положить (напоминаем, что $\hat{P}(B)$ — левое значение многочлена $P(\lambda)$, а $Q(B)$ — правое значение многочлена $Q(\lambda)$ после замены λ на B):

$$P = \hat{P}(B), \quad Q = Q(B). \tag{4.24}$$

Приравнивая в обеих частях равенства (4.23) коэффициенты при нулевой и при первой степенях λ, получим:

$$B = PAQ, \quad E = PQ,$$

т.е.

$$B = T^{-1}AT,$$

где

$$T = Q = P^{-1}$$

Теорема доказана.

Замечание. Попутно нами установленное следующее предложение, что мы сформулируем как

***Добавление к теореме* 5**. *Если* $A = \| a_{ik} \|^n_1$ *и* $B = \| b_{ik} \|^n_1$ — *две подобные матрицы*

$$B= T^{-1}AT, \tag{4.25}$$

это в качестве преобразующей матрица T можно взять матрицу

$$T=Q(B) = [\hat{P}(B)]^{-1} \tag{4.26}$$

где $P(\lambda)$ *и* $Q(\lambda)$ — *многочленные матрицы в тождестве*

$$\lambda E-B = P(\lambda)(\lambda E - A)\, Q(\lambda),$$

связывающем эквивалентные характеристические матрицы $\lambda E — A$ *и* $\lambda E — B$; *в формуле* (4.26) $Q(B)$ *обозначает правое значение матричного многочлена* $Q(\lambda)$, *а* $\hat{P}(B)$ — *левое значение матричного многочлена* $P(\lambda)$ *при замене аргумента* λ *матрицей B.*

4.4. Нормальные формы матрицы

1. Пусть дан некоторый многочлен с коэффициентами из поля *К*

$$g(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_{m-1}\lambda + \alpha_m .$$

Рассмотрим квадратную матрицу m-го порядка

$$L = \begin{Vmatrix} 0 & 0 & \dots & 0 & -\alpha_m \\ 1 & 0 & \dots & 0 & -\alpha_{m-1} \\ 0 & 1 & \dots & 0 & -\alpha_{m-2} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & \dots & 1 & -\alpha_1 \end{Vmatrix}. \tag{4.27}$$

Нетрудно проверить, что многочлен $g(\lambda)$ является характеристическим многочленом матрицы L:

$$|\lambda E - L| = \begin{vmatrix} \lambda & 0 & 0 & \dots & 0 & \alpha_m \\ -1 & \lambda & 0 & \dots & 0 & \alpha_{m-1} \\ 0 & -1 & \lambda & \dots & 0 & \alpha_{m-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & -1 & \alpha_1 + \lambda \end{vmatrix} = g(\lambda).$$

С другой стороны, минор элемента α_m в характеристическом определителе равен $\pm$ 1. Поэтому $D_{m-1}(\lambda) = 1$ и

$$i_1(\lambda) = \frac{D_m(\lambda)}{D_{m-1}(\lambda)} = D_m(\lambda) = g(\lambda),\ i_2(\lambda) = \ldots = i_n(\lambda) = 1.$$

Таким образом, матрица L имеет единственный отличный от единицы инвариантный многочлен, равный $g(\lambda)$.

Матрицу L мы будем называть *сопровождающей* матрицей для многочлена $g(\lambda)$.

Пусть дана матрица $A = \| a_{ik} \|^n_1$ с инвариантными многочленами

$$i_1(\lambda), i_2(\lambda), \ldots, i_t(\lambda), i_{t+1}(\lambda) = 1, \ldots, i_n(\lambda) = 1. \tag{4.28}$$

Здесь все многочлены $i_1(\lambda), i_2(\lambda)$, ..., $i_t(\lambda)$ имеют степень выше нулевой, причем каждый из этих многочленов, начиная со второго, является делителем предыдущего. Сопровождающие матрицы для этих многочленов обозначим через $L_1, L_2, \ldots, L_t$.

Тогда квазидиагональная матрица n-го порядка

$$L_I = \{L_1, L_2, \ldots, L_t\} \tag{4.29}$$

имеет своими инвариантными многочленами многочлены (4.28) (см. теорему 2). Поскольку матрицы A и L_I, имеют одни и те же инвариантные многочлены, они подобны, т.е. существует всегда такая неособенная матрица U ($| U | \neq 0$), что

$$A = U L_I U^{-1} \tag{4.30}$$

Матрица L_I называется *первой естественной нормальной формой* для матрицы *A*. Эта нормальная форма характеризуется: 1) квазидиагональным видом (4.29), 2) специальной структурой диагональных клеток (4.27) и 3) дополнительным условием: в ряде характеристических многочленов диагональных клеток каждый многочлен, начиная со второго, является делителем предыдущего. (Из условий 1), 2), 3) автоматически следует, что характеристические многочлены диагональных клеток в L_I являются инвариантными многочленами матрицы L_I и, следовательно, матрицы *A*).

2. Обозначим теперь через

$$\chi_1(\lambda), \chi_2(\lambda), \ldots, \chi_\mu(\lambda) \qquad (4.31)$$

элементарные делители матрицы $A= \|a_{ik}\|^n_1$ в числовом поле *K*. Соответствующие сопровождающие матрицы обозначим через

$$L^{(1)}, L^{(2)}, \ldots, L^{(u)}$$

Поскольку $\chi_j(\lambda)$ - единственный элементарный делитель матрицы $L^{(j)}$ (j=1, 2, ..., u) ($\chi_j(\lambda)$— единственный инвариантный многочлен матрицы $L^{(j)}$ и в тот же время $\chi_j(\lambda)$ есть степень неприводимого в поле *K* многочлена), то согласно теореме 3 квазидиагональная матрица

$$L_{II}= \{ L^{(1)}, L^{(2)}, \ldots, L^{(u)}\} \qquad (4.32)$$

имеет своими элементарными делителями многочлены (4.31).

Матрицы *A* и L_{II} имеют одни и те же элементарные делители в поле *K*. Поэтому эти матрицы подобны, т.е. существует всегда такая неособенная матрица V ($|V| \neq 0$), что

$$A = V L_{II} V^{-1} \qquad (4.33)$$

Матрица L_{II} называется *второй естественной нормальной формой* для матрицы *A*. Эта нормальная форма характеризуется: 1) квазидиагональным видом (4.32), 2) специальной структурой диагональных клеток (4.27) и 3) дополнительным условием: характеристический многочлен каждой диагональной клетки представляет собой степень неприводимого в поле *K* многочлена.

Замечание. Элементарные делители матрицы *A* на отличие от инвариантных многочленов существенно связаны с данным числовым полем *K*. Если мы вместо исходного числового поля *K* возьмем другое числовое поле (которому также принадлежат элементы данной матрицы *A),* то элементарные делители могут измениться. Вместе с элементарными делителями изменится и вторая естественная нормальная форма матрицы.

Так, например, пусть дана матрица $A=\|a_{ik}\|^n_1$ с вещественными элементами. Характеристический многочлен этой матрицы будет иметь вещественные коэффициенты. В то же время этот многочлен может иметь комплексные корни. Если *K* — поле вещественных чисел, то среди элементарных делителей могут быть и степени квадратных неприводимих трехчленов с вещественными коэффициентами. Если *К* — поле комплексных чисел, то каждый элементарный делитель имеет вид $(\lambda—\lambda_0)^{p.}$

3. Предположим теперь, что числовое поле *К* содержит не только элементы матрицы *A,* но и все характеристические числа этой матрицы (это всегда имеет место для любой матрицы *A,* если *К* — поле комплексных чисел). Тогда элементарные делители матрицы *A* имеют вид (среди чисел $\lambda_{1,}$ λ_2, ..., λ_u могут быть и равные между собой)

$$(\lambda-\lambda_1)^{p_1},\ (\lambda-\lambda_2)^{p_2},\ \ldots,\ (\lambda-\lambda_u)^{p_u}\qquad (p_1+p_2+\ldots+p_u=n). \tag{4.34}$$

Рассмотрим один из таких элементарных делителей

$$(\lambda—\lambda_0)^p$$

и поставим ему в соответствие следующую матрицу порядка *p:*

$$\begin{Vmatrix} \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_0 \end{Vmatrix} = \lambda_0 E^{(p)} + H^{(p)}. \tag{4.35}$$

Нетрудно проверить, что эта матрица имеет только один элементарный делитель $(\lambda—\lambda_0)^{p.}$ Матрицу (4.35) мы будем называть *жордановой клеткой,* соответствующей элементарному делителю $(\lambda—\lambda_0)^{p.}$

Жордановы клетки, которые соответствующие элементарным делителям (4.34), обозначим через

$$J_1,\ J_2,\ldots,J_{u.}$$

Тогда квазидиагональная матрица

$$J = \{\ J_1,\ J_2,\ldots,J_u\ \}$$

имеет своими элементарными делителями степени (4.34).

Матрицу *J* можно еще записать так:

$$J=\{\lambda_1E_1+H_1,\lambda_2E_2+H_2,\ldots,\lambda_uE_u+H_u\},$$

где

$$E_k=E^{(p_k)},\quad H_k=H^{(p_k)}\qquad (k=1,2,\ldots,u).$$

Поскольку матрицы *A* и *J* имеют одни и те же элементарные делители, они подобны между собой, т.е. существует такая неособенная матрица $T(|T|\neq 0)$, что

$$A = TJT^{-1} = T\{\lambda_1 E_1 + H_1,\ \lambda_2 E_2 + H_2,\ \ldots,\ \lambda_u E_u + H_u\}\, T^{-1}. \tag{4.36}$$

Матрица J называется *жордановой нормальной формой* или просто *жордановой формой* матрицы A. Жорданова форма характеризуется квазидиагональным видом и специальной структурой (4.35) диагональных клеток.

На нижеследующей схеме выписана жорданова матрица J при элементарных делителях $(\lambda - \lambda_1)^2$, $(\lambda - \lambda_2)^3$, $\lambda - \lambda_3$, $(\lambda - \lambda_4)^2$:

$$J = \left\| \begin{array}{cc|ccc|c|cc} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \end{array} \right\|. \tag{4.37}$$

Если все элементарные делители матрицы A — первой степени (и только в этом случае), жорданова форма является диагональной матрицей, и в этом, случая мы имеем:

$$A = T\{\lambda_1; \lambda_2, \ldots, \lambda_n\}\, T^{1} \tag{4.38}$$

Таким образом, *матрица A имеет простую структуру в том и только в том случае, когда все ее элементарные делители имеют первую степень* (часто вместо «элементарные делители первой степени» говорят «линейные элементарные делители» или «простые элементарные делители»).

Иногда вместо жорданової клетки (4.36) рассматривают *«нижнюю» жорданову клетку p*-го порядка

$$\begin{Vmatrix} \lambda_0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 1 & \lambda_0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & \cdot & \cdot & & & \cdot & \cdot \\ \cdot & & \cdot & \cdot & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \lambda_0 & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 1 & \lambda_0 \end{Vmatrix} = \lambda_0 E^{(p)} + F^{(p)}.$$

Эта матрица также имеет только один элементарный делитель $(\lambda—\lambda_0)^p$. Элементарным делителям (4.34) соответствует *«нижняя» жорданова матрица* (в отличие от нижней жордановой матрицы $J_{(1)}$ матрицу J иногда называют *верхней жордановой матрицей*)

$$J_{(1)} = \{\lambda_1 E_1 + F_1, \lambda_2 E_2 + F_2, \ldots, \lambda_u E_u + F_u\}$$
$$(E_k = E^{(p_k)}, F_k = F^{(p_k)}; k = 1, 2, \ldots, u).$$

Произвольная матрица A, имеющая элементарные делители (4.34), всегда подобна матрице $J_{(1)}$, т.е. существует такая неособенная матрица $T_1(|T_1|\neq 0)$, что

$$A = T_1 J_{(1)} T_1^{-1} = T_1 \{\lambda_1 E_1 + F_1, \lambda_2 E_2 + F_2, \ldots, \lambda_u E_u + F_u\} T_1^{-1}. \tag{4.39}$$

Заметим еще, что если $\lambda_0\neq 0$, то каждая из матриц

$$\lambda_0 (E^{(p)} + H^{(p)}), \quad \lambda_0 (E^{(p)} + F^{(p)})$$

имеет только один элементарный делитель: $(\lambda—\lambda_0)^p$. Поэтому для неособенной матрицы A, имеющей элементарные делители (4.34), наряду с (4.36) и (4.39) имеют место представления

$$A = T_2 \{\lambda_1 (E_1 + H_1), \lambda_2 (E_2 + H_2), \ldots, \lambda_u (E_u + H_u)\} T_2^{-1}, \tag{4.40}$$

$$A = T_3 \{\lambda_1 (E_1 + F_1), \lambda_2 (E_2 + F_2), \ldots, \lambda_u (E_u + F_u)\} T_3^{-1}. \tag{4.41}$$

4.5. Элементарные делители матрицы *f(A)*

1. В этом разделе рассмотрим следующую задачу:

Даны элементарные делители (в поле комплексных чисел) матрицы $A=||a_{ik}||^n_1$ *и дана функция* $f(\lambda)$, *определенная на спектре матрицы A. Требуется определить элементарные делители (в поле комплексных чисел) матрицы f(A).*

Обозначим через

$$(\lambda - \lambda_1)^{p_1}, \quad (\lambda - \lambda_2)^{p_2}, \ldots, (\lambda - \lambda_u)^{p_u},$$

элементарные делители матрицы A (среди чисел $\lambda_1, \lambda_2, ..., \lambda_u$ могут быть и равные между собой). Тогда матрица A подобна жордановой матрицы J

$$A = TJT^{-1}$$

и, следовательно (см. 2°),

$$f(A) = Tf(J)\,T^{-1}$$

При этом

$$J = \{J_1, J_2, \ldots, J_u\}, \quad J_i = \lambda_i E^{(p_i)} + H^{(p_i)} \qquad (i = 1, 2, \ldots, u),$$

а

$$f(J) = \{f(J_1), f(J_2), \ldots, f(J_u)\}, \tag{4.42}$$

где

$$f(J_i) = \begin{Vmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \ldots & \frac{f^{(p_i-1)}(\lambda_i)}{(p_i-1)!} \\ 0 & f(\lambda_i) & \ddots & \vdots \\ \vdots & & \ddots & \frac{f'(\lambda_i)}{1!} \\ 0 & 0 & \ldots & f(\lambda_i) \end{Vmatrix}. \quad (i = 1, \ldots, u). \tag{4.43}$$

Поскольку подобные матрицы $f(A)$ и $f(J)$ имеют одни и те же элементарные делители, то в дальнейшем вместо матрицы $f(A)$ мы будем рассматривать матрицу $f(J)$.

2. Определим сначала дефект d матрицы $f(A)$ или, что то же самое, матрицы $f(J)$ ($d=n-r$, где r — ранг матрицы $f(A)$). Дефект квазидиагональной матрицы равен сумме дефектов отдельных диагональных клеток, а дефект матрицы $f(J_1)$ [см. (4.43)] равен наименьшему из чисел k_i и p_i, где k_i — кратность λ_i как корня $f(\lambda)$ (в общем случае, когда $f(\lambda)$ — не многочлен, под кратностью корня λ_i функции $f(\lambda)$ понимают целое число k_i, определяемое из условий (4.44). k_i может и равняться нулю; в этом случае $f(\lambda_i)\neq 0$), поскольку

$$f(\lambda_i) = f'(\lambda_i) = \ldots = f^{(k_i-1)}(\lambda_i) = 0, \quad f^{(k_i)}(\lambda_i) \neq 0 \qquad (i = 1, 2, \ldots, u). \tag{4.44}$$

Мы пришли к теореме:

Теорема 6. *Дефект матрицы f(A), где матрица A имеет элементарные делители*

$$(\lambda-\lambda_1)^{p_1},\quad (\lambda-\lambda_2)^{p_2},\ \ldots,(\lambda-\lambda_u)^{p_u},$$

(4.45)

определяется формулой

$$d=\sum_{i=1}^{u}\min\,(k_i, p_i);$$

(4.46).

здесь k_i *— кратность* λ_i *как корня* $f(\lambda)$ $(i = 1,2,..., u)$.

В качестве приложения доказанной теоремы определим все элементарные делители произвольной матрицы $A=\|a_{ik}\|^n_1$, соответствующие характеристическому числу λ_0:

$$\underbrace{\lambda-\lambda_0,\ \ldots,\lambda-\lambda_0}_{g_1};\quad \underbrace{(\lambda-\lambda_0)^2,\ \ldots,(\lambda-\lambda_0)^2}_{g_2};\ \ldots;\quad \underbrace{(\lambda-\lambda_0)^m,\ \ldots,(\lambda-\lambda_0)^m}_{g_m},$$

где $g_i\geq 0$ $(i = 1, 2,..., m — 1)$, $g_m > 0$, в случае, когда даны дефекты

$$d_1, d_2, ...,d_m$$

матриц

$$A - \lambda_0E,\ (A - \lambda_0E)^2,...,(A- \lambda_0E)^m.$$

Для этого заметим, что $(A — \lambda_0E)^j = f_j (A)$, где $f_j (\lambda) = (\lambda — \lambda_0)^j$ $(j = 1, 2,... ..., m)$. Поэтому для определения дефекта матрицы $(A — \lambda_0E)^j$ следует в формуле (4.46) положить $k_i=j$ для элементарных делителей, соответствующих характеристическому числу λ_0, и $k_i=0$ во всех других слагаемых $(j = 1,2, ..., m)$.

Таким образом, получим формулы

$$\left.\begin{array}{l} g_1+\ g_2+\ g_3+\ldots+\ g_m=d_1,\\ g_1+2g_2+2g_3+\ldots+2g_m=d_2,\\ g_1+2g_2+3g_3+\ldots+3g_m=d_3,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ g_1+2g_2+3g_3+\ldots+mg_m=d_m.\end{array}\right\}$$

(4.47)

Отсюда (если задан ряд чисел $d_1, d_2, d_3, ...$, где d_j — дефект степени $(A — \lambda_0E)^j$ $(j = 1, 2, 3, ...)$, то число m — наибольший из показателей степеней элементарных делителей вида $(\lambda — \lambda_0)^v$ — определяется как индекс, при котором $d_{m-1} < d_m = d_{m+1}$)

$$g_j = 2d_j - d_{j-1} - d_{j+1} \qquad (j = 1, 2, ..., m;\ d_0 = 0,\ d_{m+1} = d_m). \qquad (4.48)$$

3. Вернемся к основному вопросу об определении элементарных делителей матрицы $f(A)$. Как мы уже отмечали выше, элементарные делители $f(A)$ совпадают с элементарными делителями $f(J)$, а элементарные делители квазидиагональной матрицы состоят из элементарных делителей диагональных клеток (см. теорему 3). Поэтому вопрос сводится к поиску элементарных делителей матрицы *C*, имеющей правильную треугольную форму:

$$C = \sum_{k=0}^{p-1} a_k H^k = \left\| \begin{matrix} a_0 & a_1 & \cdot & \cdot & \cdot & a_{p-1} \\ 0 & a_0 & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & a_1 \\ 0 & 0 & \cdot & \cdot & \cdot & a_0 \end{matrix} \right\|. \tag{4.49}$$

Рассмотрим отдельно два случая:

1° $a_1 \neq 0$. Характеристический многочлен матрицы *C*, очевидно, равен

$$D_p(\lambda) = (\lambda - a_0)^p.$$

Тогда, поскольку $D_p(\lambda)$ делится на $D_{p-1}(\lambda)$ без остатка, то

$$D_{p-1}(\lambda) = (\lambda - a_0)^g \qquad (g \leq p).$$

Здесь через $D_{p-1}(\lambda)$ мы обозначили наибольший общий делитель миноров $(p — 1)$-го порядка характеристической матрицы

$$\lambda E - C = \left\| \begin{matrix} \lambda - a_0 & -a_1 & \cdot & \cdot & \cdot & -a_{p-1} \\ 0 & \lambda - a_0 & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & -a_1 \\ \overset{+}{0} & 0 & \cdot & \cdot & \cdot & \lambda - a_0 \end{matrix} \right\|.$$

Легко видеть, что минор нулевого элемента, отмеченного значком «+», равен $(-a_1)^{p-1}$ и, следовательно, в нашем случае отличен от нуля. Поэтому здесь $g = 0$. Но тогда из

$$D_p(\lambda) = (\lambda - a_0)^p, \quad D_{p-1}(\lambda) = 1$$

следует, что матрица *C* имеет только один элементарный делитель $(\lambda — a_0)^p$.

2° $a_1 = ... = a_k = 0$, $a_k \neq 0$. В этом случае

$$C = a_0 E + a_k H^k + ... + a_{p-1} H^{p-1} \qquad (H = H^{(p)}).$$

Поэтому при любом целом положительном *j* дефект матрицы

$$(C - a_0 E)^j = a_k^j H^{kj} + ...$$

определится равенством

$$d_j = \begin{cases} kj & \text{при} \quad kj \leq p, \\ p & \text{при} \quad kj > p. \end{cases}$$

Положим

$$p = qk + h \qquad (0 \leq h < k). \tag{4.50}$$

Тогда (в этом случае число q+1 играет роль числа m в формулах (4.47) и (4.48))

$$d_1 = k,\ d_2 = 2\,k,\ ...,\ d_q = qk,\ d_{q+1} = p. \tag{4.51}$$

Поэтому согласно формулам (4.52) имеем:

$$g_1 = \ldots = g_{q-1} = 0, g_q = k-h, g_{q+1} = h...$$

Таким образом, матрица C имеет элементарные делители

$$\underbrace{(\lambda - a_0)^{q+1}, \ldots, (\lambda - a_0)^{q+1}}_{h}, \quad \underbrace{(\lambda - a_0)^{q}, \ldots, (\lambda - a_0)^{q}}_{k-h}, \tag{4.52}$$

где целые числа $q > 0$ и $h \geq 0$ определяются из (4.50).

4. Теперь мы уже имеем возможность выяснить, какие элементарные делители имеет матрица $f(J)$ [см. формулы (4.42) и (4.43)]. Каждому элементарному делителю матрицы A

$$(\lambda - \lambda_0)^p$$

отвечает в матрице $f(J)$ диагональная клетка

$$f(\lambda_0 E + H) = \sum_{i=0}^{p-1} \frac{f^{(i)}(\lambda_0)}{i!} H^i = \left\| \begin{matrix} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \cdots & \frac{f^{(p-1)}(\lambda_0)}{(p-1)!} \\ 0 & f(\lambda_0) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \frac{f'(\lambda_0)}{1!} \\ 0 & 0 & \cdots & f(\lambda_0) \end{matrix} \right\|. \tag{4.53}$$

Очевидно, вопросы сводится к нахождению элементарных делителей клеток вида (4.53). Но матрица (4.53) имеет правильную треугольную форму (4.49), причем здесь

$$a_0 = f(\lambda_0),\ a_1 = f'(\lambda_0),\ a_2 = \frac{f''(\lambda_0)}{2!},\ \ldots$$

Таким образом, приходим к теореме:

Теорема 7. *Элементарные делители матрицы $f(A)$ получаются из элементарных делителей матрицы A следующим образом: элементарному делителю*

$$(\lambda - \lambda_0)^p \tag{4.54}$$

матрицы А при p=1 *или при* p>1 *и* $f'(\lambda_0)\neq 0$ *отвечает один элементарный делитель*

$$(\lambda - f(\lambda_0))^p \tag{4.55}$$

матрицы f(A); при p>1, $f'(\lambda_0)= \ldots =f^{(k-1)}(\lambda_0)$, $f^{(k)}(\lambda_0)\neq 0$ $(k<p)$ *элементарному делителю* (4.54) *матрицы А отвечают следующие элементарные делители матрицы f(A):*

$$\underbrace{(\lambda - f(\lambda_0))^{q+1}, \ldots, (\lambda - f(\lambda_0))^{q+1}}_{h}, \quad \underbrace{(\lambda - f(\lambda_0))^{q}, \ldots, (\lambda - f(\lambda_0))^{q}}_{k-h}, \tag{4.56}$$

где

$$p = qk + h, \quad 0 \leq q, \quad 0 \leq h < k;$$

наконец, при

$$p > 1, \; f'(\lambda_0) = f''(\lambda_0) = \ldots = f^{(p-1)}(\lambda_0) = 0$$

элементарному делителю (4.54) *отвечают р элементарных делителей первой степени матрицы f(A)* ((4.55) получается из (4.56), если положить k = 1; (4.57) выходит из (4.56), если положить $k = p$ или $k > p$):

$$\lambda - f(\lambda_0), \ldots, \lambda - f(\lambda_0). \tag{4.57}$$

Отметим следующие частные положения, которые содержатся в этой теореме.

1. *Если* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *— характеристические числа матрицы А, то* $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ *суть характеристические числа матрицы f(A)* (как в первом, так и во втором рядах чисел каждое характеристическое число повторяется столько раз, какова его кратность как корня характеристического уравнения).

2. *Если производная* $f'(\lambda)$ *не равна нулю на спектре матрицы А* (т.е. $f'(\lambda_i)\neq 0$ для тех λ_i, которые являются кратными корнями минимального многочлена), *то при переходе от матрицы А к матрице f(A) элементарные делители не «расщепляются», т.е. если матрица А имеет элементарные делители*

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \ldots, (\lambda - \lambda_n)^{p_n},$$

это матрица f(A) имеет элементарные делители

$$(\lambda - f(\lambda_1))^{p_1}, (\lambda - f(\lambda_2))^{p_2}, \ldots, (\lambda - f(\lambda_n))^{p_n}.$$

4.6. Общий метод построения преобразующей матрицы

Во многих вопросах теории матриц и их приложениях достаточно знать только нормальную форму, к которой приводится данная матрица $A = \| a_{ik} \|^n_1$ преобразованием подобия. Нормальная форма вполне определяется инвариантными многочленами характеристической матрицы $\lambda E — A$. Для нахождения последних можно воспользоваться определяющими формулами [см. (4.10)] или приведением характеристической матрицы $\lambda E — A$ при помощи элементарных преобразований к канонической диагональной форме.

В некоторых же вопросах необходимо знать не только нормальную форму $\tilde{A}$ данной матрицы A, но и преобразующую неособенную матрицу T.

Непосредственный способ определения матрицы T заключается в следующем.

Равенство

$$A = T\tilde{A}T^{-1}$$

переписывается так:

$$AT - T\tilde{A} = 0.$$

Это матричное уравнение относительно T равносильно системе n^2 линейных однородных уравнений относительно n^2 неизвестных коэффициентов матрицы T. Определение преобразующей матрицы сводится к решению этой системы из n^2 уравнений. При этом из множества решений необходимо выбрать такое решение, для которого $|T| \neq 0$. Существование такого решения обеспечено тем, что матрицы A и $\tilde{A}$ имеют одни и те же инвариантные многочлены (из этого факта следует подобие матриц A и $\tilde{A}$).

Заметим, что в то время как нормальная форма определяется однозначно заданием данной матрицы A (это утверждение безоговорочно справедливо в отношении первой естественной нормальной формы. Что же касается второй нормальной формы или жордановой нормальной формы, то они определяются однозначно с точностью до порядка диагональных клеток), для преобразующей матрицы T мы всегда имеем бесчисленное множество значений, охватываемых формулой

$$T = UT_1, \tag{4.58}$$

где T_1 — одна из преобразующих матриц, а U — произвольная матрица, перестановочная с A (формулу (4.58) можно заменить формулой $T = T_1V$, где V — произвольная матрица, перестановочна с $\tilde{A}$).

Предложенный выше способ определения преобразующей матрицы T очень прост по своей идее, но пригоден при использовании вычислительной техники, так как связан с очень большими объемами вычислений (так, уже при $n = 4$ он требует решения системы из 16 линейных уравнений).

Рассмотрим более эффективный метод построения преобразующей матрицы T. Этот метод опирается в добавление к теореме 5. Согласно этому добавлению в качестве преобразующей матрицы можно взять матрицу

$$T=Q(\tilde{A}), \tag{4.59}$$

при этом

$$\lambda E - \tilde{A} = P(\lambda)(\lambda E - A)Q(\lambda).$$

Последнее равенство выражает собой эквивалентность характеристических матриц $\lambda E — A$ и $\lambda E — \tilde{A}$. Здесь $P(\lambda)$ и $Q(\lambda)$ — многочленные матрицы с постоянными отличными от нуля определителями.

Для конкретного нахождения матрицы $Q(\lambda)$ мы приводим к канонической диагональной форме каждую из λ-матриц $\lambda E\text{-}A$ и $\lambda E\text{-}\tilde{A}$ при помощи соответствующих элементарных преобразований (здесь существенно лишь то, что обе λ-матрицы $\lambda E — A$ и $\lambda E — \tilde{A}$ приводятся к одному и тому же виду. Мы избрали канонический диагональный вид, поскольку существует алгоритм, который обеспечивает приведение к такому виду):

$$\{i_n(\lambda), i_{n-1}(\lambda), \ldots, i_1(\lambda)\} = P_1(\lambda)(\lambda E - A)\,Q_1(\lambda), \tag{4.60}$$

$$\{i_n(\lambda), i_{n-1}(\lambda), \ldots, i_1(\lambda)\} = P_2(\lambda)(\lambda E - \tilde{A})\,Q_2(\lambda), \tag{4.61}$$

где

$$Q_1(\lambda) = T_1T_2 \ldots T_{p_1}, \quad Q_2(\lambda) = T_1^*T_2^* \ldots T_{p_2}^*, \tag{4.62}$$

а $T_1,\ldots,T_{p_1},T^*_1,\ldots,T^*_{p_2}$ — элементарные матрицы, соответствующие элементарным операциям над столбцами λ-матриц $\lambda E — A$ и $\lambda E — \tilde{A}$. Из (4.60), (4.61) и 4.(62) следует:

$$\lambda E - \tilde{A} = P(\lambda)(\lambda E - A)Q(\lambda),$$

где

$$Q(\lambda) = Q_1(\lambda)\, Q_2^{-1}(\lambda) = T_1 T_2 \ldots T_{p_1} T^{*-1}_{p_2} T^{*-1}_{p_2-1} \ldots T^{*-1}_1. \qquad (4.63)$$

Матрицу $Q(\lambda)$ вычисляем, применив последовательно к столбцам единичной матрицы E элементарные операции с матрицами $T_1, ..., T_{p_1}, T^{*-1}{}_1, ..., T^{*-1}{}_1$. После этого [согласно формуле (4.59)] заменяем в $Q(\lambda)$ аргумент λ матрицей $\mathcal{A}$.

Пример.

$$\begin{Vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{Vmatrix}.$$

Введем символические обозначения для левых и правых элементарных операций и соответствующих матриц:

$$S' = \{(c)\, i\}, \quad S'' = \{i + (b(\lambda))\, j\}, \quad S''' = \{ij\},$$
$$T' = [(c)\, i], \quad T'' = [i + (b(\lambda))\, j], \quad T''' = [ij].$$

Читатель легко проверит, что характеристическая матрица

$$\lambda E - A = \begin{Vmatrix} \lambda - 1 & 0 & -1 \\ 0 & \lambda - 1 & 1 \\ 1 & 1 & \lambda - 1 \end{Vmatrix}$$

приводится к каноническому диагональному виду

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda - 1)^3 \end{Vmatrix}$$

с помощью следующих последовательно выполненных элементарных операций:

$$[1 + (\lambda - 1)\, 3], \quad \{2 + 1\}, \quad \{3 + (\lambda - 1)\, 1\}, \quad \{(-1)\, 1\}, \quad [1 - 2],$$
$$[1 - (\lambda^2 - 2\lambda + 1)\, 2], \quad \{2 - (\lambda - 1)\, 1\}, \quad \{(-1)\, 2\}, \quad [13], \quad \{23\}. \qquad (4.64)$$

Из канонического диагонального вида матрицы $\lambda E - A$ усматриваем, что матрица A имеет только один элементарный делитель $(\lambda - 1)^3$. Поэтому соответствующей жордановой формой будет матрица

$$J = \begin{Vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{Vmatrix}.$$

Нетрудно видеть, что характеристическая матрица λE — J приводится к тому же каноническому диагональному виду при помощи элементарных операций

$$\{3+(\lambda-1)\,2\},\quad \{3+(\lambda^2-2\lambda+1)^3\,1\},\quad [2+(\lambda-1)\,3],\quad [1+(\lambda-1)\,2],$$
$$\{(-1)\,1\}\,\{(-1)\,2\},\,[13],\,\{12\}.$$

(4.65)

Выбрасывая из (4.64) и (4.65) левые элементарные операции, которые обозначены символом {…}, мы в соответствии с формулами (4.62), (4.63) получим

$$Q(\lambda)=Q_1(\lambda)\,Q_2^{-1}(\lambda)=$$
$$=[1+(\lambda-1)\,3]\,[1-2]\,[1-(\lambda^2-2\lambda+1)\,2]\,[13]\,[13]\,[1-(\lambda-1)\,2]\,[2-$$
$$-(\lambda-1)\,3]=[1+(\lambda-1)\,3]\,[1-(\lambda^2-\lambda+1)\,2]\,[2-(\lambda-1)\,3].$$

Применим к единичной матрице последовательно эти правые элементарные операции:

$$E=\begin{Vmatrix}1&0&0\\0&1&0\\0&0&1\end{Vmatrix}\to\begin{Vmatrix}1&0&0\\0&1&0\\\lambda-1&0&1\end{Vmatrix}\to\begin{Vmatrix}1&0&0\\-\lambda^2+\lambda-1&1&0\\\lambda-1&0&1\end{Vmatrix}\to$$
$$\to\begin{Vmatrix}1&0&0\\-\lambda^2+\lambda-1&1&0\\\lambda-1&-\lambda+1&1\end{Vmatrix}=Q(\lambda).$$

Таким образом,

$$Q(\lambda)=\begin{Vmatrix}0&0&0\\-1&0&0\\0&0&0\end{Vmatrix}\lambda^2+\begin{Vmatrix}0&0&0\\1&0&0\\1&-1&0\end{Vmatrix}\lambda+\begin{Vmatrix}1&0&0\\-1&1&0\\-1&1&1\end{Vmatrix}.$$

Замечая, что

$$J^2=\begin{Vmatrix}1&2&1\\0&1&2\\0&0&1\end{Vmatrix},$$

Находим

$$T = Q(J) = \begin{Vmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix} \begin{Vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{Vmatrix} + \begin{Vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 0 \end{Vmatrix} \begin{Vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{Vmatrix} +$$

$$+ \begin{Vmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{Vmatrix}.$$

Проверка.

$$AT = \begin{Vmatrix} 1 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{Vmatrix}, \quad TJ = \begin{Vmatrix} 1 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{Vmatrix},$$

$$|T| = \begin{vmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{vmatrix} = 1.$$

Следовательно, $AT=TJ(|T|\neq 0)$, т.е. $A = TJT^{-1}$

4.7. Частный метод построения преобразующей матрицы

1. Изложим частный метод построения преобразующей матрицы, который часто приводит к меньшим вычислениям, чем метод предыдущего раздела. Однако этот метод может быть применим лишь тогда, когда нормальная форма жорданова и известны элементарные делители

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \ldots \qquad (4.66)$$

данной матрицы A.

Пусть $A = TJT^{-1}$, где

$$J = \{\lambda_1 E^{(p_1)} + H^{(p_1)},\quad \lambda_2 E^{(p_2)} + H^{(p_2)}, \ldots\} = \left\| \begin{array}{cc} \overbrace{\begin{matrix} \lambda_1 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & 1 \\ 0 & \ldots & & \lambda_1 \end{matrix}}^{p_1} & \\ & \overbrace{\begin{matrix} \lambda_2 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & 1 \\ 0 & \ldots & & \lambda_2 \end{matrix}}^{p_2} \\ & & \ddots \end{array} \right\|.$$

Тогда, обозначая через t_k k-й столбец матрицы T (k = 1, 2, ..., n), матричное равенство

$$AT = TJ$$

заменим эквивалентной системой равенств

$$At_1 = \lambda_1 t_1, At_2 = \lambda_1 t_2 + t_1, \ldots, At_{p_1} = \lambda_1 t_{p_1} + t_{p_1-1}, \tag{4.67}$$

$$At_{p_1+1} = \lambda_2 t_{p_1+1}, At_{p_1+2} = \lambda_2 t_{p_1+2} + t_{p_1+1}, \ldots, At_{p_1+p_2} = \lambda_2 t_{p_1+p_2} + t_{p_1+p_2-1},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{4.68}$$

которую перепишем еще так:

$$(A - \lambda_1 E)\, t_1 = 0,\quad (A - \lambda_1 E)\, t_2 = t_1, \ldots, (A - \lambda_1 E)\, t_{p_1} = t_{p_1-1}, \tag{4.69}$$

$$(A - \lambda_2 E)\, t_{p_1+1} = 0,\quad (A - \lambda_2 E)\, t_{p_1+2} = t_{p_1+1}, \ldots, (A - \lambda_2 E)\, t_{p_1+p_2} = t_{p_1+p_2-1},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \tag{4.70}$$

Таким образом, все столбцы матрицы T разбиваются на «жордановы цепочки» столбцов: [t_1, t_2,….., t_{p1}], [t_{p1+1}, t_{p1+2}, ..., t_{p1+p2}] ,…Каждой жордановой клетке в J [или, что то же самое, каждому элементарному делителю (4.66)] соответствует своя жорданова цепочка столбцов. Каждая жорданова цепочка столбцов характеризуется системой уравнений типа (4.67), (4.68) и т.п.

Нахождение преобразующей матрицы T сводится к поиску жордановых цепочек, которые в совокупности давали бы n линейно независимых столбцов.

Покажем, что эти жордановы цепочки столбцов можно определить при помощи приведенной матрицы $C(\lambda)$.

Для матрицы $C(\lambda)$ имеем тождество

$$(\lambda E — A)\ C(\lambda)= \psi(\lambda)\ E, \qquad (4.71)$$

где $\psi(\lambda)$ — минимальный многочлен матрицы A.

Пусть

$$\psi(\lambda)=(\lambda - \lambda_0)^m \chi(\lambda) \qquad (\chi(\lambda_0)\neq 0).$$

Продифференцируем почленно последовательно m—1 раз тождество (4.71):

$$\left.\begin{array}{c} (\lambda E - A)\, C'(\lambda) + C(\lambda) = \psi'(\lambda)\, E, \\ (\lambda E - A)\, C''(\lambda) + 2C'(\lambda) = \psi''(\lambda)\, E, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ (\lambda E - A)\, C^{(m-1)}(\lambda) + (m-1)\, C^{(m-2)}(\lambda) = \psi^{(m-1)}(\lambda)\, E. \end{array}\right\} \qquad (4.72)$$

Подставляя λ_0 вместо λ в (4.71), (4.72) и замечая, что правые части при этом обращаются в нуль, получим:

$$(A - \lambda_0 E)\, C = 0,\ (A - \lambda_0 E)\, D = C,\ (A - \lambda_0 E)\, F = D, \dots, (A - \lambda_0 E)\, K = G, \qquad (4.73)$$

где

$$\left.\begin{array}{c} C = C(\lambda_0), \quad D = \frac{1}{1!} C'(\lambda_0), \quad F = \frac{1}{2!} C''(\lambda_0), \dots, G = \frac{1}{(m-2)!}\, C^{(m-2)}(\lambda_0), \\ K = \frac{1}{(m-1)!}\, C^{(m-1)}(\lambda_0). \end{array}\right\} \qquad (4.74)$$

Заменим в равенствах (4.73) матрицы (4.74) их k-мы столбцами (k=1,2,..., n). Получим:

$$(A - \lambda_0 E)\, C_k = 0, \quad (A - \lambda_0 E)\, D_k = C_k, \dots, (A - \lambda_0 E)\, K_k = G_k$$
$$(k = 1, 2, \dots, n). \qquad (4.75)$$

Поскольку $C= C(\lambda_0) \neq 0$ (из $C(\lambda_0) = 0$ следовало бы, что все элементы $C(\lambda)$ имеют общий делитель степени выше нулевой, что противоречит определению $C(\lambda_0)$), можно выбрать такое k ($\leq n$), что

$$C_k \neq 0. \qquad (4.76)$$

Тогда m столбцов

$$C_k, D_k, F_k, \dots, G_k, K_k \qquad (4.77)$$

линейно независимы. В самом деле, пусть

$$\gamma C_k + \delta D_k + \dots + \varkappa K_k = 0. \qquad (4.78)$$

Поножая обе части (4.78) последовательно на $A — \lambda_0 E, \dots, (A — \lambda_0 E)^{m-1}$, получим:

$$\delta C_k + \dots + \varkappa G_k = 0, \dots, \varkappa C_k = 0. \qquad (4.79)$$

Из (4.78) и (4.79) в силу (4.76) находим:

$$\gamma = \delta = \ldots = \varkappa = 0.$$

Поскольку линейно независимые столбцы (4.77) удовлетворяют системе уравнений (4.75), они образуют жорданову цепочку векторов, отвечающую элементарному делителю $(\lambda - \lambda_{0)})^m$ [ср. (4.75) с (4.69)].

Если при некотором k $C_k = 0$, но $D_k \neq 0$, то столбцы D_k, ..., G_k, K_k образуют жорданову цепочку из $m - 1$ векторов и т.д.

2. Покажем сначала, как построить преобразующую матрицу T в том случае, когда матрица A имеет попарно взаимно простые элементарные делители:

$$(\lambda - \lambda_1)^{m_1},\ (\lambda - \lambda_2)^{m_2}, \ldots, (\lambda - \lambda_s)^{m_s} \qquad (\lambda_i \neq \lambda_j \text{ при } i \neq j;\quad i,\ j = 1, 2, \ldots, s).$$

Элементарному делителю $(\lambda - \lambda_j)^{mj}$ ставим в соответствие жорданову цепочку столбцов $C^{(j)}$,$D^{(j)}$, ..., $G^{(j)}$, $K^{(j)}$ построенную по указанному выше способу. Тогда

$$(A - \lambda_j E)\, C^{(j)} = 0,\ (A - \lambda_j E)\, D^{(j)} = C^{(j)}, \ldots, (A - \lambda_j E)\, K^{(j)} = G^{(j)}. \tag{4.80}$$

Давая j значение 1,2, …...,.s, получим s жордановых цепочек, которые содержат в совокупности n столбцов. Эти столбцы линейно независимы.

Действительно, пусть

$$\sum_{j=1}^{s} [\gamma_j C^{(j)} + \delta_j D^{(j)} + \ldots + \varkappa_j K^{(j)}] = 0. \tag{4.81}$$

Умножим обе части равенства (4.81) слева на произведение

$$(A - \lambda_1 E)^{m_1} \ldots (A - \lambda_{j-1} E)^{m_{j-1}}\, (A - \lambda_j E)^{m_j - 1}\quad (A - \lambda_{j+1} E)^{m_{j+1}} \ldots (A - \lambda_s E)^{m_s}. \tag{4.82}$$

Получим:

$$\varkappa_j = 0.$$

Заменяя в (4.82) $m_j - 1$ последовательно на $m_j - 2$, $m_j - 3$, ..., найдем:

$$\gamma_j = \delta_j = \ldots = \varkappa_j = 0 \quad (j = 1, 2, \ldots, s),$$

что и требовалось доказать.

Матрицу T определим формулой

$$T = (C^{(1)}, D^{(1)}, \ldots, K^{(1)};\quad C^{(2)}, D^{(2)}, \ldots, K^{(2)}; \ldots;\quad C^{(s)}, D^{(s)}, \ldots, K^{(s)}). \tag{4.83}$$

Пример.

$$
\begin{array}{c}
\begin{matrix} \mathit{8} & \mathit{3} & \mathit{-10} & \mathit{-3} \end{matrix} \\
A = \begin{Vmatrix} 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{Vmatrix};
\end{array}
$$

$$
\begin{array}{cccc}
\cdots & \cdots & \cdots & \cdots \\
3 & 2 & 2 & 1 \\
\cdots & \cdots & \cdots & \cdots \\
\mathbf{1} & \mathbf{4} & \mathbf{0} & \mathbf{2}
\end{array}
$$

$$\psi(\lambda) = \Delta(\lambda) = (\lambda - 1)^2 (\lambda + 1)^2 = \lambda^4 - 2\lambda^2 + 1,$$

элементарные делители:

$$(\lambda - 1)^2, (\lambda + 1)^2,$$

$$\Psi(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu^3 + \lambda\mu^2 + (\lambda^2 - 2)\mu + \lambda^3 - 2\lambda,$$

$$C(\lambda) = \Psi(\lambda E, A) = A^3 + \lambda A^2 + (\lambda^2 - 2) A + (\lambda^3 - 2\lambda) E.$$

Составим первый столбец $C_1(\lambda)$:

$$C_1(\lambda) = [A^3]_1 + \lambda [A^2]_1 + (\lambda^2 - 2) A_1 + (\lambda^3 - 2\lambda) E_1.$$

Для вычисления первого столбца матрицы A^2 умножим все строки матрицы A на первый столбец матрицы A. Получим (столбец, на который умножаем строки, подписываем под строками матрицы A. Курсивом набранны элементы контрольной суммарной строки): $[A^2]_1 = (1, 4, 0, 2)$. Умножая на этот столбец все строки матрицы A, найдем $[A^3]_1 = (3, 6, 2, 3)$. Поэтому

$$C_1(\lambda) = \begin{Vmatrix} 3 \\ 6 \\ 2 \\ 3 \end{Vmatrix} + \lambda \begin{Vmatrix} 1 \\ 4 \\ 0 \\ 2 \end{Vmatrix} + (\lambda^2 - 2) \begin{Vmatrix} 3 \\ 2 \\ 2 \\ 1 \end{Vmatrix} + (\lambda^3 - 2\lambda) \begin{Vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{Vmatrix} = \begin{Vmatrix} \lambda^3 + 3\lambda^2 - \lambda - 3 \\ 2\lambda^2 + 4\lambda + 2 \\ 2\lambda^2 - 2 \\ \lambda^2 + 2\lambda + 1 \end{Vmatrix}.$$

Отсюда $C_1(1) = (0, 8, 0, 4)$ и $C'_1(1) = (8, 8, 4, 4)$. Поскольку $C_1(-1) = (0, 0, 0, 0)$, то переходим ко вторым столбцам и, действуя аналогично предыдущему, находим: $C_2(-1) = (-4, 0, -4, 0)$ и $C'_2(-1) = (4, -4, 4, -4)$. Составляем матрицу:

$$(C_1(1), C'_1(1); C_2(-1), C'_2(-1)) = \begin{Vmatrix} 0 & 8 & -4 & 4 \\ 8 & 8 & 0 & -4 \\ 0 & 4 & -4 & 4 \\ 4 & 4 & 0 & -4 \end{Vmatrix}.$$

Сокращаем (при умножении всех столбцов жордановоц цепочки на число $c\neq 0$ цепочка остается жордановой) первые два столбца на 4 и вторые два столбца на —4.

$$T=\begin{Vmatrix} 0 & 2 & 1 & -1 \\ 2 & 2 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{Vmatrix}.$$

Предлагаем читателю проверить, что

$$AT=T\cdot\begin{Vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{Vmatrix}.$$

3. Переходя к общему случаю, будем разыскивать жордановы цепочки векторов, которые отвечают характеристическому числу λ_0, которому соответствуют p элементарных делителей $(\lambda-\lambda_0)^m$, q элементарных делителей $(\lambda-\lambda_0)^{m-1}$, r делителей $(\lambda-\lambda_0)^{m-2}$ и т.д.

Установим предварительно некоторые свойства матриц

$$C=C(\lambda_0),\quad D=C'(\lambda_0),\quad F=\frac{1}{2!}C''(\lambda_0),\ldots,K=\frac{1}{(m-1)!}C^{(m-1)}(\lambda_0).$$

(4.84)

1° *Матрицы* (4.84) *могут быть представлены в виде многочленов от A:*

$$C=h_1(A),\quad D=h_2(A),\ldots,K=h_m(A), \tag{4.85}$$

где

$$h_i(\lambda)=\frac{\psi(\lambda)}{(\lambda-\lambda_0)^i}\qquad (i=1,2,\ldots,m).$$

(4.86)

В самом деле,

$$C(\lambda)=\Psi(\lambda E,A),$$

где

$$\Psi(\lambda,\mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}.$$

Поэтому

$$\frac{1}{k!}C^{(k)}(\lambda_0)=\frac{1}{k!}\Psi^{(k)}(\lambda_0 E,A),$$

(4.87)

где

$$\frac{1}{k!}\Psi^{(k)}(\lambda_0,\mu)=\frac{1}{k!}\left[\frac{\partial^k}{\partial\lambda^k}\psi(\lambda,\mu)\right]_{\lambda=\lambda_0}=$$

$$=\frac{1}{k!}\left[\frac{\partial^k}{\partial\lambda^k}\frac{\psi(\mu)}{\mu-\lambda}\right]_{\lambda=\lambda_0}=\frac{\psi(\mu)}{(\mu-\lambda_0)^{k+1}}=h_{k+1}(\mu). \quad (4.88)$$

Из (4.84), (4.87) и (4.88) следует (4.85).

2° *Матрицы* (4.84) *имеют соответственно ранги*

$$p, 2p+q, 3p+2q+r, \ldots$$

Это свойство матриц (4.84) непосредственно получается из 1° и теоремы 6, если положить ранг равным $n - d$ и воспользоваться формулой (4.46) для дефекта функции от A.

3° *В ряду матриц* (4.84) *столбец каждой матрицы является линейной комбинацией столбцов любой следующей матрицы.*

Возьмем две матрицы $h_i(A)$ и $h_k(A)$ в ряду (4.84) (см. 1°). Пусть $i < k$. Тогда из (4.86) следует:

$$h_i(A)=h_k(A)(A-\lambda_0 E)^{k-i}.$$

Отсюда j-й столбец y_j ($j=1,2,\ldots,n$) матрицы $h_i(A)$ линейно выражается через столбцы $z_1, z_2, \ldots, z_n$ матрицы $h_k(A)$:

$$y_j=\sum_{g=1}^{n}\alpha_g z_g,$$

где $\alpha_1, \alpha_2, \ldots, \alpha_n$ — элементы j -го столбца матрицы $(A-\lambda_0 E)^{k-i}$.

4° *Не меняя основных формул* (4.73), *можно в матрице C любой столбец заменить произвольной линейной комбинацией всех столбцов, сделав соответствующую замену в D, ..., K.*

Теперь перейдем к построению жордановых цепочек столбцов для элементарных делителей

$$\underbrace{(\lambda-\lambda_0)^m,\ldots,(\lambda-\lambda_0)^m}_{p};\ \underbrace{(\lambda-\lambda_0)^{m-1},\ldots,(\lambda-\lambda_0)^{m-1}}_{q};\ \ldots$$

Пользуясь свойствами 2° и 4°, мы матрицу C преобразуем к виду

$$C=(C_1,C_2,\ldots,C_p;0,0,\ldots,0), \quad (4.89)$$

где столбцы $C_1, C_2, \ldots, C_p$ линейно независимы между собой. При этом

$$D=(D_1,D_2,\ldots,D_p;D_{p+1},\ldots,D_n).$$

Согласно 3° для любого i ($1\le i\le p$) столбец C_i есть линейная комбинация столбцов $D_1,D_2,\ldots,D_n$:

$$C_i=\alpha_1 D_1+\ldots+\alpha_p D_p+\alpha_{p+1}D_{p+1}+\ldots+\alpha_n D_n. \quad (4.90)$$

Умножим обе части этого равенства на $A-\lambda_0 E$. Тогда, замечая, что [см. (4.75)]

$(A-\lambda_0 E)C_i=0$ $(u= 1,2, ...,p)$, $(A-\lambda_0 E)D_j=C_j$ $(j= 1, 2, ..., n)$,
получим в силу (4.89):

$$0 = \alpha_1 C_1 + \alpha_2 C_2 + ... + \alpha_p C_p,$$

откуда в (4.90)

$$\alpha_1 = ... = \alpha_p = 0.$$

Поэтому столбцы $C_1, C_2,...,C_p$ представляют собой линейно независимые комбинации столбцов $D_{p+1},\ldots,D_n$, и потому, согласно 4° и 2°, не меняя матрицы C, можно вместо $D_{p+1}, ..., D_{2p}$ взять столбцы $C_1, ..., C_p$, а вместо $D_{2p+q+1}, ..., D_n$ — нули.

Тогда матрица D примет вид

$$D = (D_1, \ldots, D_p;\ C_1, C_2, \ldots, C_p;\ D_{2p+1}, \ldots, D_{2p+q};\ 0, 0, \ldots, 0). \tag{4.91}$$

Таким же образом, сохраняя вид (4.89) и (4.91) для матриц C и D, мы следующую матрицу F представим в виде

$$\left.\begin{array}{l} F = (F_1, \ldots, F_p;\ D_1, \ldots, D_p;\ F_{2p+1}, \ldots, F_{2p+q};\ C_1, \ldots, C_p; \\ D_{2p+1}, \ldots, D_{2p+q}, F_{3p+2q+1}, \ldots, F_{3p+2q+r}, 0, \ldots, 0) \end{array}\right\} \tag{4.92}$$

и т.д.

Формулы (4.75) дадут нам жордановы цепочки

$$\left.\begin{array}{l} \underbrace{\overbrace{(C_1, D_1, \ldots, K_1)}^{m}, \ldots, \overbrace{(C_p, D_p, \ldots K_p)}^{m}}_{p}; \\ \underbrace{\overbrace{(D_{2p+1}, F_{2p+1}, \ldots, K_{2p+1})}^{m-1}, \ldots, \overbrace{(D_{2p+q}, F_{2p+q}, \ldots, K_{2p+q})}^{m-1}}_{q}; \ldots \end{array}\right\} \tag{4.93}$$

Эти жордановы цепочки независимы между собой. Действительно, все столбцы C_i в цепочках (4.93) независимы, так как они образуют p независимых столбцов матрицы C. Все столбцы C_i, D_j в (4.93) независимы, так как они образуют $2p + q$ независимых столбцов в матрице D и т.д.; в конце концов, *все столбцы* в (4.93) *независимы, так как они образуют* $n_0 = mp + (m - 1)\, q + ...$ *независимых столбцов в матрице* K. Число столбцов в (4.93) равно сумме степеней элементарных делителей, соответствующих данному характеристическому числу λ_0.

Пусть матрица $A = \| a_{ik} \|^n_1$ имеет s различных характеристических чисел λ_j [$j = 1,2, ..., s$;

$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}...(\lambda - \lambda_s)^{n_s}$; $\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_s)^{m_s}$].

Для каждого характеристического числа λ_j составим свою систему независимых жордановых цепочек (4.93); число столбцов в этой системе будет равно n_j ($j = 1, 2, ..., s$). Все полученные таким образом цепочки содержат $n = n_1 + n_2 + \ldots + n_s$ столбцов.

Эти n столбцов линейно независимы и составляют одну из искомых преобразующих матриц T.

Доказательство линейной независимости полученных n столбцов проводится следующим образом.

Любая линейная комбинация этих n столбцов может быть представлена в виде

$$\sum_{j=1}^{s} H_j = 0, \tag{4.94}$$

где H_j — линейная комбинация столбцов в жордановых цепочках (4.93), соответствующих характеристическому числу λ_j ($j = 1,2, ...,s$). Но любой столбец в жордановоц цепочке, соответствующей характеристическому числу λ_j, удовлетворяет уравнению

$$(A - \lambda_j E)^{m_j} x = 0.$$

Поэтому

$$(A - \lambda_j E)^{m_j} H_j = 0. \tag{4.95}$$

Возьмем фиксированное число j ($1 \le j \le s$) и построим интерполяционный многочлен Лагранжа-Сильвестра $r(\lambda)$ по следующим значениям на спектре матрицы:

$$r(\lambda_i) = r'(\lambda_i) = \ldots = r^{(m_i - 1)}(\lambda_i) = 0 \text{ при } i \neq j$$

и

$$r(\lambda_j) = 1,\ r'(\lambda_j) = \ldots = r^{(m_j - 1)}(\lambda_j) = 0.$$

Тогда при любом $i \neq j$ $r(\lambda)$ делится на $(\lambda - \lambda_i)^{m_i}$ без остатка; поэтому в силу (4.95)

$$r(A)H_i = 0 \quad (i \neq j). \tag{4.96}$$

Точно так же разность $r(\lambda) - 1$ делится на $(\lambda - \lambda_i)^{m_j}$ без остатка; поэтому

$$r(A)H_j = H_j. \tag{4.97}$$

Помножая обе части (4.94) на $r(A)$, мы согласно (4.96) и (4.97) получим:

$$H_j = 0.$$

Это справедливо для каждого j = 1,2, ..., s. Но H_j есть линейная комбинация независимых столбцов, отвечающих одному и тому же характеристическому числу λ_j (j = 1, 2, ..., s). Поэтому все коэффициенты в линейной комбинации H_j (j= 1, 2, ..., s) и, следовательно, все коэффициенты в (4.94) равны нулю.

Замечание. Укажем на некоторые преобразования над столбцами матрицы T, при которых она остается преобразующей к той же жордановой форме (при том же расположении жордановых диагональных клеток):

I. *Умножение всех столбцов какой-либо жордановой цепочки на произвольное число, отличное от нуля.*

II. *Прибавление к каждому (начиная соз второго) столбца жордановой цепочки предыдущего столбца той же цепочки, предварительно умноженного на то же произвольное число.*

III. *Прибавление ко всем столбцам жордановой цепочки соответствующих столбцов другой цепочки, которая содержит такое же или большее число столбцов и отвечающей тому же характеристическому числу.*

Пример 1.

$$A=\left\|\begin{matrix}1 & 0 & 0 & 1 & -1\\ 0 & 1 & -2 & 3 & -3\\ 0 & 0 & -1 & 2 & -2\\ 1 & -1 & 1 & 0 & 1\\ 1 & -1 & 1 & -1 & 2\end{matrix}\right\|;$$

$$\Delta(\lambda)=(\lambda-1)^4(\lambda+1),$$

$$\psi(\lambda)=(\lambda-1)^2(\lambda+1)=\lambda^3-\lambda^2-\lambda+1.$$

Элементарные делители матрицы A

$$(\lambda-1)^2,\ (\lambda-1)^2,\ \lambda+1.$$

$$J=\left\|\begin{matrix}1 & 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & -1\end{matrix}\right\|,$$

$$\Psi(\lambda,\mu)=\frac{\psi(\mu)-\psi(\lambda)}{\mu-\lambda}=\mu^2+(\lambda-1)\mu+\lambda^2-\lambda-1,$$

$$C(\lambda)=\Psi(\lambda E, A)=A^2+(\lambda-1)A+(\lambda^2-\lambda-1)E.$$

Вычисляем последовательно столбцы матрицы A^2 и соответствующие столбцы матриц $C(\lambda)$, $C(1)$, $C'(\lambda)$, $C(1)$, $C(-1)$. Нам нужно получить два линейно независимых столбца матрицы $C(1)$ и один отличный от нуля столбец матрицы $C(-1)$.

$$C(\lambda)=\begin{Vmatrix}1&0&0&2&*\\0&1&0&2&*\\0&0&1&0&*\\2&-2&2&-1&*\\2&-2&2&-2&*\end{Vmatrix}+(\lambda-1)\begin{Vmatrix}1&0&0&1&*\\0&1&-2&3&*\\0&0&-1&2&*\\1&-1&1&0&*\\1&-1&1&-1&*\end{Vmatrix}+(\lambda^2-\lambda-1)\begin{Vmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{Vmatrix},$$

$$C(1)=\begin{Vmatrix}0&0&0&2&*\\0&0&0&2&*\\0&0&0&0&*\\2&-2&2&-2&*\\2&-2&2&-2&*\end{Vmatrix},\quad C'(\lambda)=\begin{Vmatrix}1&0&0&1&*\\0&1&-2&3&*\\0&0&-1&2&*\\1&-1&1&0&*\\1&-1&1&-1&*\end{Vmatrix}+(2\lambda-1)\begin{Vmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{Vmatrix},$$

$$C'(1)=\begin{Vmatrix}2&*&*&1&*\\0&*&*&3&*\\0&*&*&2&*\\1&*&*&1&*\\1&*&*&-1&*\end{Vmatrix},\quad C(-1)=\begin{Vmatrix}0&0&0&*&*\\0&0&4&*&*\\0&0&4&*&*\\0&0&0&*&*\\0&0&0&*&*\end{Vmatrix}.$$

Поэтому (здесь нижний индекс означает номер столбца; например $C'_4(1)$ обозначает четвертый столбец матрицы $C'(1)$)

$$T=(C_1(1),C_1'(1),C_4(1),C_4'(1),C_3(-1))=\begin{Vmatrix}0&2&2&1&0\\0&0&2&3&4\\0&0&0&2&4\\2&1&-2&1&0\\2&1&-2&-1&0\end{Vmatrix}.$$

Матрицу T можно немного упростить. Последовательно

1) разделим пятый столбец на 4;
2) к третьему столбцу прибавим первый, к четвертому - второй;
3) из четвертого столбца вычтем третий;
4) разделим первый и второй столбцы на 2;
5) вычтем из второго столбца первый, предварительно умноженный на 1/2 . Получим матрицу

$$T_1=\begin{Vmatrix}0&1&2&1&0\\0&0&2&1&1\\0&0&0&2&1\\1&0&0&2&0\\1&0&0&0&0\end{Vmatrix}.$$

Предлагаем читателю проверить, что $AT_1 = T_1J$ и $|\ T_1\ | \neq 0$.

Пример 2.

$$A=\begin{Vmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{Vmatrix};$$

$$\Delta(\lambda)=(\lambda+1)^4,$$
$$\psi(\lambda)=(\lambda+1)^3.$$

Элементарные делители: $(\lambda+1)^3$, $\lambda+1$.

$$J=\begin{Vmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Vmatrix}.$$

Составляем многочлены

$$h_1(\lambda)=\frac{\psi(\lambda)}{\lambda+1}=(\lambda+1)^2,\quad h_2(\lambda)=\frac{\psi(\lambda)}{(\lambda+1)^2}=\lambda+1,\quad h_3(\lambda)=\frac{\psi(\lambda)}{(\lambda+1)^3}=1$$

и матрицы (так как имеется только один элементарный делитель наивысшей степени, то ранг матрицы C должен быть равен единице. Поэтому достаточно, например, вычислить 7 элементов, которые стоят в первом столбце и во второй строке матрицы C. Тогда сразу определятся другие элементы матрицы C)

$$C=h_1(A)=(A+E)^2,\ D=h_2(A)=A+E,\ F=E:$$

$$C=\begin{Vmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & -1 & 1 \end{Vmatrix},\ D=\begin{Vmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{Vmatrix},\ F=\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

В качестве первых трех столбцов матрицы T возьмем третьи столбцы этих матриц: $T=(C_3, D_3, F_3, *)$. В матрицах C, D, F из первого столбца вычтем удвоенный третий, а ко второму и четвертому столбцам прибавим третий. Получим:

$$\tilde{C}=\begin{Vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{Vmatrix},\ \tilde{D}=\begin{Vmatrix} 0 & 0 & 1 & 0 \\ 7 & -1 & -5 & -1 \\ 0 & 0 & 4 & 0 \\ -7 & 1 & 11 & 1 \end{Vmatrix},\ \tilde{F}=\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

В матрицах $\tilde{D}, \tilde{F}$ к первому столбцу прибавим четвертый столбец, умноженный на 7, а из второго столбца вычтем четвертый. Получим:

$$\tilde{C} = \begin{Vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{Vmatrix}, \quad \tilde{\tilde{D}} = \begin{Vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -5 & -1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 11 & 1 \end{Vmatrix}, \quad \tilde{\tilde{F}} = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 5 & 0 & 1 & 1 \\ 7 & -1 & 0 & 1 \end{Vmatrix}.$$

В качестве последнего столбца в T берем первый столбец в $\tilde{\tilde{F}}$.
Имеем:

$$T = (C_3, D_3, F_3, \tilde{\tilde{F}}_1) = \begin{Vmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & 0 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 7 \end{Vmatrix}.$$

Для контроля можно проверить, что $AT = TJ$ и $| T | \neq 0$.

Микромодуль 11
Матричные уравнения

В этом микромодуле будут рассмотрены некоторые типы матричных уравнений, которые встречаются в разнообразных вопросах теории матриц и ее приложениях.

4.8. Уравнение AX = *XB*

Пусть дано уравнения

$$AX = XB, \tag{4.98}$$

где A и B — две заданные квадратные матрицы (вообще говоря, разных порядков)

$$A = ||a_{ij}||^m_1, \quad B = ||b_{ki}||^n_1,$$

а X — искомая прямоугольная матрица размером $m \times n$:

$$X = ||x_{jk}|| \quad (j = 1,2, ...,m;\ k = 1,2, ...,n).$$

Выпишем элементарные делители матриц A и B (в поле комплексных чисел):

$$(A)\colon\ (\lambda - \lambda_1)^{p_1},\ \ (\lambda - \lambda_2)^{p_2},\ \ldots,\ (\lambda - \lambda_u)^{p_u}\quad (p_1 + p_2 + \ldots + p_u = m),$$

$$(B)\colon\ (\lambda - \mu_1)^{q_1},\ \ (\lambda - \mu_2)^{q_2},\ \ldots,\ (\lambda - \mu_v)^{q_v}\quad (q_1 + q_2 + \ldots + q_v = n).$$

В соответствии с этими элементарным делителям приведем матрицы A и B к нормальной жордановой форме

$$A = U\tilde{A}U^{-1} \quad B = V\tilde{B}V^{-1} \tag{4.99}$$

где U и V — квадратные неособенные матрицы соответственно порядка m и n, а $\tilde{A}$ и $\tilde{B}$ — жордановы матрицы:

$$\tilde{A} = \{\lambda_1 E^{(p_1)} + H^{(p_1)},\quad \lambda_2 E^{(p_2)} + H^{(p_2)},\ \ldots,\ \lambda_u E^{(p_u)} + H^{(p_u)}\},$$

$$\tilde{B} = \{\mu_1 E^{(q_1)} + H^{(q_1)},\quad \mu_2 E^{(q_2)} + H^{(q_2)},\ \ldots,\ \mu_v E^{(q_v)} + H^{(q_v)}\}. \tag{4.100}$$

Подставляя в уравнение (4.98) вместо A и B их выражения (4.99), получим:

$$U\tilde{A}U^{-1}X = XV\tilde{B}V^{-1}.$$

Умножим обе части этого равенства слева на U^{-1}, а справа — на V:

$$\tilde{A}U^{-1}XV = U^{-1}XV\tilde{B}. \tag{4.101}$$

Вводя вместо искомой матрицы X новую искомую матрицу $\tilde{X}$ (тех же размеров $m\times n$)

$$\tilde{X} = U^{-1}XV, \tag{4.102}$$

мы уравнение (4.101) запишем так:

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{B}. \tag{4.103}$$

Мы заменили матричное уравнение (4.98) уравнением (4.103) того же вида, но в котором заданные матрицы имеют нормальную жорданову форму.

В соответствии с квазидиагональным видом матриц $\tilde{A}$ и $\tilde{B}$ разобьем матрицу $\tilde{X}$ на блоки:

$$\tilde{X} = (X_{\alpha\beta}) \quad (\alpha = 1,2, \ldots, u;\quad \beta = 1,2, \ldots, v)$$

(здесь $X_{\alpha\beta}$ — прямоугольная матрица размером $p_\alpha \times q_\beta$; $\alpha = 1, 2, \ldots, u$; $\beta = 1,2, \ldots, v$).

Используя правило умножения блочной матрицы на квазидиагональную, произведем умножение матриц в левой и правой частях уравнения (4.103). Тогда это уравнение распадается на uv матричных уравнений

$$[\lambda_\alpha E^{(p_\alpha)} + H^{(p_\alpha)}]\, X_{\alpha\beta} = X_{\alpha\beta}\, [\mu_\beta E^{(q_\beta)} + H^{(q_\beta)}]$$

$$(\alpha = 1, 2, \ldots, u;\quad \beta = 1, 2, \ldots, v),$$

которые перепишем еще так:

$$(\mu_\beta - \lambda_\alpha)\, X_{\alpha\beta} = H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta \quad (\alpha = 1, 2, \ldots, u;\ \beta = 1, 2, \ldots, v); \tag{4.104}$$

при этом мы ввели сокращенные обозначения

$$H_\alpha = H^{(p_\alpha)},\quad G_\beta = H^{(q_\beta)} \quad (\alpha = 1, 2, \ldots, u;\quad \beta = 1, 2, \ldots, v).$$

(4.105)

Возьмем какое-нибудь из уравнений (4.104). Могут представиться два случая:

1. $\lambda_\alpha \neq \mu_\beta$. Проитерируем r — 1 раз равенство (4.104) (обе части равенства (4.104) множаем на $\mu_\beta - \lambda_\alpha$ и в каждом члене правой части заменяем $(\mu_\beta - \lambda_\alpha)X_{\alpha\beta}$ на $H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta$. Этот процесс повторяем r — 1 раз):

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = \sum_{\sigma+\tau=r} (-1)^\tau H_\alpha^\sigma X_{\alpha\beta} G_\beta^\tau. \tag{4.106}$$

Заметим, что в силу (4.105)

$$H_\alpha^{p_\alpha} = G_\beta^{q_\beta} = 0. \tag{4.107}$$

Если в (4.106) взять $r \geq p_\alpha + q_\beta - 1$, то в каждом члене суммы, которая стоит в правой части равенства (4.106), выполняется по крайней мере одно из соотношений

$$\sigma \geq p_\alpha, \quad \tau \geq q_\beta$$

и потому в силу (4.107) либо $H_\alpha^\sigma = 0$, либо $G_\beta^\tau = 0$. Так как, кроме того, в рассматриваемом случае $\lambda_\alpha \neq \mu_\beta$, то из (4.16) находим:

$$X_{\alpha\beta} = 0. \tag{4.108}$$

2. $\lambda_\alpha = \mu_\beta$. В этом случае уравнения (4.104) принимает вид

$$H_\alpha X_{\alpha\beta} = X_{\alpha\beta} G_\beta. \tag{4.109}$$

В матрицах H_α и G_β элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Учитывая эту специфическую структуру матриц H_α и G_β и полагая

$$X_{\alpha\beta} = \|\xi_{ik}\| \quad (i = 1, 2, \ldots, p_\alpha;\ k = 1, 2, \ldots, q_\beta),$$

мы заменим матричное уравнение (4.109) следующей эквивалентной ему системой скалярных соотношений (из структуры матриц H_α и G_β следует, что произведение $H_\alpha X_{\alpha\beta}$ получается из $X_{\alpha\beta}$ сдвигом всех строк на одно место вверх и заполнением последней строки нулями: аналогично $X_{\alpha\beta}\, G_\beta$ получается из $X_{\alpha\beta}$ сдвигом всех столбцов на одно место вправо и заполнением первого столбца нулями. Для сокращения обозначений мы не пишем при ξ_{ik} дополнительных индексов α, β):

$$\xi_{i+1,\,k} = \xi_{i,\,k-1}\ (\xi_{i0} = \xi_{p_\alpha+1,\,k} = 0;\ i = 1, 2, \ldots, p_\alpha;\ k = 1, 2, \ldots, q_\beta). \tag{4.110}$$

Равенства (4.110) означают:

1) В матрице $X_{\alpha\beta}$ на каждой линии, которая параллельна главной диагонали, стоят разные между собой элементы,

2) $\xi_{21} = \xi_{31} = \ldots = \xi_{p_\alpha 1} = \xi_{p_\alpha 2} = \ldots = \xi_{p_\alpha,\, q_\beta - 1} = 0.$

Пусть $p_\alpha = q_\beta$. В этом случае $X_{\alpha\beta}$ — квадратная матрица. Из 1), 2) следует, что в матрице $X_{\alpha\beta}$ все элементы, которые расположены под главной диагональю, равны нулю, все элементы главной диагонали равны некоторому числу $c_{\alpha\beta}$, все элементы первой наддиагонали равны некоторому числу $c'_{\alpha\beta}$ и т.д., т.е.

$$X_{\alpha\beta} = \begin{Vmatrix} c_{\alpha\beta} & c'_{\alpha\beta} & \dots & c_{\alpha\beta}^{(p_\alpha-1)} \\ 0 & c_{\alpha\beta} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & c'_{\alpha\beta} \\ 0 & \dots & 0 & c_{\alpha\beta} \end{Vmatrix} = T_{p_\alpha};$$

$$(p_\alpha = q_\beta)$$

(4.111)

здесь

$$c_{\alpha\beta},\ c'_{\alpha\beta},\ \dots,\ c_{\alpha\beta}^{(p_\alpha-1)}$$

— произвольные параметры (уравнение (4.109) не накладывают никаких ограничений на значение этих параметров).

Легко видеть, что при $p_\alpha < q_\beta$

$$X_{\alpha\beta} = (\overbrace{\ 0\ }^{q_\beta - p_\alpha},\ T_{p_\alpha}),$$

(4.112)

а при $p_\alpha > q_\beta$

$$X_{\alpha\beta} = \begin{pmatrix} T_{\alpha\beta} \\ 0 \end{pmatrix} \}p_\alpha - q_\beta\ .$$

(4.113)

О матрицах (4.111), (4.112) и (4.113) мы будем говорить, что они имеют *правильную верхнюю треугольную форму*. Число произвольных параметров в $X_{\alpha\beta}$ равно наименьшему из чисел p_α и q_β, Приведенная ниже схема показывает структуру матрицы $X_{\alpha\beta}$ при $\lambda_\alpha = \mu_\beta$ (произвольные параметры здесь обозначены через *a, b, c, d):*

$$X_{\alpha\beta}=\begin{Vmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{Vmatrix}, \quad X_{\alpha\beta}=\begin{Vmatrix} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{Vmatrix}, \quad X_{\alpha\beta}=\begin{Vmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & 0 \end{Vmatrix}.$$

$$(p_\alpha = q_\beta = 4) \qquad (p_\alpha = 3,\ q_\beta = 5) \qquad (p_\alpha = 5,\ q_\beta = 3)$$

Для того чтобы при подсчете произвольных параметров в матрице $\tilde{X}$ охватить и случай 1, обозначим через $d_{\alpha\beta}(\lambda)$ наибольший общий делитель элементарных делителей $(\lambda-\lambda_\alpha)^{p_\alpha}$ и $(\lambda-\mu_\beta)^{q_\beta}$, а через $\delta_{\alpha\beta}$ — степень многочлена $d_{\alpha\beta}(\lambda)$ ($\alpha = 1, 2, ..., u$; $\beta = 1, 2, ..., v$). В случае 1 $\delta_{\alpha\beta} = 0$; в случае 2 имеем: $\delta_{\alpha\beta} = \min(p_\alpha, q_\beta)$. Таким образом, в обеих случаях число произвольных параметров в $X_{\alpha\beta}$ равно $\delta_{\alpha\beta}$. Число произвольных параметров в $\tilde{X}$ определяется формулой

$$N=\sum_{\alpha=1}^{u}\sum_{\beta=1}^{v}\delta_{\alpha\beta}.$$

В дальнейшем нам удобно будет общее решение уравнения (4.103) обозначить через $X_{\tilde{A}\tilde{B}}$ (до сих пор мы это решение обозначали буквой $\tilde{X}$).

Полученные в этом пункте результаты можно сформулировать в виде следующей теоремы:

Теорема 1. *Общее решение матричного уравнения*

$$AX = XB,$$

где

$$A=\|a_{ik}\|_1^m=U\tilde{A}U^{-1}=U\{\lambda_1E^{(p_1)}+H^{(p_1)},\ \ldots,\ \lambda_uE^{(p_u)}+H^{(p_u)}\}\,U^{-1},$$

$$B=\|b_{ik}\|_1^n=V\tilde{B}V^{-1}=V\{\mu_1E^{(q_1)}+H^{(q_1)},\ \ldots,\ \mu_vE^{(q_v)}+H^{(q_v)}\}\,V^{-1},$$

задается формулой

$$X=U\,X_{\tilde{A}\tilde{B}}\,V^{-1}. \tag{4.114}$$

Здесь $X_{\tilde{A}\tilde{B}}$ *- общее решение уравнения* $\tilde{A}\tilde{X}=\tilde{X}\tilde{B}$ *— имеет следующую структуру:* $X_{\tilde{A}\tilde{B}}$ *разбивается на блоки*

$$X_{\tilde{A}\tilde{B}}=(\overbrace{X_{\alpha\beta}}^{q_\beta})\}p_\alpha \quad (\alpha=1, 2, \ldots, u;\ \beta=1, 2, \ldots, v);$$

если $\lambda_\alpha \neq \mu_\beta$, *то на месте* $X_{\alpha\beta}$ *стоит нулевая матрица, если же* $\lambda_\alpha = \mu_\beta$, *то на месте* $X_{\alpha\beta}$ *стоит произвольная правильная верхняя треугольная матрица.*

$X_{\alpha\beta}$, *а следовательно, и* X *зависят линейно от* N *произвольных параметров* $c_1, c_2, ..., c_N$:

$$X = \sum_{i=1}^{N} c_j X_j,$$

(4.115)

где N *определяется формулой*

$$N = \sum_{\alpha=1}^{u} \sum_{\beta=1}^{v} \delta_{\alpha\beta} \tag{4.116}$$

[*здесь* $\delta_{\alpha\beta}$ *обозначает степень наибольшего общего делителя* $(\lambda - \lambda_\alpha)^{p_\alpha}$ и $(\lambda - \mu_\beta)^{q_\beta}$,].

Заметим, что матрицы $X_1, X_2, .. ., X_N$, фигурирующие в формуле (4.115), суть решения исходного уравнения (4.98) (матрица X_j получается из X, если параметру c_j дать значение единица, а другим параметрам — нулевые значения; $j = 1,2, ..., N$). Эти решения линейно независимы, так как в противном случае при некоторых значениях параметров $c_1, c_2, ..., c_N$, не равных одновременно нулю, матрица X, а следовательно, и $X_{\alpha\beta}$ равнялись бы нулю, что невозможно. Таким образом, равенство (4.116) показывает, что любое решение исходного уравнения представляет собой линейную комбинацию N линейно независимых решений.

Если матрицы A и B *не имеют общих характеристических чисел* (характеристические многочлены $|\lambda E — A|$ и $|\lambda E — B|$ взаимно просты), *то* $N = N = \sum_{\alpha=1}^{u} \sum_{\alpha=1}^{v} \delta_{\alpha\beta} = 0$ и, следовательно, $X = 0$, т.е. *в этом случае уравнение* (4.98) *имеет только тривиальное нулевое решение* $X = 0$.

Замечание. Пусть элементы матриц A и B принадлежат некоторому числовому полю K. Тогда нельзя утверждать, что элементы матриц U, V, $X_{\alpha\beta}$, фигурирующие в формуле (4.114), также принадлежат полю K. Элементы этих матриц можно выбрать в расширенном поле K_1, которое получается из поля K путем приобщения к последнему корней характеристических уравнений $|\lambda E — A| = 0$ и $|\lambda E — B| = 0$. С такого рода расширением основного поля всегда приходится иметь дело, когда пользуются приведением заданных матриц к нормальной жордановой форме.

Однако матричное уравнение (4.98) эквивалентно системе *mn* линейных однородных уравнений, где неизвестными служат элементы x_{ik} (j = 1, 2, ..., m; k = 1,2, ..., n) искомой матрицы X:

$$\sum_{j=1}^{m} a_{ij}x_{jk} = \sum_{h=1}^{n} x_{ih}b_{hk} \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n). \tag{4.117}$$

Нами доказано, что эта система имеет N линейно независимых решений, где N определяется формулой (4.116). Но известно, что базисные линейно независимые решения можно выбрать в основном поле K, которому принадлежат коэффициенты уравнений (4.117). Таким образом, в формуле (4.115) матрицы X_1, X_2, ……,X_N можно выбрать так, чтобы их элементы принадлежали полю K. Тогда, придавая в формуле (4.115) произвольным параметрам всевозможные значения из поля K, мы получим все матрицы X с элементами из K, удовлетворяющие уравнению (4.98) (матрицы $A=\|a_{ij}\|^n_1$ и $B=\|b_{ik}\|^n_1$ определяют линейный оператор $\tilde{F}(X)=AX - XB$ в пространстве прямоугольных матриц X с размерами $m \times n$).

4.9. Различные матричные уравнения

1. Частный случай: *A* = *B*. Перестановочные матрицы

Рассмотрим частный случай уравнения (4.98) - уравнение

$$AX = XA, \tag{4.118}$$

где $A = \|a_{ik}\|^n_1$ — заданная, а $X = \| x_{ik} \|^n_1$ — искомая матрица. Мы пришли к задаче Фробениуса: определить все матрицы X, перестановочные с данной матрицей A.

Приведем матрицу A к нормальной жордановой форме:

$$A = U\tilde{A}U^{-1} = U\{\lambda_1 E^{(p_1)} + H^{(p_1)}, \ldots, \lambda_u E^{(p_u)} + H^{(p_u)}\}\, U^{-1}. \tag{4.119}$$

Тогда, полагая в формуле (4.114) $V=U$, $\tilde{B} = \tilde{A}$ и обозначая $X_{\tilde{A}\tilde{A}}$ сокращенно через $X_{\tilde{A}}$, получим все решения уравнения (4.118), т.е. все матрицы, которые являются перестановочными с A, в следующем виде:

$$X = U X_{\tilde{A}} U^{-1} \tag{4.120}$$

где $X_{\tilde{A}}$ обозначает произвольную матрицу, перестановочною с A. Как было выяснено в предыдущем пункте, $X_{\tilde{A}}$ разбивается на u^2 блоков

$$X_{\tilde{A}} = (X_{\alpha\beta})^{n}{}_{1}$$

В соответствии с разбиением на блоки жордановой матрицы $\tilde{A}$; $X_{\alpha\beta}$ — нулевая матрица дибо произвольная правильная верхняя треугольная матрица в зависимости от того, $\lambda_\alpha \neq \lambda_\beta$ или $\lambda_\alpha = \lambda_\beta$.

Для примера выпишем элементы матрицы $X_{\tilde{A}}$ для случая, когда матрица A имеет следующие элементарные делители:

$$(\lambda - \lambda_1)^4, \quad (\lambda - \lambda_1)^3, \quad (\lambda - \lambda_2)^2, \quad \lambda - \lambda_2 \quad (\lambda_1 \neq \lambda_2).$$

В этом случае $X_{\tilde{A}}$ имеет такой вид:

$$\left\| \begin{array}{cccc|ccc|cc|c}
a & b & c & d & e & f & g & 0 & 0 & 0 \\
0 & a & b & c & 0 & e & f & 0 & 0 & 0 \\
0 & 0 & a & b & 0 & 0 & e & 0 & 0 & 0 \\
0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & h & k & l & m & p & q & 0 & 0 & 0 \\
0 & 0 & h & k & 0 & m & p & 0 & 0 & 0 \\
0 & 0 & 0 & h & 0 & 0 & m & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & r & s & t \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & r & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w & z
\end{array} \right\|$$

($a,b,\dots,z$ — произвольные параметры).

Число параметров в $X_{\tilde{A}}$ равно N, где $N = \sum_{\alpha,\beta=1}^{u} \delta_{\alpha\beta}$; здесь $\delta_{\alpha\beta}$ обозначает степень наибольшего общего делителя многочленов $(\lambda - \lambda_\alpha)^{p_\alpha}$ и $(\lambda - \mu_\beta)^{q_\beta}$, .

Введем в рассмотрение инвариантные многочлены матрицы A: $i_1(\lambda)$, $i_2(\lambda)$,... ..., $i_t(\lambda)$; $i_{t+1}(\lambda) = \dots = i_n(\lambda) = 1$. Степени этих многочленов обозначим через $n_1 \geq n_2 \geq \dots \geq n_t > n_{t+1} = \dots = 0$. Так как каждый инвариантный многочлен является произведением нескольких попарно взаимно простых элементарных делителей, то формулу для N можно записать и так:

$$N = \sum_{g,j=1}^{t} \varkappa_{gj}, \tag{4.121}$$

где $\varkappa_{gj}$ — степень наибольшего общего делителя многочленов $i_g(\lambda)$ и $i_j(\lambda)$ $(g, j = 1, 2, ..., t)$. Но наибольшим общим делителем многочленов $i_g(\lambda)$ и $i_j(\lambda)$ является один из этих же многочленов и потому $\varkappa_{gj}= \min(n_q, n_j)$. Отсюда получаем:

$$N = n_i + 3n_2 + ... + (2t-1)n_t.$$

Число N является числом линейно независимых матриц, перестановочных с матрицей A (можно считать, что элементы этих матриц принадлежат основному полю K, содержащему элементы матрицы A; см. замечание в конце предыдущего пункта). Мы пришли к теореме:

Теорема 2. *Число линейно независимых матриц перестановочных с матрицей* $A = \| a_{ik} \|^n_1$, *определяется формулой*

$$N = n_i + 3n_2 + ... + (2t-1)n_t, \tag{4.122}$$

где $n_1, n_2, .. .n_t$ *— степени непостоянных инвариантных многочленов* $i_1(\lambda)$, $i_2(\lambda)$, … $i_t(\lambda)$ *матрицы* A.

Заметим, что

$$n = n_1 + n_2 + ... + n_t \tag{4.123}$$

Из (4.122) и (4.123) вытекает:

$$N \geq n, \tag{4.124}$$

причем знак = имеет место в том и только в том случае, когда $t=1$, т.е. когда все элементарные делители матрицы A попарно взаимно просты.

Пусть $g(\lambda)$ — некоторый многочлен от λ. Тогда матрица $g(A)$ перестановочна с A. Возникает обратный вопрос: в каком случае любая матрица, перестановочная с A, может быть представлена как многочлен от A? В этом случае любая матрица, перестановочная с A, была бы линейной комбинацией линейно независимых матриц

$$E, A, A^2 ..., A.^{n_i - 1}$$

В рассматриваемом случая $N=n_1 \leq n$; сопоставляя с (4.124), получаем: $N = n_1 =$ п*г.* Таким образом, мы получили следующее следствие.

Следствие 1 из теоремы 2. *Все матрицы, перестановочные с A, представляются как многочлены от A в том и только в том случае, когда* $n_1 = n$, *т.е. когда все элементарные делители матрицы A попарно взаимно просты.*

Многочлены от матрицы, перестановочной A, также перестановочны с A. Поставим вопросы: в каком случае все матрицы, перестановочнык с A, представляются в виде многочленов от некоторой (одной и той же) матрицы C? Предположим, что такой случай имеет место. Тогда, так как матрица C удовлетворяет в силу теоремы Гамильтона-Кели своему характеристическому уравнению, то любая матрица, перестановочная с C, выразится линейно через матрицы

$$E, C, C^2 \ldots, C^{n-1}.$$

Поэтому в рассматриваемом случая $N \leq n$. Сопоставляя с (4.124), находим $N = n$. Но тогда из (4.122) и (4.123) и $n_1 = n$.

Следствие 2 из теоремы 2. *Все матрицы, перестановочные с A, представляются в виде многочленов от одной и той же матрицы C в том и только в том случае, когда $n_1 = n$, т.е. когда все элементарные делители в λE — A взаимно просты. В этом случае все матрицы, перестановочные с A, представляются и в виде многочленов от A.*

Отметим еще одно свойство перестановочных матриц.

Теорема 3. *Если две матрицы $A = \|a_{ik}\|^n_1$, $B = \| b_{ik} \|^n_1$ перестановочны и одна из них, например A, имеет квазидиагональный вид*

$$A = \{\overbrace{A_1}^{s_1}, \overbrace{A_2}^{s_2}\}, \tag{4.125}$$

где матрицы A_1 и A_2 не имеют общих характеристических чисел, то и другая матрица имеет такой же квазидиагональный вид

$$B = \{\overbrace{B_1}^{s_1}, \overbrace{B_2}^{s_2}\}. \tag{4.126}$$

Доказательство. Разобьем матрицу B на блоки в соответствии квазидиагональным видом (4.125):

$$B = \begin{pmatrix} \overbrace{B_1}^{s_1} & \overbrace{X}^{s_2} \\ Y & B_2 \end{pmatrix}.$$

Записывая, что $AB = BA$, получим четыре матричных равенства:

1. $A_1B_1 = B_1A_1$, 2. $A_1X = XA_2$, 3. $A_2Y = YA_1$, 4. $A_2B_2 = B_2A_2$. (4.127)

Второе и третье из уравнений (4.127), как было выяснено в п. 4.9, имеют только нулевые решения $X = 0$, $Y = 0$, поскольку матрицы A_1 и A_2 не имеют общих характеристических чисел. Таким образом, наше предложение доказано. Первое и четвертое из уравнений (4.127) выражают перестановочность матриц A_1 и B_1, A_2 и B_2.

Доказанное предложение в геометрической формулировке гласит:

Теорема 3'. *Если*

$$\boldsymbol{R} = \boldsymbol{I}_1 + \boldsymbol{I}_2$$

*есть расщепление всего пространства **R** на инвариантные относительно оператора **A** подпространства **I**$_1$ и **I**$_2$ и минимальные многочлены этих подпространств (относительно **A**) взаимно просты,* *то эти подпространства* I_1 *и* I_2 *инвариантны относительно любого линейного оператора* B*, перестановочного с* A.

Из доказанной теоремы вытекает следующее следствие:

Следствие 1. *Если линейные операторы* $A, B, ..., L$ *попарно перестановочны, то можно расщепить все пространство* R *на инвариантные относительно всех операторов* $A, B, ..., L$ *подпространства*

$$R=I_1 + I_2+ ...+I_W$$

так, чтобы минимальный многочлен каждого из этих подпространств относительно любого из операторов $A, B, ..., L$ *был степенью неприводимого многочлена.*

Как частный случай отсюда получим

Следствие 2. *Если линейные операторы* $A, B, ..., L$ *попарно перестановочны и все характеристические числа этих операторов принадлежат основному полю К, то можно расщепить все пространство* R *на инвариантные относительно всех операторов* $A, B, ..., L$ *подпространства* $I_1, I_2, ..., I_w$*, в каждом из которых любой из операторов* $A, B, ..., L$ *имеет равные характеристические числа.*

И, наконец, отметим частный случай этого предложения:

Следствие 3. *Если операторы простой структуры* $A, B, ..., L$ *попарно перестановочны, то можно составить базис пространства из общих собственных векторов этих операторов.*

Дадим матричную формулировку последнего предложения: *Перестановочные матрицы простой структуры можно одновременно, т.е. одним и тем же преобразованием подобия, привести к диагональному виду.*

2. Матричное уравнение $AX - XB = C$. Пусть дано матричное уравнение

$$AX-XB = C, \tag{4.128}$$

где $A = \|a_{ij}\|^n_1$, $B = \| b_{kl} \|^n_1$ — заданные квадратные матрицы порядков m и n, $C = \|C_{jk} \|$ — заданная, а $X=\| x_{jk} \|$ — искомая прямоугольная матрица размером $m \times n$. Уравнение (4.128) эквивалентно системе mn скалярных уравнений относительно элементов матрицы X:

$$\sum_{j=1}^{m} a_{ij}x_{jk} - \sum_{l=1}^{n} x_{il}b_{lk} = c_{ik} \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n).$$

Соответствующая однородная система уравнений

$$\sum_{j=1}^{m} a_{ij}x_{jk} - \sum_{l=1}^{n} x_{il}b_{lk} = 0 \quad (i = 1, 2, \ldots, m;\ k = 1, 2, \ldots, n)$$

в матричном виде записывается так:

$$AX\text{-}XB = 0. \tag{4.129}$$

Таким образом, если уравнение (4.129) имеет только нулевое решение, то уравнение (4.128) имеет один-единственное решение. Но в п. 4.8 было установлено, что уравнение (4.129) имеет только нулевое решение тогда и только тогда, когда матрицы A и B не имеют общих характеристических чисел. Следовательно, *если матрицы A и B не имеют общих характеристических чисел, то уравнение* (4.128) *имеет одно и только одно решение; если же матрицы A и B имеют общие характеристические числа, то в зависимости от «свободного члена» C могут представиться два случая: либо уравнение* (4.128) *противоречиво, либо оно имеет бесчисленное множество решений, которые задаются формулой*

$$X=X_0+X_1,$$

где X_0 — фиксированное частное решения уравнения (4.128), X_1 — общее решение однородного уравнения (4.129) (структура X_1 была выяснена в п.4.9).

3. Скалярное уравнение *f(X)* =0. Рассмотрим сначала уравнение

$$G(X) = 0, \tag{4.130}$$

где

$$g(\lambda) = (\lambda - \lambda_1)^{a_1} (\lambda - \lambda_2)^{a_2} \dots (\lambda - \lambda_h)^{a_h}$$

— заданный многочлен переменной λ, а X — искомая квадратная матрица порядка n. Так как минимальный многочлен матрицы X, т.е. первый инвариантный многочлен $i_1(\lambda)$, должен быть делителем многочлена $g(\lambda)$, то элементарные делители матрицы X должны иметь такой вид:

$$(\lambda - \lambda_{i_1})^{p_{i_1}}, (\lambda - \lambda_{i_2})^{p_{i_2}}, \dots, (\lambda - \lambda_{i_\nu})^{p_{i_\nu}} \begin{pmatrix} i_1, i_2, \dots, i_\nu = 1, 2, \dots, h, \\ p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_h} \leq a_{i_h}, \\ p_{i_1} + p_{i_2} + \dots + p_{i_\nu} = n \end{pmatrix}$$

(среди индексов $i_1, i_2, \dots, i_\nu$ могут быть и равные, n — заданный порядок искомой матрицы X).

Искомая матрица X представится в виде

$$X = T\{\lambda_{i_1}E^{(p_{i_1})} + H^{(p_{i_1})}, \dots, \lambda_{i_\nu}E^{(p_{i_\nu})} + H^{(p_{i_\nu})}\}T^{-1}, \tag{4.131}$$

где T — произвольная неособенная матрица порядка n. Множество решений уравнения (4.130) с заданным порядком искомой матрицы распадается согласно формуле (4.131) на конечное число классов подобных между собой матриц.

Пример 1. Дано уравнения

$$X^m = 0. \qquad (4.132)$$

Если некоторая степень матрицы равна нулю, то матрица называется *нильпотентной*. Наименьший из показателей, при которых степень матрицы равен нулю, называется *индексом нильпотентности* данной матрицы.

Очевидно, решениями уравнения (4.132) есть все нильпотентные матрицы с индексом нильпотентности $\mu \leq m$. Формула, которая охватывает все решения данного порядка n, выглядит так:

$$X = T\{H^{(p_1)}, H^{(p_2)}, \ldots, H^{(p_\nu)}\}T^{-1} \begin{pmatrix} p_1 \cdot p_2, \ldots, p_\nu \leqslant m, \\ p_1 + p_2 + \ldots + p_\nu = n \end{pmatrix} \qquad (4.133)$$

(T— произвольная неособенная матрица).

Пример 2. Дано уравнения

$$X^2 = X. \qquad (4.134)$$

Матрица, которая удовлетворяет этому уравнению, называется *идемпотентной*. Элементарными делителями идемпотентной матрицы могут быть только λ либо λ - 1. Поэтому идемпотентную матрицу можно определить как матрицу простой структуры (т.е. такую, которая приводится к диагональной форме) с характеристическими числами, равными нулю или единице. Формула, которая охватывает все идемпотентные матрицы данного порядка, имеет вид

$$X = T\underbrace{\{1, 1, \ldots, 1, 0, \ldots, 0\}}_{n} T^{-1}, \qquad (4.135)$$

где T — произвольная неособенная матрица порядка n.

Рассмотрим теперь большее общее уравнение

$$f(X) = 0, \qquad (4.136)$$

где $f(\lambda)$ — регулярная функция в некоторой области G плоскости комплексного аргумента λ. От искомого решения $X = \| x_{ik} \|^n_1$ будем требовать, чтобы характеристические числа его принадлежали области G. Выпишем все нули функции $f(\lambda)$, лежащие в области G, и их кратности:

$$\lambda_1, \lambda_2, \ldots,$$
$$a_1, a_2, \ldots$$

Как и в предыдущем случае, каждый элементарный делитель матрицы X должен иметь вид

$$(\lambda - \lambda_i)^{p_i} \, (p_i \leqslant a_i)$$

и потому

$$X = T\{\lambda_{i_1}E^{(p_{i_1})} + H^{(p_{i_1})}, \ldots, \lambda_{i_\nu}E^{(p_{i_\nu})} + H^{(p_{i_\nu})}\}\, T^{-1}$$
$$(i_1, i_2, \ldots, i_\nu = 1, 2, \ldots; \ p_{i_1} \leqslant a_{i_1}, p_{i_2} \leqslant a_{i_2}, \ldots, p_{i_\nu} \leqslant a_{i_\nu};$$
$$p_{i_1} + p_{i_2} + \ldots + p_{i_\nu} = n) \qquad (4.137)$$

(T - произвольная неособенная матрица).

4. Матричное многочленное уравнение. Рассмотрим уравнение

$$A_0X^m + A_1X^{m-1} + \ldots + A_m = 0, \qquad (4.138)$$

$$Y^mA_0 + Y^{m-1}A_1 + \ldots + A_m = 0, \qquad (4.139)$$

где $A_0, A_1, \ldots, A_m$ — заданные, а X и Y — искомые квадратные матрицы порядка n. Уравнение (4.130) представляет собой весьма частный (можно сказать, тривиальный) случай уравнений (4.138), (4.139) и получается из последних, если положить $A_i = \alpha_i E$, где α_i — число и $i = 1, 2, \ldots, m$.

Следующая теорема устанавливает связь между уравнениями (4.138), (4.139) и (4.130).

Теорема 4. *Каждое решение матричного уравнения*

$$A_0X^m + A_iX^{m-1} + \ldots + A_m = 0$$

удовлетворяет скалярному уравнению

$$g(X) = 0, \qquad (4.140)$$

где

$$g(X) = |A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m|. \qquad (4.141)$$

Этому же скалярному уравнению удовлетворяет и любое решение Y матричного уравнения

$$Y^mA_0 + Y^{m-1}A_1 + \ldots + A_m = 0.$$

Доказательство. Обозначим через $F(\lambda)$ матричный многочлен

$$F(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \ldots + A_m.$$

Тогда уравнения (4.138) и (4.139) запишутся так:

$F(X) = 0, \quad F$€$(Y) = 0.$

Согласно обобщенной теореме Безу, если X и Y — решение этих уравнений, то матричный многочлен $F(\lambda)$ делится справа на $\lambda E - X$ и слева на $\lambda E - Y$:

$$F(\lambda) = Q(\lambda)(\lambda E - X) = (\lambda E - Y)\, Q_1(\lambda).$$

Отсюда

$$g(\lambda) = |F(\lambda)| = |Q(\lambda)|\,\Delta(\lambda) = |Q_1(\lambda)|\,\Delta_1(\lambda), \qquad (4.142)$$

где $\Delta(\lambda) = |\lambda E - X|$ и $\Delta_1(\lambda) = |\lambda E - Y|$ — характеристические многочлены матриц X и Y. По теореме Гамильтона-Кели

$$\Delta(X) = 0, \quad \Delta_1(Y) = 0.$$

Поэтому из (4.142) вытекает:

$$g(X) = g(Y)=0.$$

Теорема доказана.

Мы доказади, что каждое решение уравнения (4.138) удовлетворяет скалярному уравнению (степени $\leq mn$)

$$g(\lambda) =0.$$

Но множество матричных решений этого уравнения с заданным порядком n распадается на конечное число классов подобных между собой матриц. Поэтому все решения уравнения (4.138) приходится искать среди матриц вида

$$Te\, D_i\, Te^{-1} \qquad (4.143)$$

(здесь D_i — известные матрицы; при желании можно считать, что D_i имеют нормальную жорданову форму; T_i — произвольные неособенные матрицы n-го порядка; i = 1, 2, ..., h). Подставим в (4.138) вместо X матрицу (4.143) и подберем T_i так, чтобы удовлетворялось уравнение (4.138). Для каждого T_i получим свое линейное уравнение

$$A_0T_iD_i^m + A_1T_iD_i^{m-1} + \ldots + A_mT_i = 0 \quad (i = 1, 2, ..., h). \qquad (4.144)$$

Единственный способ, который можно предложить для нахождения решения T_i уравнения (4.144), заключается в замене матричного уравнения системой линейных однородных скалярных уравнений относительно элементов искомой матрицы T_i. Каждое неособенное решение T_i уравнения (4.144), будучи подставлено в (4.143), дает решение данного уравнения (4.138). Аналогичные соображения могут быть проведены для уравнения (4.139).

В следующих двух разделах мы рассмотрим частные случаи уравнения (4.138), связанные с извлечением корня m-й степени из матрицы.

Заметим, что теорема Гамильтона-Кели является частичным случаем теоремы 4. В самом деле, любая квадратная матрица A, будучи подставлена вместо λ, удовлетворяет уравнению

$$\lambda E — A = 0.$$

Поэтому в силу доказанной теоремы

$$\Delta(A) = 0,$$

где $\Delta (\lambda) = | \lambda E - A|$.

Теорема 4 может быть обобщена следующим образом:

Теорема 5 (Филлипса). *Если попарно перестановочные между собой квадратные матрицы n-го порядка X_0, X_1, ..., X_m удовлетворяют матричному уравнению*

$$A_0X_0 + A_1X_1+\ldots+ A_mX_m = 0 \qquad (4.145)$$

(A_0, A_1,...,A_m — заданные квадратные матрицы n-го порядка), то эти же матрицы X_0, X_1, ..., X_m удовлетворяют скалярному уравнению

$$g(X_0,X_1,\ldots,X_m)=0, \qquad (4.146)$$

где

$$g(\xi_0, \xi_1, \ldots, \xi_m) = |A_0\xi_0 + A_1\xi_1 + \ldots + A_m\xi_m|. \tag{4.147}$$

Доказательство. Положим ($f_{ik}(\xi_0, \xi_1, \ldots, \xi_m)$ -линейные формы от $\xi_0, \xi_1, \ldots, \xi_m$ $(i, k = 1, 2, \ldots, n)$.

$$F(\xi_0, \xi_1, \ldots, \xi_m) = \| f_{ik}(\xi_0, \xi_1, \ldots, \xi_m) \|_1^n = A_0\xi_0 + A_1\xi_1 + \ldots + A_m\xi_m;$$

$\xi_0, \xi_1, \ldots, \xi_m$ – скалярные переменные.

Обозначим через

$$\widehat{F}(\xi_0, \xi_1, \ldots, \xi_m) = \| \widehat{f}_{ik}(\xi_0, \xi_1, \ldots, \xi_m) \|_1^n$$

присоединенную матрицу для матрицы F [$\widehat{f}_{ik}$ есть алгебраическое дополнение (адъюнкта) элемента f_{ki} в определителе

$|F(\xi_0, \xi_1, \ldots, \xi_m)| = |f_{ik}|_1^n$ $(i, k = 1, 2, \ldots, n)$].

Тогда каждый элемент $\widehat{f}_{ik}$ $(i, k — 1, 2, ..., n)$ матрицы $\widehat{F}$ есть однородный многочлен относительно $\xi_0, \xi_1, \ldots, \xi_{m-1}$ степени $m — 1$, и потому матрицу $\widehat{F}$ можно представить в виде

$$\widehat{F} = \sum_{j_0 + j_1 + \ldots + j_m = n-1} F_{j_0 j_1 \ldots j_m} \xi_0^{j_0} \xi_1^{j_1} \ldots \xi_m^{j_m},$$

где $F_{j_0 j_1 \cdots j_m}$ - некоторые постоянные матрицы порядка n.

Из определения матрицы $\widehat{F}$ следует тождество

$$\widehat{F}F = g(\xi_0, \xi_1, \ldots, \xi_m) E.$$

Запишем это тождество следующим образом:

$$\sum_{j_0 + j_1 + \ldots + j_m = n-1} F_{j_0 j_1 \ldots j_m} (A_0\xi_0 + A_1\xi_1 + \ldots + A_m\xi_m) \xi_0^{j_0} \xi_1^{j_1} \ldots \xi_m^{j_m} = \\ = g(\xi_0, \xi_1, \ldots, \xi_m) E. \tag{4.148}$$

Переход от левой части к правой части в тождестве (4.148) осуществляется путем раскрытия скобок и приведение подобных членов. При этом *приходится переставлять местами переменные* $\xi_0, \xi_1, \ldots, \xi_m$ *между собой и не приходится переставлять местами переменные* $\xi_0, \xi_1, \ldots, \xi_m$ *с матричными коэффициентами* A_i *и* $F_{j_0 j_1 \cdots j_m}$.

Поэтому равенство (4.148) не нарушится, если мы вместо переменных $\xi_0, \xi_1, \ldots, \xi_m$ подставим попарно перестановочные между собой матрицы $X_0, X_1, \ldots, X_m$:

$$\sum_{j_0+j_1+\ldots+j_m=n-1} F_{j_0 j_1 \ldots j_m}(A_0X_0 + A_1X_1 + \ldots + A_mX_m)\, X_0^{j_0} X_1^{j_1} \ldots X_m^{j_m} = \\ = g(X_0, X_1, \ldots, X_m).$$

(4.149)

Но по условию $A_0X_0 + A_1X_1 + ... + A_mX_m = 0$. Тогда из (4.149) находим: $g(X_0, X_1, ..., X_m) = 0$, что и требовалось доказать.

Замечание 1. Теорема 5 сохраняет свою силу, если уравнение (4.145) заменить уравнениям

$$X_0A_0 + X_1A_1 + ... + X_mA_m = 0. \qquad (4.150)$$

Действительно, теорему 5 можно применить к уравнению

$$A'_0X_0 + A'_1X_1 + ... + A'_mX_m = 0$$

и потом перейти в этом уравнении почленно к транспонированным матрицам.

Замечание 2. Теорема 4 получится как частный случай теоремы 5, если в качестве $X_0, X_1, ..., X_m$ взять

$$X^m, X^{m-1} ..., X, E.$$

4.10. Извлечение корня *m*-й степени из матриц

1. Извлечение корня *m*-й степени из неособенной матрицы

Этот и следующий пункт мы посвятим уравнению

$$X^m = A, \qquad (4.151)$$

где A — заданная, а X- искомая матрицае (обе порядка n), m — данное целое положительное число.

В данном пункте мы рассмотрим случай, когда $|A| \neq 0$ (A — неособенная матрица). В этом случае все характеристические числа матрицы A отличны от нуля (ибо $|A|$ равен произведению этих характеристических чисел).

Обозначим через

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \ldots, (\lambda - \lambda_u)^{p_u} \qquad (4.152)$$

элементарные делители матрицы A и приведем матрицу A к жордановой форме (здесь $E_j = E^{(p_j)}, H_j = H^{(p_j)}$ ($j = 1,2 ...u$):

$$A = U\tilde{A}U^{-1} = U\{\lambda_1E_1 + H_1, \ldots, \lambda_uE_u + H_u\}\, U^{-1}. \qquad (4.153)$$

Так как характеристические числа искомой матрицы X при возведении в m-ю степень дают характеристические числа матрицы A, то и в матрице X все характеристические числа отличны от нуля. Поэтому на этих характеристических числах производная от

$$f(\lambda) = \lambda^{m}$$

не обращается в нуль. Но в таком случае элементарные делители матрицы X не «расщепляются» при возведении матрицы X в m-ю степень. Из сказанного следует, что элементарными делителями матрицы X будут:

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \ldots, (\lambda - \xi_u)^{p_u}, \tag{4.154}$$

где

$$\lambda_j \,(\xi_j = \sqrt[m]{\lambda_j};$$

т.е. ξ_j является одним из корней в m-й степени из

$$\lambda_j \,(\xi_j = \sqrt[m]{\lambda_j};$$

(j=1, 2,...,u).

Определим теперь

$$\sqrt[m]{\lambda_j E_j + H_j}$$

следующим образом. Возьмем в λ-плоскости круг с центром в точке λ_j, не захватывающий нуля. В этом кругу мы имеем m раздельных ветвей функции $\sqrt[m]{\lambda}$. Эти ветви можно отличать одну от другой по значениям, которые они принимают в центре круга, в точке λ_j.

Обозначим через $\sqrt[m]{\lambda}$ ту ветвь, значение которой в точке λ_j совпадает с характеристическим числом ξ_j искомой матрицы X, и, исходя из этой ветви, определим функцию от матрицы $\sqrt[m]{\lambda_j E_j + H_j}$ с помощью обрівающегося ряда

$$\sqrt[m]{\lambda_j E_j + H_j} = \lambda_j^{\frac{1}{m}} E_j + \frac{1}{m}\lambda_j^{\frac{1}{m}-1} H_j + \frac{1}{2!}\,\frac{1}{m}\left(\frac{1}{m}-1\right)\lambda_j^{\frac{1}{m}-2} H_j^2 + \ldots \tag{4.155}$$

Так как производная от рассматриваемой функции $\sqrt[m]{\lambda}$ в точке λ_j не равна нулю, то матрица (4.155) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$, где $\xi_j = \sqrt[m]{\lambda_j}$ (здесь j = 1, 2, ..., u). Отсюда следует, что квазидиагональная матрица

$$\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\}$$

имеет элементарные делители (4.154), т.е. те же элементарные делители, что и искомая матрица X. Поэтому существует такая неособенная матрица T ($|T| \neq 0$), что

$$X = T\left\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\right\} T^{-1}. \tag{4.156}$$

Для определения матрицы T заметим, что, подставляя в обе части тождества

$$(\sqrt[m]{\lambda})^m = \lambda$$

вместо λ матрицу $\lambda_j E_j + H_j$ ($j = 1, 2, \ldots, u$), получим:

$$(\sqrt[m]{\lambda_j E_j + H_j})^m = \lambda_j E_j + H_j \quad (j = 1, 2, \ldots, u).$$

Теперь из (4.151) и (4.156) следует:

$$A = T\{\lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \ldots, \lambda_u E_u + H_u\} T^{-1}. \tag{4.157}$$

Сопоставляя (4.153) и (4.157), найдем:

$$T = UX_{\tilde{A}}, \tag{4.158}$$

где $X_{\tilde{A}}$ — произвольная неособенная матрица, перестановочная с A (структура матрицы $X_{\tilde{A}}$ детально описана в п. 4.9).

Подставляя в (4.156) вместо T выражение $UX_{\tilde{A}}$, получаем формулу, которая охватывает все решения уравнения (4.151):

$$X = UX_{\tilde{A}}\left\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\right\} X_{\tilde{A}}^{-1} U^{-1}. \tag{4.159}$$

Многозначность правой части этой формулы имеет как дискретный, так и континуальный характер: дискретный (в этом случае и конечный) характер этой многозначности получается за счет выбора разных ветвей функции $\sqrt[m]{\lambda}$ в разных клетках квазидиагональной матрицы (при этом даже при $\lambda_j = \lambda_k$ ветви $\sqrt[m]{\lambda}$ в j-й и в k-й диагональных клетках могут быть различными); континуальный характер многозначности получается за счет произвольных параметров, которые содержатся в матрице $X_{\tilde{A}}$.

Все решения уравнения (4.151) мы будем называть *корнями m-й степени из матрицы A* и обозначать многозначным символом $\sqrt[m]{A}$. Обращаем внимание на то, что $\sqrt[m]{A}$ в общем случае не является функцией от матрицы A (т.е. не представляется в виде многочлена от A).

Замечание. Если все элементарные делители матрицы A попарно взаимно просты, т.е. числа $\lambda_1, \lambda_2, \ldots, \lambda_u$ все различны, то матрица $X_{\tilde{A}}$ имеет квазидиагональный вид $_t$

$$X_{\tilde{A}} = \{X_1, X_2, \ldots, X_u\},$$

где матрица X_j перестановочна с $\lambda_j E_j + H_j$ и, следовательно, перестановочна с любой функцией от $\lambda_j E_j + H_j$, в частности с

$$\sqrt[m]{\lambda_j E_j + H_j} \quad (j = 1, 2, \ldots, u).$$

Поэтому в рассматриваемом случае формула (4.159) принимает вид

$$X = U\left\{\sqrt[m]{\lambda_1 E_1 + H_1},\ \sqrt[m]{\lambda_2 E_2 + H_2},\ \ldots,\ \sqrt[m]{\lambda_u E_u + H_u}\right\} U^{-1}.$$

Таким образом, если элементарные делители матрицы A попарно взаимно просты, то в формуле для $X = \sqrt[m]{A}$ имеется только дискретная многозначность. В этом случае любое значение $\sqrt[m]{A}$ можно представить как многочлен от A.

Пример. Пусть нужно найти все квадратные корни из матрицы

$$A = \begin{Vmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix},$$

т.е. все решения уравнения

$$X^2 = A.$$

В этом случае матрица A уже имеет нормальную жорданову форму. Поэтому в формуле (4.159) можно положить $A = \tilde{A}$, $U=E$. Матрица $X_{\tilde{A}}$ в этом случае выглядит так:

$$X_{\tilde{A}} = \begin{Vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{Vmatrix},$$

где a, b, c, d, e — произвольные параметры.

Формула (4.159), дающая все искомые решения X, в этом случае принимает следующий вид:

$$X = \begin{Vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{Vmatrix} \begin{Vmatrix} \varepsilon & \frac{\varepsilon}{2} & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \eta \end{Vmatrix} \begin{Vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{Vmatrix}^{-1} \qquad (\varepsilon^2 = \eta^2 = 1). \tag{4.160}$$

Не изменяя X, мы можем в формуле (4.159) умножить $X_{\tilde{A}}$ на такой скаляр, чтобы $| X_{\tilde{A}} | = 1$. В этом случае это приведет к равенству $a^2 e = 1$, откуда $e = a^{-2}$.

Вычислим элементы матрицы $X^{-1}_{\tilde{A}}$. Для этого выпишем линейное преобразование с матрицей коэффициентов $X_{\tilde{A}}$:

$$y_1 = ax_1 + bx_2 + cx_3,$$
$$y_2 = ax_2,$$
$$y_3 = dx_2 + a^{-2}x_3.$$

Разрешим эту систему уравнений относительно x_1, x_2, x_3. Тогда получим преобразование с обратной матрицей $X^{-1}_{\tilde{A}}$:

$$x_1 = a^{-1}y_1 — (a^{-2}b — cd)y_2 — acy_3,$$
$$x_2 = a^{-1}y_2,$$
$$x_3 = —ady_2 + a^2 y_3.$$

Отсюда находим:

$$X_{\tilde{A}}^{-1} = \left\| \begin{matrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{matrix} \right\|^{-1} = \left\| \begin{matrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^2 \end{matrix} \right\|.$$

Формула (4.160) дает:

$$X = \left\| \begin{matrix} \varepsilon & (\varepsilon - \eta)\, a\, cd + \dfrac{\varepsilon}{2} & a^2 c\, (\eta - \varepsilon) \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta)\, da^{-1} & \eta \end{matrix} \right\| =$$

$$= \left\| \begin{matrix} \varepsilon & (\varepsilon - \eta)\, vw + \dfrac{\varepsilon}{2} & (\eta - \varepsilon)\, v \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta)\, w & \eta \end{matrix} \right\| \quad (v = a^2 c;\ w = a^{-1} d).$$

(4.161)

Решение X зависит от двух произвольных параметров v и w и от двух произвольных знаков ε и η).

2. Извлечение корня *m*-й степени из особенной матрицы. Переходим к разбору случая, когда $| A | = 0$ (A — особенная матрица). Как и в первом случае, приведем матрицу A к нормальной жордановой форме:

$$A = U \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \ldots, \lambda_u E^{(p_u)} + H^{(p_u)};\ \ H^{(q_1)}, H^{(q_2)}, \ldots, H^{(q_t)}\}\, U^{-1};$$

(4.162)

здесь мы через

$$(\lambda - \lambda_1)^{p_1}, \ \ldots, \ (\lambda - \lambda_u)^{p_u}$$

обозначили элементарные делители матрицы A, отвечающие ненулевым характеристическим числам, а через $\lambda^{q_1}, \lambda^{q_2}, \ldots, \lambda^{q_t}$ - элементарные делители с нулевыми характеристическими числами.

Тогда

$$A = U\{A_1, A_2\}U^{-1} \qquad (4.163)$$

где

$$A_1 = \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \ldots, \lambda_u E^{(p_u)} + H^{(p_u)}\}, A_2 = \{H^{(q_1)}, H^{(q_2)}, \ldots, H^{(q_t)}\}. \qquad (4.164)$$

Заметим, что A_1 — неособенная матрица ($|A_1| \neq 0$), а A_2 — нильпотентная матрица с индексом нильпотентности

$$\mu = \max(q_1, q_2, \ldots, q_t) \ (A_2^\mu = 0).$$

Из исходного уравнения (4.151) следует перестановочность матрицы A с искомой матрицей X, а следовательно, и перестановочность подобных им матриц

$$U^{-1}AU = \{A_1, A_2\} \text{ и } U^{-1}XU. \qquad (4.165)$$

Как было доказано в п. 4.9 (теорема 3), из перестановочности матриц (4.165) и из того факта, что матрицы A_1 и A_2 не имеют общих характеристических чисел, следует, что и вторая из матриц (4.165) имеет соответствующую квазидиагональную форму

$$U^{-1}XU = \{X_1, X_2\}. \qquad (4.166)$$

Заменяя в уравнении (4.151) матрицы A и X подобными им матрицами

$$\{A_1, A_2\} \text{ и } \{X_1, X_2\},$$

мы заменим уравнение (4.151) двумя уравнениями:

$$X^m{}_1 = A_1 \qquad (4.167)$$

$$X^m{}_2 = A_2. \qquad (4.168)$$

Так как $|A_1| \neq 0$, то к уравнению (4.167) применимы результаты предыдущего пункта. Поэтому X_1 находим по формуле (4.159):

$$X_1 = X_{A_1}\left\{\sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \ldots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}}\right\} X_{A_1}^{-1}. \qquad (4.169)$$

Таким образом, остается рассмотреть уравнение (4.168), т.е. заняться нахождением всех корней m-й степени из нильпотентной матрицы A_2, уже имеющей нормальную жорданову форму:

$$A_2 = \{H^{(q_1)}, H^{(q_2)}, \ldots, H^{(q_t)}\}. \qquad (4.170)$$

$\mu = \max(q_1, q_2, \ldots, q_t)$ — индекс нильпотентности матрицы A_2. С $A^\mu{}_2 = 0$ и из (4.168) находим:

$$A_2^{m\mu} = 0.$$

Последнее равенство показывает, что искомая матрица X_2 также является нильпотентной с индексом нильпотентности v, где $m(\mu — 1)<v\leq m\mu$. Приведем матрицу X_2 к жордановой форме:

$$X_2 = T\{H^{(v_1)}, H^{(v_2)}, \ldots, H^{(v_s)}\}T^{-1}$$

$$(v_1, v_2, \ldots, v_s \leqslant \nu). \tag{4.171}$$

Возведем теперь обе части последнего равенства в m-ю степень. Получим:

$$A_2 = X_2^m = T\{[H^{(v_1)}]^m, [H^{(v_2)}]^m, \ldots, [H^{(v_s)}]^m\}T^{-1}. \tag{4.172}$$

Выясним теперь, какие элементарные делители имеет матрица $[H^{(v)}]^m$. Обозначим через $\boldsymbol{H}$ линейный оператор, который задается матрицей $H^{(v)}$ в v-мерном векторном пространстве с базисом $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_v$. Тогда из вида матрицы $H^{(v)}$ (в матрице $H^{(v)}$ все элементы первой наддиагонали равны единице и все остальные элементы равны нулю) следует, что

$$\boldsymbol{He}_1 = 0, \quad \boldsymbol{He}_2 = \boldsymbol{e}_1, \ldots, \boldsymbol{He}_v = \boldsymbol{e}_{v-1}. \tag{4.173}$$

Эти равенства показывают, что для оператора $\boldsymbol{H}$ векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_v$ образуют жорданову цепочку векторов, соответствующую элементарному делителю λ^v.

Равенства (4.173) запишем так:

$$\boldsymbol{He}_j = \boldsymbol{e}_{j-1} \quad (j = 1, 2, \ldots, v;\ \boldsymbol{e}_0 = 0).$$

Очевидно, что

$$\boldsymbol{H}^m\boldsymbol{e}_j = e_{j-m} \quad (j = 1, 2, \ldots, v;\ \boldsymbol{e}_0 = \boldsymbol{e}_{-1} = \ldots = \boldsymbol{e}_{-m+1} = 0). \tag{4.174}$$

Представим число v в виде

$$v = km + r \qquad (r < m),$$

где k, r — целые неотрицательные числа. Расположим базисные векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_v$ следующим образом:

$$\begin{array}{lll}
\boldsymbol{e}_1, & \boldsymbol{e}_2, & \ldots, \boldsymbol{e}_m, \\
\boldsymbol{e}_{m+1}, & \boldsymbol{e}_{m+2}, & \ldots, \boldsymbol{e}_{2m}, \\
\cdots & \cdots & \cdots \\
\boldsymbol{e}_{(k-1)m+1}, & \boldsymbol{e}_{(k-1)m+2}, & \ldots, \boldsymbol{e}_{km}, \\
\boldsymbol{e}_{km+1}, \ \ldots, & \boldsymbol{e}_{km+r}. &
\end{array} \tag{4.175}$$

В этой таблице мы имеем m столбцов: первые r содержат по $(k + 1)$ векторов в каждом, остальные — по k векторов. Равенство (4.175) показывает, что векторы каждого столбца образуют жорданову цепочку векторов относительно оператора $\boldsymbol{H}^m$**.** Если вместо последовательной нумерации векторов (4.175) по строкам занумеровать их по столбцам, то в полученном таким образом новом базисе матрица оператора $\boldsymbol{H}^m$ будет иметь следующую нормальную жорданову форму:

$$\{\underbrace{H^{(k+1)}, \ldots, H^{(k+1)}}_{r}, \quad \underbrace{H^{(k)}, \ldots, H^{(k)}}_{m-r}\}$$

(В случае k=0 клетки $\underbrace{H^{(k)}, \ldots, H^{(k)}}_{m-r}$ отсутствуют, и матрица имеет вид $\underbrace{\{H^{(1)}, \ldots, H^{(1)}\}}_{r}$)

и, следовательно,

$$[H^{(v)}]^m = P_{v,m}\{\underbrace{H^{(k+1)}, \ldots, H^{(k+1)}}_{r}, \quad \underbrace{H^{(k)}, \ldots, H^{(k)}}_{m-r}\} P_{v,m}^{-1},$$

(4.176)

где матрица $P_{v,m}$ (матрица перехода от одного базиса к другому) имеет вид

$$P_{v,m} = \left\| \begin{matrix} \overbrace{1 \quad 0 \;\ldots\; 0}^{m} & 0 & \ldots \\ 0 \quad 0 \;\ldots\; 0 & 1 & \ldots \\ \cdots\cdots\cdots & \cdots & \cdots \\ \cdots\cdots\cdots & \cdots & \cdots \\ 0 \quad 0 \;\ldots\; 0 & & \\ 0 \quad 1 \;\ldots\; 0 & & \\ \cdots\cdots\cdots & \cdots & \cdots \end{matrix} \right\| \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} m \\ \\ \\ \\ \end{matrix} .$$

(4.177)

Матрица $H^{(v)}$ имеет один элементарный делитель $\lambda^{(v)}$. При возведении матрицы $H^{(v)}$ к m-ю степень этот элементарный делитель «расщепляется». Как показывает формула (4.176), матрица $[H^{(v)}]^m$ имеет элементарные делители:

$$\underbrace{\lambda^{k+1}, \ldots, \lambda^{k+1}}_{r}, \quad \underbrace{\lambda^{k}, \ldots, \lambda^{k}}_{m-r}.$$

Возвращаясь теперь к равенству (4.172), положим:

$$v_i = k_i m + r_i \quad (0 \leqslant r_i < m, \quad k_i \geqslant 0;\ i = 1, 2, \ldots, s).$$

(4.178)

Тогда в силу (4.176) равенство (4.172) перепишется так:

$$A_2 = X_2^m = TP\{\underbrace{H^{(k_1+1)}, \ldots, H^{(k_1+1)}}_{r_1}, \quad \underbrace{H^{(k_1)}, \ldots, H^{(k_1)}}_{m-r_1},$$

$$\underbrace{H^{(k_2+1)}, \ldots, H^{(k_2+1)}}_{r_2}, H^{(k_2)}, \ldots\} P^{-1}T^{-1},$$

(4.179)

где

$$P = \{P_{v_1,m}, P_{v_2,m}, \ldots, P_{v_s,m}\}.$$

Сопоставляя (4.179) с (4.170), видим, что клетки

$$H^{(k_1+1)}, \ldots, H^{(k_1+1)}, H^{(k_1)}, \ldots, H^{(k_1)}, H^{(k_2+1)}, \ldots, H^{(k_2+1)}, \ldots$$

(4.180)

с точностью до порядка должны совпасть с клетками

$$H^{(q_1)}, H^{(q_2)}, \ldots, H^{(q_t)}.$$

(4.181)

Условимся систему элементарных делителей $\lambda^{v_1}, \lambda^{v_2}, \ldots, \lambda^{v_s}$ называть *возможной* для X_2, если после возведения матрицы в m-ю степень эти элементарные делители, расщепляясь, порождают заданную систему элементарных делителей матрицы A_2: $\lambda^{q_1}, \lambda^{q_2}, \ldots, \lambda^{q_t}$. Число возможных систем элементарных делителей всегда конечно, поскольку

$$\max(v_1, v_2, \ldots, v_s) \leqslant m\mu, \quad v_1 + v_2 + \ldots + v_s = n_2$$

(4.182)

(n_2 — степень матрицы A_2).

В каждом конкретном случае возможные системы элементарных делителей для X_2 могут быть легко определенны путем конечного числа испытаний.

Покажем, что для каждой возможной системы элементарных делителей $\lambda^{v_1}, \lambda^{v_2}, \ldots, \lambda^{v_s}$ существуют соответствующие решения уравнения (4.168), и определим все эти решения. В этом случае существует преобразующая матрица Q такая, что

$$\{H^{(k_1+1)}, \ldots, H^{(k_1+1)}, H^{(k_1)}, \ldots, H^{(k_1)}, H^{(k_2+1)}, \ldots\} = Q^{-1}A_2Q.$$

(4.183)

Матрица Q осуществляет перестановку клеток в квазидиагональной матрице, что достигается надлежащей перенумерацией базисных векторов. Поэтому матрицу Q можно считать известной. Используя (4.183), мы из (4.179) получим:

$$A_2 = TPQ^{-1}A_2QP^{-1}T^{-1}.$$

Отсюда

$$TPQ^{-1}=X_{A_2},$$

или

$$T = X_{A_2}QP^{-1}, \tag{4.184}$$

где X_{A_2} — произвольная матрица, перестановочная с A_2.

Подставляя выражение (4.184) для T в (4.171), будем иметь:

$$X_2 = X_{A_2}QP^{-1}\left\{H^{(v_1)}, H^{(v_2)}, \ldots, H^{(v_s)}\right\} PQ^{-1}X_{A_2}^{-1}. \tag{4.185}$$

Из (4.166), (4.169) и (4.185) получим общую формулу, охватывающую все искомые решения:

$$X = U\left\{X_{A_1}, X_{A_2}QP^{-1}\right\}\left\{\sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \ldots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}},\right.$$
$$\left. H^{(v_1)}, \ldots, H^{(v_s)}\right\} \cdot \left\{X_{A_1}^{-1}, PQ^{-1}X_{A_2}^{-1}\right\} U^{-1}. \tag{4.186}$$

Обращаем внимание читателя на то, что корень m-й степени из особенной матрицы не всегда существует. Его существование связано с существованием системы возможных элементарных делителей для матрицы X_2.

Легко видеть, например, что уравнение

$$X^m = H^{(p)}$$

не имеет решений при $m > 1$, $p > 1$.

Пример. Требуется извлечь корень квадратный из матрицы

$$A = \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix},$$

т.е. найти все решения уравнения

$$X^2=A.$$

В этом случае $A = A_2$, $X = X_2$, $m = 2$, $t= 2$, $q_1 = 2$, $q_2 = 1$. Матрица X может иметь только один элементарный делитель λ^3. Поэтому $s = 1$, $v_1 = 3$, $k_1 = 1$, $r_1 = 1$ и [см. (4.177)]

$$P = P_{3,2} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{Vmatrix} = P^{-1}, \quad Q = E.$$

Кроме того, можно положить в формуле (4.185):

$$X_{A_1} = \begin{Vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{Vmatrix}, \quad X_{A_1}^{-1} = \begin{Vmatrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^2 \end{Vmatrix}.$$

Из этой формулы получим:

$$X = X_2 = X_{A_1} P^{-1} H^{(3)} P X_{A_1}^{-1} = \begin{Vmatrix} 0 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & \beta^{-1} & 0 \end{Vmatrix},$$

где $\alpha = ca^{-1} - a^2 d$ и $\beta = a^3$ — произвольные параметры.

4.11. Логарифм матрицы

1. Рассмотрим матричное уравнение

$$e^X = A. \qquad (4.187)$$

Все решения этого уравнения будем называть логарифмами (натуральными) матрицы A и обозначать через $\ln A$.

Характеристические числа λ_j матрицы A связаны с характеристическими числами ξ_j матрицы X формулой $\lambda_j = e^{\xi_j}$; поэтому, если уравнение (4.187) имеет решение, то все характеристические числа матрицы A отличны от нуля и матрица A является неособенной ($|A| \neq 0$). Таким образом, условие $|A| \neq 0$ является необходимым для существования решения уравнения (4.187). Ниже мы увидим, что это условие является и достаточным.

Итак, пусть $|A| \neq 0$. Выпишем элементарные делители матрицы A:

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \ldots, (\lambda - \lambda_u)^{p_u}$$
$$(\lambda_1 \lambda_2 \ldots \lambda_u \neq 0, p_1 + p_2 + \ldots + p_u = n). \qquad (4.188)$$

В соответствии и этими элементарными делителями приведем матрицу A к нормальной жордановой форме:

$$A = U\tilde{A}U^{-1} =$$
$$= U\{\lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \ldots, \lambda_u E^{(p_u)} + H^{(p_u)}\} U^{-1}. \qquad (4.189)$$

Так как производная от функции e^ξ отлична от нуля при всех значениях ξ, то при переходе от матрицы X к матрице $A = e^X$ элементарные делители не расщепляются, т.е. матрица X имеет элементарные делители

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \ldots, (\lambda - \xi_u)^{p_u}, \qquad (4.190)$$

где $e^{\xi_j} = \lambda_j$ $(j = 1,2, ..., u)$, т.е. ξ_j есть одно из значений $\ln \lambda_j$ $(j = 1,2,...,u)$. Возьмем в плоскости комплексного переменного λ круг с центром в точке λ_j радиуса $< |\lambda_j|$ и обозначим через $f_j(\lambda) = \ln \lambda$ ту из ветвей функции $\ln \lambda$ в рассмотренном круге, которая в точке λ_j принимает значение, равное характеристическому числу ξ_j матрицы X $(j = 1,2, ...,u)$. После этого полагаем:

$$\ln (\lambda_j E^{(p_j)} + H^{(p_j)}) = f_j(\lambda_j E^{(p_j)} + H^{(p_j)}) = \ln \lambda_j E^{(p_j)} + \lambda_j^{-1} H^{(p_j)} + \dots \tag{4.191}$$

Так как производная от $\ln \lambda$ нигде не обращается в нуль (в конечной части плоскости λ), то матрица (4.191) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$. В силу этого квазидиагональная матрица

$$\{\ln (\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln (\lambda_2 E^{(p_2)} + H^{(p_2)}), \ \dots, \ \ln (\lambda_u E^{(p_u)} + H^{(p_u)})\} \tag{4.192}$$

имеет те же элементарные делители, что и искомая матрица X. Поэтому существует такая матрица $T(|T| \neq 0)$, что

$$X = T \{\ln (\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln (\lambda_u E^{(p_u)} + H^{(p_u)}\} T^{-1}. \tag{4.193}$$

Для определения матрицы T заметим, что

$$A = e^X = T \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)}\} T^{-1}. \tag{4.194}$$

Сопоставляя (4.194) с (4.189), находим:

$$T = UX_{\tilde{A}}, \tag{4.195}$$

где $X_{\tilde{A}}$ — произвольная матрица, перестановочная с матрицей $\tilde{A}$. Подставляя выражение для T из (4.195) в (4.193), получим общую формулу, которая охватывает все логарифмы матрицы:

$$X = UX_{\tilde{A}} \{\ln (\lambda_1 E^{(p_1)} + H^{(p_1)}),$$
$$\ln (\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln (\lambda_u E^{(p_u)} + H^{(p_u)})\} X_{\tilde{A}}^{-1} U^{-1}. \tag{4.196}$$

Замечание. Если все элементарные делители матрицы A взаимно просты, то в правой части формулы (4.196) можно выбросить множители $X_{\tilde{A}}$ и $X_{\tilde{A}}^{-1}$.

2. Выясним, когда вещественная неособенная матрица A имеет вещественный логарифм X. Пусть искомая матрица имеет несколько элементарных делителей, отвечающих характеристическому числу вида

$$\rho + i\pi:\ (\lambda - \rho - i\pi)^{q_1}, \ldots, (\lambda - \rho - i\pi)^{q_t}.$$

Поскольку матрица X вещественна, то она имеет и сопряженные элементарные делители:

$$(\lambda - \rho - i\pi)^{q_1}, \ldots, (\lambda - \rho - i\pi)^{q_t}.$$

При переходе от матрицы X к матрице A элементарные делители не расщепляются, но характеристические числа $\rho+i\pi$, $\rho-i\pi$ заменяются в них числами $e^{\rho+i\pi}$ = —μ, $e^{\rho-i\pi}$ = —μ, где $\mu=e^{\rho}>0$. Поэтому в системе элементарных делителей матрицы A каждый элементарный делитель, соответствующий отрицательному характеристическому числу (если такие существуют), повторяется четное число раз. Докажем теперь, что это необходимое условие является и достаточным, т.е. что *вещественная неособенная матрица A тогда и только тогда имеет вещественный логарифм X, когда у матрицы A либо совсем нет элементарных делителей, соответствующих отрицательным характеристическим числам* (в этом случае существует вещественный $\ln A = r(A)$, где $r(\lambda)$ — надлежащий интерполяционный многочлен для $\ln_0 \lambda$ (см. р. 3.8.), *либо каждый такой элементарный делитель повторяется четное число раз* (это условие, в частности, выполняется, когда $A=B^2$, где B — вещественная матрица).

Действительно, пусть это условие выполнено. Тогда в квазидиагональной матрице (4.192) в соответствии с формулой (4.191) в тех клетках, где λ_i - вещественно и положительно, возьмем для $\ln \lambda_i$ вещественное значение; если же в какой-нибудь клетке имеется комплексное λ_n, то найдется другая клетка такого же размера с $\lambda_g = \overline{\lambda}_n$.
В этих клетках возьмем комплексно сопряженные значения для
$\ln \lambda_n$ и $\ln \lambda_q$. Каждая же клетка по условию повторяется в (4.195) четное число раз с сохранением размера клетки. Тогда в половине этих клеток положим $\ln \lambda_k = \ln|\lambda_k| + i\pi$, а в другой половине возьмем $\ln \lambda_k = \ln|\lambda_k| - i\pi$. Тогда в квазидиагональной матрице (4.195) диагональные клетки либо будут вещественными, либо будут попарно комплексно сопряженными. Но *такая квазидиагональная матрица всегда подобна вещественной матрице*.
(Для того чтобы убедиться в этом, достаточно показать, что квазидиагональная матрица

$$D = \begin{pmatrix} B & 0 \\ 0 & \overline{B} \end{pmatrix}$$

всегда подобна некоторой вещественной матрице. Здесь $B= U+i$, $\overline{B} = U — iV$, где U *и* V— вещественные матрицы. Обозначая через E единичную матрицу тех же размеров, что и B, и полагая

$$T = \begin{pmatrix} E & iE \\ E & -iE \end{pmatrix},$$

легко проверим, что

$$T^{-1} = \begin{pmatrix} \frac{1}{2}E & \frac{1}{2}E \\ \frac{1}{2i}E & -\frac{1}{2i}E \end{pmatrix}, \quad T^{-1}DT = \begin{pmatrix} U & -V \\ V & U \end{pmatrix}.$$

).

Поэтому существует такая неособенная матрица T_1 ($|T_1| \neq 0$), что матрица

$$X_1 = T_1 \{\ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ldots, \ln(\lambda_n E^{(p_n)} + H^{(p_n)})\} T_1^{-1}$$

вещественна. Но тогда будет вещественной и матрица

$$A_1 = e^{X_1} = T_1 \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \ldots, \lambda_n E^{(p_n)} + H^{(p_n)}\} T^{-1}. \tag{4.197}$$

Сопоставляя формулу (4.197) с формулой (4.189), заключаем, что матрицы A и A_1 подобны между собой (поскольку они подобны одной и той же жордановой матрице). Но две подобные вещественные матрицы могут быть преобразованы друг в друга с помощью некоторой неособенной вещественной матрицы $W(|W| \neq 0)$:

$$A = WA_1W^{-1} = We^{X_1}W^{-1} = e^{WX_1W^{-1}}.$$

Тогда матрица $X= WX_1W^{-1}$ и будет искомым вещественным логарифмом матрицы A.

Микромодуль 12
Пространство переменных состояний

4.12. Переменные состояния.

При рассмотрении физической системы как объекта исследования или проектирования целесообразно распределить все переменные, характеризующиет систему или имеющие к ней какое-либо отношение, на три множества:

1) *входные переменные* $v_1, v_2, ..., v_m,$ характеризующие внешние воздействия на входы системы;

2) *переменные состояния* x_1 x_2, ..., x_n — внутренние (промежуточные) переменные, совокупность которых полностью характеризует свойства системы;

3) *выходные переменные* $y_1, y_2, ..., y_r$, представляющие те реакции на внешние воздействия и те состояния системы, которые интересны для исследователя или конструктора.

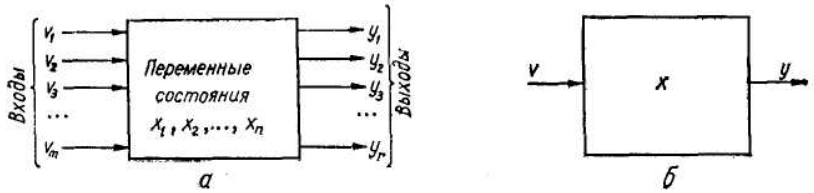


Рис. 4.1. Система с *m* входами и *r* выходами (*а*) и ее обобщенное представление (*б*).

После упорядочения (нумерации) элементов этих множеств получаем соответственно три вектора: входной (задающий) вектор $v = (v_1, v_2, ..., v_m)$, вектор состояний $x = (x_1, x_2, ..., x_n)$ и выходной вектор (вектор реакций) $y = (y_1, y_2, ..., y_r)$. Сама система в общем виде представляется «черным ящиком» с *m* входами и *r* выходами, с каждым из которых связана соответствующая переменная (рис. 4.1, *а*). Можно рассматривать совокупность входов как один обобщенный вход, на который воздействует входной вектор *v,* а совокупность выходов — как обобщенный выход, который характеризуется выходным вектором *у* (рис. 4.1, *б*). Переменные состояния связаны с внутренними свойствами системы и поэтому указываются внутри «черного ящика».

Собственно система, ее входы и выходы — это три взаимозависимых объекта, которые в каждой конкретной ситуации определяются соответственно описанием системы (структура и свойства компонент или математическая модель системы), а также заданием множеств входных и выходных переменных. В зависимости от того, какой из этих объектов подлежит определению (при остальных двух заданных) различают три типа задач исследования и проектирование систем:

Тип задачи	Входы	Система	Выходы
Анализ	×	×	?
Синтез	×	?	×
Измерение	?	×	×

Решение любой из этих задач непосредственно связано с исследованием состояний системы, множество которых образует пространство состояний.

1. Основные уравнения. Непрерывные детерминированные системы в каждый момент времени t можно описать парой матричных уравнений:

$$\frac{dx(t)}{dt} = F[x(t), \quad v(t)]; \quad y(t) = \varphi[x(t), \quad v(t)].$$

Первое из них является *уравнением состояния* системы, решение которого, удовлетворяющее начальному условию $x_0 = x(t_0)$, дает вектор состояния $x(t) = \psi[x(t_0), \quad v(t)]$.

Второе уравнение определяет выходные переменные в зависимости от $x(t)$ и $v(t)$, и потому оно называется *выходным уравнением.*

В частных случаях эти уравнении принимают специфическую форму в соответствии со свойствами системы. Для линейных систем имеем:

$$\frac{dx(t)}{dt} = A(t)x(t) + B(t)v(t);$$
$$y(t) = C(t)x(t) + D(t)v(t),$$

где A — *матрица системы* (квадратная n-го порядка); $B(t)$ — *матрица управления* размера $(n \times m)$; $C(t)$ — *матрица выхода* размера $r \times n$ и $D(t)$ — *матрица входа* размера $r \times m$.

Если элементы этих матриц зависят от времени t, то система называется *линейной нестационарной* (или *параметрической).* Для *линейных стационарных* систем (часто их называют просто *линейными системами)* элементы матриц A, B, C, D выражаются постоянными числами, которые являются функциями параметров компонент системы.

Наиболее сложную структуру имеют уравнения *нелинейных систем,* компоненты которых характеризуются нелинейными зависимостями между переменными на их входах и выходах. В ряде практически важных случаев *уравнения состояния нелинейной системы* можно представить в виде:

$$\frac{dx(t)}{dt} = Ax(t) + Bv(t) + Fz(t); \quad f[x(t), \quad z(t), \quad v(t)] = 0,$$

где A, B и F — постоянные матрицы; $f(x, z, v) = 0$ — нелинейное алгебраическое уравнение, решение которого относительно вектора z позволяет исключить этот вектор из дифференциального уравнения. В более сложных случаях элементы матриц A, B и F могут зависеть от состояния системы.

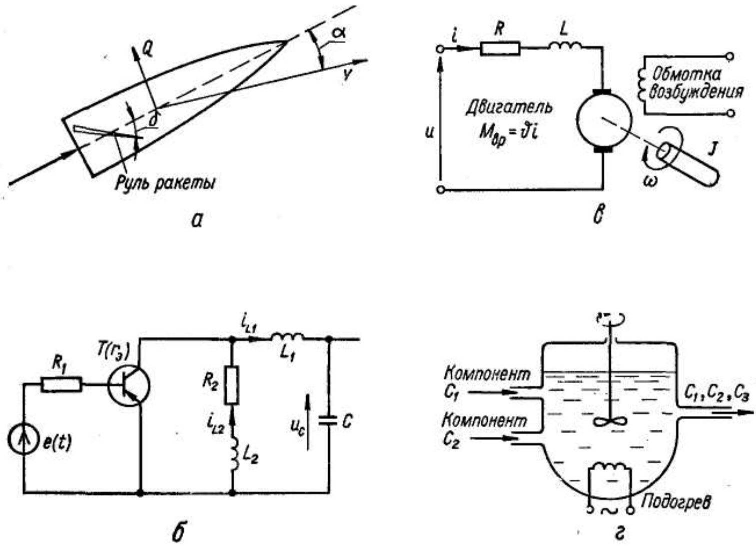


Рис. 4.2. Примеры физических систем:
а — управляемая ракета; *б* — электронная схема; *в* — двигатель постоянного тока; *г* — химический реактор.

2. Примеры физических систем. Математическое моделирование в пространстве переменных состояний широко используется при исследовании физических систем. Приведем несколько примеров, которые относятся к системам различной природы.

Управляемая ракета (рис. 4.2, *а)* при малых значениях угла атаки α и угла отклонения управляемой поверхности δ описывается линеаризированной системой уравнений:

$$\frac{d}{dt}\begin{bmatrix}\omega\\ \alpha\end{bmatrix}=\begin{bmatrix}0 & \frac{Q}{J}C_{m\alpha}\\ 1 & -\frac{1}{mv}(QC_{L\alpha}+T)\end{bmatrix}\begin{bmatrix}\omega\\ \alpha\end{bmatrix}+\begin{bmatrix}\frac{Q}{J}C_{m\delta}\\ -\frac{Q}{mv}C_{L\delta}\end{bmatrix}\delta;$$

$$\alpha_N=\frac{QC_{N\alpha}}{mg}\alpha+\frac{QC_{N\delta}}{mg}\delta .$$

Переменными состояния являются угловая скорость ω и угол атаки α, а задающой величиной - угол отклонения руля δ. Поперечное ускорение ракеты α_N (в единицах *g)* является выходной величиной. Момент инерции *J* относительно оси тангажа, масса *m* и скорость *v* ракеты предполагаются постоянными. Динамические коэффициенты $C_{m\alpha}$, $C_{m\delta}$, $C_{L\alpha}$, $C_{L\delta}$, $C_{N\alpha}$, $C_{N\delta}$ также постоянны (их значения получают в результате аэродинамических продувок).

Электронная схема (транзисторный усилитель со сложной коррекцией, рис. 4.2, *б)* в квазилинейном режиме на низких частотах описывается системой уравнений:

$$\frac{d}{dt}\begin{bmatrix}u_C\\ i_{L1}\\ i_{L2}\end{bmatrix}=\begin{bmatrix}0 & \frac{1}{C} & 0\\ -\frac{1}{L_1} & -\frac{r}{L_1} & -\frac{r}{L_1}\\ 0 & -\frac{r}{L_2} & -\frac{1}{L_2}(r+R_2)\end{bmatrix}\begin{bmatrix}u_C\\ i_{L1}\\ i_{L2}\end{bmatrix}+$$

$$+\begin{bmatrix}0\\ -\frac{r_{21}}{L_1(r_{11}+R_1)}\\ -\frac{r_{21}}{L_2(r_{11}+R_1)}\end{bmatrix}e(t),$$

где

$$r=\frac{(r_{11}+R_1)\,r_{22}-r_{12}r_{21}}{r_{11}+R_1}.$$

Переменными состояния являются напряжение u_C на емкости и токи i_{L1}, i_{L2} в индуктивностях, а задающей функцией, - напряжение *e*(*t*) источника на входе. Низкочастотные параметры транзистора r_{11}, r_{12}, r_{21} и r_{22} определены по схеме с общим емиттером.

Двигатель постоянного тока с независимым возбуждением (рис. 4.2, *в)* в предположении, что поток возбуждения постоянен, описывается уравнениями:

$$\frac{d}{dt}\begin{bmatrix}\omega\\ i\end{bmatrix}=\begin{bmatrix}0 & \frac{\theta}{J}\\ -\frac{\lambda}{L} & -\frac{R}{L}\end{bmatrix}\begin{bmatrix}\omega\\ i\end{bmatrix}+\begin{bmatrix}0\\ \frac{1}{L}\end{bmatrix}u.$$

Здесь переменные состояния — угловая скорость двигателя ω и ток i в цепи якоря, а задающая величина - напряжение u на входе двигателя. Параметры двигателя определяются сопротивлением R и индуктивностью L цепи якоря, моментом инерции J и коэффициентом пропорциональности θ между вращающим моментом и током i в цепи якоря.

Химический реактор с мешалкой и подогревом (рис. 4.2, *г),* в котором протекают одновременно две реакции: экзотермическая $C_1 \to C_2$ со скоростью v_1 и теплотой h_1 и эндотермическая $C_2 \to C_3$ со скоростью v_2 и теплотой h_2, описывается системой нелинейных дифференциальных уравнений:

$$\left.\begin{aligned}\frac{dx_1}{dt} &= v_1-(v_1-v_2-\nu_1)x_1\\ \frac{dx_3}{dt} &= \nu_1(1-x_1-x_3)-(v_1+v_2)x_3\\ \frac{d\tau}{dt} &= h_1\nu x_1+h_2\nu_2(1-x_1-x_3)+v_1(\tau_1-\tau)+v_2(\tau_2-\tau)+c\end{aligned}\right\}.$$

Переменные состояния x_1 и x_3 означают весовые доли компонентов C_1 и C_3, которые связаны с весовой долью x_2 компоненты C_2 зависимостью $x_1 + x_2 + x_3 = 1$. Третья переменная состояния t — температура реактора. Уравнение составлено в предположении, что компоненты C_1, C_2, C_3 имеют одинаковые теплоемкости и плотности, молекулярный вес компонентов в ходе реакции не изменяется. Скорости v_1 и v_2 поступление компонент C_1 и C_2 соответственно и их температуры τ_1 и τ_2 являются задающими воздействиями, а c - скорость теплообмена реактора. Скорости реакций выражаются экспоненциальными функциями

$$\nu_1 = x_1 \exp\frac{\alpha_1-\beta_1}{\tau} \quad \text{и} \quad \nu_2 = x_2 \exp\frac{\alpha_2-\beta_2}{\tau},$$

где α_1, α_2, β_1, β_2 — константы.

Обозначив вектор переменных состояния $x = (x_1, x_3, \tau)$ и вектор воздействий — $(v_1, v_2, \tau_1, \tau_2)$, нелинейное уравнение можно записать в матричной форме через вектор-функцию f:

$$\frac{dx}{dt} = f(x, v, t).$$

Приведенные примеры иллюстрируют высокую степень всеобщности представления физических систем в пространстве переменных состояния.

3. Линейные (стационарные) системы. Уравнение состояния линейной стационарной системы $\dot{x}(t) = Ax + Bv$ представляет собой матричную запись системы дифференциальных уравнений с постоянными коэффициентами в нормальной форме. Его решение, которое удовлетворяет начальным условиям $x_0 = x(0)$, для вектора состояния и выходного вектора имеет вид:

$$x(t) = \Phi(t)\, x(0) + \int_0^t \Phi(t-\tau)\, Bv(\tau)\, d\tau;$$

$$y(t) = C\Phi(t)\, x(0) + \int_0^t C\Phi(t-\tau)\, Bv(\tau)\, d\tau + Dv(t).$$

Первые слагаемые соответствуют реакции, которая зависит от начальных условий (свободное движение системы), а остальные слагаемые - реакции на входные воздействия (вынужденное движение). Фундаментальная матрица

$$\Phi(t) = e^{At} = \exp At$$

называется *переходной матрицей состояния* системы. Она осуществляет линейное преобразование, которое переводит начальное состояние $x(0)$ системы в некоторое состояние для момента времени t (при нулевых входах), т.е.

$$x(t) = \Phi(t)\, x(0).$$

При нулевых начальных условиях $x(0)=0$ и $D=0$ связь между реакцией на выходах и воздействими на входах описывается соотношением

$$y(t) = \int_0^t C\Phi(t-\tau)\, Bv(\tau)\, d\tau = \int_0^t g(t-\tau)\, v(\tau)\, d\tau.$$

Матрица $g(t) = C\Phi(t)B$ представляет собой обобщенную характеристику системы относительно ее входов и выходов. Реакция на i-м выходе

$$y_i(t) = \int_0^t [g_{i1}(t-\tau) v_1(\tau) + g_{i2}(t-\tau) v_2(\tau) + \cdots +$$
$$+ g_{1m}(t-\tau) v_m(\tau)]\, d\tau,$$

где $g_{ij}(t)$ — ij-элемент матрицы $g(t)$. Каждый член подинтегральной функции отражает вклад соответствующего входного воздействия и равен реакции $y_{ij}(t)$ на i-м выходе относительно j-го входа при условии, что все остальные входы нулевые, т.е.

$$y_{ij}(t)=\int_0^t g_{ij}(t-\tau) v_j(\tau) d\tau$$

$(i=1, 2,\ldots,r;\ j=1,2,\ldots,m)$...

Правая часть этого скалярного равенства является *интегралом свертывания (сверткой функций).* Для двух интегрируемых функций $f_1(t)$ и $f_2(t)$ он определяется в общем виде соотношениями:

$$f_1(t) \ast f_2(t) = \int_0^t f_1(t-\tau) f_2(\tau)\, d\tau = \int_0^t f_1(\tau) f_2(t-\tau)\, d\tau$$

и обладает следующие свойства (верхние индексы k и —k указывают соответственно на кратность операций дифференцирования и интегрирования):

$$f_1(t) \ast f_2^{(k)}(t) = f_1^{(k)}(t) \ast f_2(t); \quad f_1(t) \ast f_2^{(-k)}(t) = f_1^{(-k)}(t) \ast f_2(t).$$

В соответствии с приведенными соотношениям выражение для $y_{ij}(t)$ можно записать четырьмя различными способами (в дальнейшем скалярные функции $y_{ij}(t)$ и $v_j(t)$ для упрощения обозначаются через $y(t)$ и $v(t)$):

$$y(t) = g(t) \ast v(t) = \int_0^t g(t-\tau) v(\tau)\, d\tau = \int_0^t g(\tau) v(t-\tau)\, d\tau =$$
$$= \int_0^t h(t-\tau) v'(\tau)\, d\tau = \int_0^t h(\tau) v'(t-\tau)\, d\tau,$$

где $h(t)$ — функция, производная которой по ее аргументу определяет $g(t)$, т.е.

$$g(t) = \frac{d}{dt} h(t); \quad g(t-\tau) = \frac{d}{d(t-\tau)} h(t-\tau).$$

Скалярные функции $g(f)$ и $h(t)$ называются соответственно *импульсной* и *переходной характеристиками.* Им можно дать наглядное истолкование, уводя в рассмотрение специальные функции.

4. Импульсная и переходная характеристики. Пусть на входе системы в момент времени τ приложен кратковременный импульс *v(t)* продолжительностью Δτ. Используя известную теорему о среднем значении, выражение для реакции на выходе можно представить в виде:

$$y(t)=\int_0^t g(t-\tau)\,v(\tau)\,d\tau = g(t-\theta)\int_0^t v(\tau)\,d\tau = g(t-\theta)\,s_u,$$

где $0<\theta<t$; s_u — площадь импульса.

Реакция равна значению импульсной характеристики $g(t-\tau)$, если $s_u=1$ и $\theta=\tau$, что соответствует на входе в момент времени τ импульса единичной площади и бесконечно малой продолжительности. В пределе при $\Delta\tau\to 0$ амплитуда такого импульса неограниченно возрастает, но его площадь остается конечной, численно равной единице. Мгновенный импульс, обладающий такими свойствами, представляет собой *единичную импульсную функцию* $\delta(t-\tau)$, действующую в момент времени τ (рис. 4.3, *а*), т.е.

$$\delta(t-\tau)=0\ \ (\text{при } t\neq\tau);\qquad \int_{\tau-\varepsilon}^{\tau+\varepsilon}\delta(t-\tau)\,d\tau=1\ \ (\text{для всех } \varepsilon>0).$$

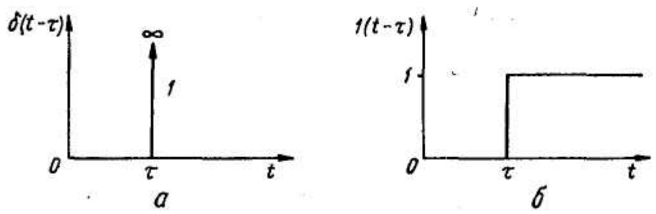


Рис. 4.3. Стандартные входные воздействия:
а — единичная импульсная функция; *б* — единичная ступенчатая функция.

Итак, $g(t-\tau)$ можно рассматривать как реакцию системы на единичную импульсную функцию $\delta(t-\tau)$, приложенную на входе в момент τ. Соответственно *импульсная характеристика g(t)* определяется как реакция на единичную импульсную функцию, воздействующую на систему в начальный момент времени $t=0$. То обстоятельство, что единичная импульсная функция не имеет конечного значения при $t=\tau$ не препятствует ее использованию, так как реакция зависит не от ее значения, а от конечной площади. Определение единичной импульсной функции не укладывается в обычные представления о функции, поэтому ее относят к классу обобщенных функций и называют также *дельта-функцией* или *функцией Дирака*.

Рассмотрим теперь интеграл свертывания в форме

$$y(t)=\int_0^t h(t-\tau)\,v'(\tau)\,d\tau$$

и выясним свойства входной функции $v(t)$, при воздействии которой реакция $y(t)$ на выходе равна переходной характеристике $h(t - \tau)$. Очевидно, производная $v'(t)$ должна представлять собой единичную импульсную функцию $\delta(t - \tau)$. Это значит, что $v(t)$ имеет скачок при $t = \tau$, а ее значения определяются интегралом

$$v(t)=\int_0^t \delta(t-\tau)\,d\tau.$$

В соответствии с определением $\delta(t - \tau)$ этот интеграл должен быть равным нулю при $t < \tau$ и единице при $t > \tau$. Такими свойствами обладает *единичная ступенчатая функция,* или *функция Хевисайда* (рис. 4.3,*б)*, определяемая как

$$1(t-\tau)=\begin{cases}0 \text{ при } t<\tau\\ 1 \text{ при } t>\tau.\end{cases}$$

Если $D \neq 0$, то

$$g(t)=Ce^{At}B+D\delta(t);\;\; h(t)=C(e^{At}-E)\,A^{-1}B+D.$$

Итак, $h(t - \tau)$ — это реакция на единичную ступенчатую функцию $1(t- \tau)$, приложенную на входе в момент времени τ. Соответственно *переходная характеристика $h(t)$* определяется как реакция на единичную ступенчатую функцию, воздействующую на систему в начальный момент времени $t = 0$.

Многомерная система характеризуется относительно ее входов и выходов матрицами $g(t)$ или $h(t)$, элементами которых являются соответственно импульсные $g_{ij}(t)$ или переходные $h_{ij}(t)$ характеристики для i-го выхода относительно j-го входа (при нулевых состояниях на всех других входах).

5. Принцип суперпозиции. Выражение для выходной функции (3) линейной системы можно рассматривать с позиций *принципа суперпозиции.* Так как $g(t - \tau)$ является реакцией на влияние единичной импульсной функции, то

$$y(t)=\int_0^t g(t-\tau)v(\tau)d\tau$$

представляет собой суммарную реакцию в момент t на элементарные воздействия, каждое из которых является импульсной функцией со значением $v(x)dx$, действующей в момент τ ($0\leq \tau\leq t$).

Другая форма интеграла свертывания

$$y\,(t)=\int_0^t g(\tau)v(t-\tau)d\tau$$

содержит входную функцию $v(t - \tau)$, сдвинутую на τ по отношению к моменту времени t. При $\tau = 0$ она равна $v(t)$ и с ростом τ функция $v(t - \tau)$ представляет собой воздействия для все более ранних моментов времени в прошлому, становясь при $\tau=t$ равной $v(0)$. Выходную реакцию $y(t)$ в момент t можно рассматривать как сумму элементарных реакций на импульсные функции со значениями $v\,(t - \tau)\,d\tau$, действующие в начальный момент времени $t = 0$.

Различие между двумя рассмотренными представлениями интеграла свертывания чисто формальное. В первом случае импульсные функции, на которые разлагается входное воздействие $v(t)$, как бы образуются стробирующим импульсом бесконечно малой длительности движущимся от 0 до t (рис. 4.4, *а*).

Во втором случае (рис. 4.4,*б)* стробирующий импульс неподвижен и расположен в точке t=0, а сама функция $v(t)$ перемещается в обратном направлении от t до 0. Ясно, что результат получим один и тот же.

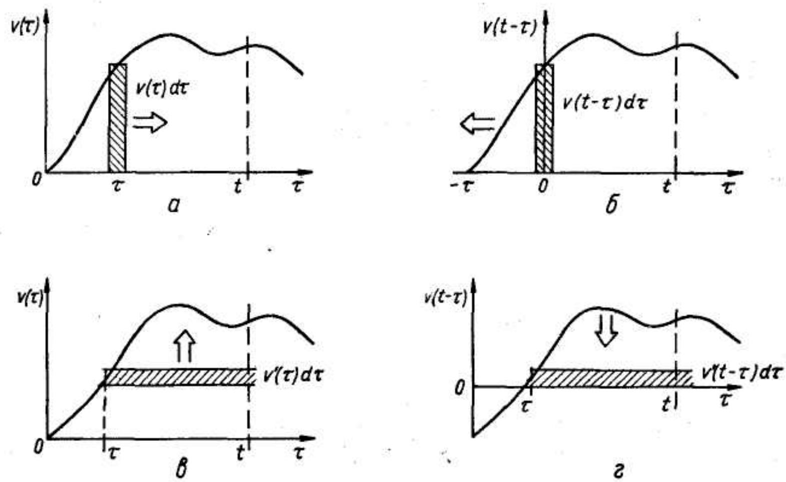


Рис. 4.4. Представление входного воздействия *v(t)* с помощью подвижного (*а*) и неподвижного (*б)* стробирующих импульсов, а также подвижной (*в)* и неподвижной *(г)* ступенчатых функций.

В двух остальных выражениях для выходной реакции через интегралы свертывания

$$y(t) = \int_0^t h(t-\tau)\, v'(\tau)\, d\tau = \int_0^t h(\tau)\, v'(t-\tau)\, d\tau$$

роль элементарных воздействий играют соответственно ступенчатые функции со значениями $v'(\tau)d\tau$ (рис. 4.4, *в)* и $v'(t-\tau)\, d\tau$ (рис. 4.4, *г).* Интерпретация интеграла свертывания на основе разложения по специальным функциям служит основанием для другого его названия — *интеграла суперпозиции.*

6. Связь с преобразованием Лапласа. Уравнение линейной стационарной системы

$$\frac{dx(t)}{dt} = Ax(t) + Bv(t); \qquad y(t) = Cx(t) + Dv(t)$$

можно представить на основе преобразования Лапласа в операторной форме

$$dx(t)/dt = Ax(t) + Bv(t); \quad y(t) = Cx(t) + Dv(t).$$

Отсюда получаем решения для вектора состояния и исходного вектора в функции комплексной частоты $p = \sigma - j\omega$:

$$X(p) = (pE - A)^{-1}[X(0) + BV(p)];$$
$$Y(p) = C(pE - A)^{-1}X(0) + [C(pE - A)^{-1}B + D]V(p).$$

Здесь матрица $\Phi(p) = (p - A)^{-1}$ является изображением переходной матрицы состояния $\Phi(t)$. Действительно, из выражения

$$X(p) = \Phi(p)X(0)+\Phi(p)BV(p)$$

на основе свойств преобразования Лапласа получаем

$$x(t) = \Phi(t)x(0) + \int_0^t \Phi(t-\tau)Bv(\tau)\,d\tau,$$

что совпадает с решением, приведенным в (3). Здесь, в частности, использовано то обстоятельство, что оригинал произведения двух функций $\Phi(p)$ и $BV(p)$ выражается через интеграл свертывания соответствующих им оригиналов $\Phi(t)$ и $Bv(t)$.

Таким образом, переходная матрица состояния $\Phi(t) = \exp(At)$ может быть вычислена путем обращения матрицы $F(p)=pE-A$ и последующего перехода от $(pE - A)^{-1}$ к ее оригиналу. Для обращения матрицы подходит, например, алгоритм Фаддеева или любой другой способ обращения матрицы. В результате имеем

$$\Phi(p) = (pE - A)^{-1} = \frac{G(p)}{\Delta(p)},$$

где $G(p)$ — присоединенная матрица для $F(p)$ и $\Delta(p)$ — характеристический многочлен n-й степени матрицы A, т.е. $\Delta(p) = \det(pE - A)$.

Если все элементы $G(p)$ имеют общие множители, то после сокращения на них $G(p)$ переходит в приведенную присоединенную матрицу $C(p)$, а $\Delta(p)$ — в минимальный многочлен $\psi(p)$.

Так как элементы матрицы $\Phi(p)$ являются дробно-рациональными функциями от p, причем степени их числителя всегда ниже степени знаменателя, то каждый из них можно разложить на простые дроби. В общем случае при наличии кратных собственных значений

$$\Delta(p) = (p-\lambda_1)^{m_1}(p-\lambda_1)^{m_2}...(p-\lambda_1)^{m_q}$$

и согласно теореме Сильвестра, имеем:

$$(pE - A)^{-1} = \frac{C(p)}{\psi(p)}\sum_{k=1}^{q}\left[\frac{Z_{k1}}{p-\lambda_k} + \frac{1!\,Z_{k2}}{(p-\lambda_k)^2} + \cdots + \frac{(m_k-1)!\,Z_{km_k}}{(p-\lambda_k)^{m_k}}\right],$$

где

$$Z_{kj} = \frac{1}{(j-1)!\,(m_k-j)!}\left[\frac{C(p)}{\psi_k(p)}\right]^{(m_k-j)}_{p=\lambda_k}; \quad \psi_k(p) = \frac{\psi(p)}{(p-\lambda)^{m_k}}.$$

Переходя от изображений к оригиналам, получаем

$$\Phi(t) = \sum_{k=1}^{q}\left[Z_{k1}e^{\lambda_k t} + Z_{k2}te^{\lambda_k t} + \dots + Z_{km_k}t^{m_k-1}e^{\lambda_k t}\right] = \\ = \sum_{k=1}^{q}\sum_{j=1}^{m_k} Z_{kj}t^{j-1}e^{\lambda_k t},$$

что совпадает с теоремой Сильвестра для экспоненциальной функции.

7. Передаточная матричная функция. При нулевых начальных условиях $x(0) = 0$ операторное выражение для выходного вектора принимает вид:

$$Y(p) = [C\Phi(p)B + D]V(p) = F(p)\,V(p).$$

Матрица $F(p)=C\Phi(p)B+D$ называется *передаточной матричной функцией.* Так как изображение i-й выходной переменной

$$Y_i(p) = \sum_{j=1}^{m} F_{ij}(p)V_j(p) \quad (i = 1,\ 2,\ \dots,\ r),$$

это элементы $F_{ij}(p)$ матрицы $F(p)$ можно рассматривать как *скалярные передаточные функции* от i -го входа к j-му выходу, причем

$$F_{ij}(p) = \frac{Y_i(p)}{V_j(p)} \quad \text{при } V_k(p) = 0 \qquad (k \neq j).$$

Зная $F_{ij}(p)$, легко получить импульсную $g_{ij}(t)$ и переходную $h_{ij}(t)$ характеристики. Для этого подставим в выражение $Y_i(p)=F_{ij}(p)V_j(p)$ вместо $V_j(p)$ соответственно изображения единичной импульсной функции $\delta(t)$, равное 1, и единичной ступенчатой функции $\delta(t)$, равное $1/p$.

Тогда получим изображение для $g_{ij}(t)$ и $h_{ij}(t)$:

$$g_{ij}(p) = F_{ij}(p); \quad h_{ij}(p) = \frac{F_{ij}(p)}{p}.$$

Отсюда видно, что передаточная функция $F_{ij}(p)$ является изображением импульсной функции, а изображение переходной характеристики равно $F_{ij}(p)$, деленной на p. Для получения $g_{ij}(t)$ и $h_{ij}(t)$ достаточно перейти от их изображений к оригиналам на основе обратного преобразования Лапласа. Определив эти характеристики для всевозможных пар «вход — выход», можно записать матрицы *g(t)* и *h(t).* Их также можно получить обратным преобразованием по Лапласу соответственно из матриц $F_{ij}(p)$ и $(1/p)F_{ij}(p)$.

Рассмотрим простой пример. Уравнения электрической схемы (рис. 4.5) для переменных состояния u_C и i_L, а также выходных переменных i_G и u_R можно получить в виде:

$$\frac{d}{dt}\begin{bmatrix} u_C \\ i_L \end{bmatrix} = \begin{bmatrix} -\frac{G}{C} & -\frac{1}{C} \\ \frac{1}{L} & -\frac{R}{L} \end{bmatrix}\begin{bmatrix} u_C \\ i_L \end{bmatrix} + \begin{bmatrix} \frac{1}{C} & 0 \\ 0 & -\frac{1}{L} \end{bmatrix}\begin{bmatrix} j(t) \\ e(t) \end{bmatrix};$$

$$\begin{bmatrix} i_G \\ u_R \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}\begin{bmatrix} u_C \\ i_L \end{bmatrix}.$$

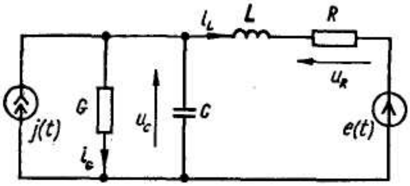


Рис. 4.5. Электрическая схема к примеру.

При C= 1Ф, L = (½)Г, G = 2(1/Ом), R = 2,5 Ом соответствующие матрицы получают следующие числовые выражения:

$$A = \begin{bmatrix} -2 & -1 \\ 2 & -5 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}; \quad C = \begin{bmatrix} 2 & 0 \\ 0 & 2{,}5 \end{bmatrix}; \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

В соответствии с изложенным выше находим:

$$\Phi(p) = \begin{bmatrix} p+2 & 1 \\ -2 & p+5 \end{bmatrix}^{-1} = \frac{1}{\Delta(p)}\begin{bmatrix} p+5 & -1 \\ 2 & p+2 \end{bmatrix};$$

$$\Delta(p) = p^2 + 7p + 12 = (p+3)(p+4);$$

$$F(p) = C\Phi(p)B + D = \begin{bmatrix} 2 & 0 \\ 0 & 2{,}5 \end{bmatrix}\frac{1}{\Delta(p)}\begin{bmatrix} p+5 & -1 \\ 2 & p+2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} =$$

$$= \frac{1}{(p+3)(p+4)}\begin{bmatrix} 2p+10 & 4 \\ 5 & -5p-10 \end{bmatrix}.$$

Разлагая элементы матрицы *F*(*p)* на простые дроби и переходя к оригиналам, получаем

$$F_{11}(p)=\frac{2p+10}{(p+3)(p+4)}=\frac{4}{p+3}-\frac{2}{p+4};\quad g_{11}(t)=4e^{-3t}-2e^{-4t};$$
$$F_{12}(p)=\frac{4}{(p+3)(p+4)}=\frac{4}{p+3}-\frac{4}{p+4};\quad g_{12}(t)=4e^{-3t}-4e^{-4t};$$
$$F_{21}(p)=\frac{5}{(p+3)(p+4)}=\frac{5}{p+3}-\frac{5}{p+4};\quad g_{21}(t)=5e^{-3t}-5e^{-4t};$$
$$F_{22}(p)=\frac{-5p-10}{(p+3)(p+4)}=\frac{5}{p+3}-\frac{10}{p+4};\quad g_{22}(t)=5e^{-3t}-10e^{-4t}.$$

Таким образом, имеем

$$g(t)=\begin{bmatrix}4e^{-3t}-2e^{-4t} & 4e^{-3t}-4e^{-4t}\\ 5e^{-3t}-5e^{-4t} & 5e^{-3t}-10e^{-4t}\end{bmatrix}.$$

Этот же результат можно получить через переходную матрицу состояния. Определив каким-либо способом $\Phi(t)=\exp(At)$, по формуле $g(t) = C\Phi(t)B+D$ находим

$$g(t)=\begin{bmatrix}2 & 0\\ 0 & 2{,}5\end{bmatrix}\begin{bmatrix}2e^{-3t}-e^{-4t} & -e^{-3t}+e^{-4t}\\ 2e^{-3t}-2e^{-4t} & -e^{-3t}+2e^{-4t}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -2\end{bmatrix}=$$
$$=\begin{bmatrix}4e^{-3t}-2e^{-4t} & 4e^{-3t}-4e^{-4t}\\ 5e^{-3t}-5e^{-4t} & 5e^{-3t}-10e^{-4t}\end{bmatrix}.$$

8. Управляемость и наблюдательность. Линейное преобразование переменных состояния $x = H\tilde{x}$ и $\tilde{x}= H^{-1}x$, где H — модальная матрица, приводит уравнение состояния $\dot{x}= Ax+Bv$ к виду

$$H\frac{d\tilde{x}}{dt}=AH\tilde{x}+Bv;\quad \frac{d\tilde{x}}{dt}=(H^{-1}AH)\tilde{x}+(H^{-1}B)v.$$

В новых координатах матрица системы A преобразуется к диагональной или жордановой форме $\tilde{A}= H^{-1}AH$, а матрица B к $\tilde{B}= H^{-1}B$. Уравнение состояния и выходное уравнение запишем следующим образом:

$$\frac{d\tilde{x}}{dt}=\tilde{A}\tilde{x}+\tilde{B}v;\quad y=\tilde{C}\tilde{x}+Dv,$$

где $\tilde{C}= CH$.

Если $\tilde{A}$ - диагональная матрица, соответствующая различным собственным значениям матрицы A, то уравнение состояния распадается на n скалярных уравнений

$$\frac{d\tilde{x}_i}{dt}=\lambda_i\tilde{x}_i+\tilde{b}_{(i)}v\quad(i=1,\ 2,\ \ldots,\ n),$$

где $b_{(i)}$ — i-я строка матрицы B.

Каждая с n преобразованных переменных $\tilde{x}_i$ связана только с одним собственным значениям.

Система называется *управляемой,* если все переменные $\tilde{x}_i$ зависят от входных воздействий v. Это значит, что переменные состояния $x = H\tilde{x}$, не содержат *свободных (неуправляемых) компонентов.* Очевидным условием управляемости является отсутствие нулевой строки в матрице B, т.е. все $\tilde{b}_{(i)}$ $(i = 1, 2, \ldots, n)$ должны быть ненулевыми векторами-строками. В общем случае кратных собственных значений доказывается необходимое и достаточное условие полной управляемости системы: матрица $[B, AB, A^2B, ..., A^{n-1}B]$ должна иметь ранг n. Полная управляемость означает, что с помощью некоторого воздействия $v(t)$, определенного на конечном интервале $0 \leq t \leq T$, система может быть переведена из заданного начального состояния $x(0)$ в конечное состояние $x(T)$.

Система называется *наблюдаемой,* если каждая из переменных $\tilde{x}_i$ $(i = 1, 2, \ldots, n)$ связана хотя бы с одним выходом (элементом выходного вектора y). Так как

$$y_t = [\tilde{c}^{(1)}, \tilde{c}^{(2)}, \ldots, \tilde{c}^{(n)}] \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \ldots \\ \tilde{x}_n \end{bmatrix} = \sum_{s=1}^{n} \tilde{c}^{(s)} \tilde{x}_s,$$

где $\tilde{c}^{(s)}$ — столбцы матрицы C, то очевидным условием наблюдательности является отсутствие в матрице C нулевого столбца. В общем случае кратных собственных значений доказывается необходимое и достаточное условие полной наблюдательности: матрица $[C^*, A^*C^*, A^{*2}C^*, \ldots, A^{*n-1}C^*]$ должна иметь ранг n (A^* и C^* — сопряженные матрицы). Полная наблюдательность означает, что существует такое воздействие $v(t)$, что по реакциям на выходах $y(t)$ на заданном интервале времени $0 \leq t \leq T$ можно определить начальное состояние $x(0)$ системы.

9. Устойчивость. Система называется *устойчивой по Ляпунову* при нулевом входе ($v = 0$), если для любого $\varepsilon > 0$ можно указать такое $\delta > 0$, что при любых начальных значениях $x_1(0)$, $x_2(0)$, ..., $x_n(0)$, меньших по модулю числа δ, переменные состояния $x_1(t)$, $x_2(t)$,…,$x_n(t)$ за все время движения ($t \geq 0$) по модулю остаются меньше числа ε, т.е. если

$| x_i(0) | < \delta$, то $| x_i(t) | < \varepsilon$ ($i = 1, 2, \ldots, n;\ t \geq 0$). Если кроме этого $x_i(t) \to 0$ при $t \to \infty$, то система называется *ассимптотически устойчивой.* Линейная система при ограниченных воздействиях устойчива, если ее реакция также ограничена.

Для линейных стационарных систем имеется непосредственная связь между ее устойчивостью и характером собственных значений матрицы системы A. Согласно теореме Сильвестра переходная матрица состояния

$$\Phi(t) = e^{At} = \sum_{k=1}^{q} (Z_{k1} + tZ_{k2} + \cdots + t^{m_k-1} Z_{km_k})\, e^{\lambda_k t},$$

где Z_{kj} ($k = 1, 2, \ldots, q;\ j = 1, 2, \ldots, m_k$) — компоненты матрицы A. Ясно, что при вещественных отрицательных собственных значениях ($\lambda_k<0$) система асимптотически устойчива, так как в этом случае $x(t) = \Phi(t)x(0)$ при $t \to \infty$ будет стремиться к нулевому вектору ($x(t) \to 0$).

Если среди собственных значений имеются комплексные, то они при вещественной матрице A могут появляться только комплексно-сопряженными парами. Пусть $\lambda'_k = \alpha + i\omega$ и $\lambda''_k = \alpha - i\omega$ - пара комплексно-сопряженных собственных значений кратности m_k. В выражении для $\Phi(t)$ им будут соответствовать слагаемые

$$t^{j-1} [Z_{kj} e^{(\alpha+i\omega)t} + \bar{Z}_{kj} e^{(\alpha-i\omega)t}] = t^{j-1} e^{\alpha t} [Z_{kj} e^{i\omega t} + \bar{Z}_{kj} e^{-i\omega t}],$$

где комплексно-сопряженные матрицы $Z_{kj} = Z'_{kj} + iZ''_{kj}$ и $\bar{Z}_{ki} = Z'_{kj} — iZ''_{kj}$. Элементарными преобразованиями это выражение приводится к виду

$$2t^{j-1} e^{\alpha t} (Z'_{kj} \cos \omega t - Z''_{kj} \sin \omega t) \qquad (j = 1, 2, \ldots, m_k).$$

Если вещественная часть комплексно-сопряженных значений $\alpha<0$, то система асимптотически устойчива. В случае чисто мнимых собственных значений $\alpha=0$ система устойчива только при отсутствии множителя t^{j-1}, что означает, что $m_k = 1$, т.е. собственные значения должны быть простыми. При $\alpha > 0$ — система всегда неустойчива.

10. Критерий Рауса-Гурвица. Для суждения об устойчивости системы совсем не обязательно вычислять собственные значения. Разработано много различных критериев устойчивоти, один из которых известен как алгебраический *критерий Рауса-Гурвица.*

Согласно критерию Рауса-Гурвица необходимое и достаточное условие устойчивоти сводится к требованию положительности определителей D_1, D_2, ..., D_n, элементами которых являются коэффициенты характеристического многочлена

$$\Delta(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n.$$

Определитель D_n образуется следующим образом: на главной диагонали располагаются коэффициенты $a_1, a_2, \ldots, a_n$, строки влево от главной диагонали заполняются коэффициентами с возрастающими индексами и вправо от главной диагонали — с убывающими индексами, а остальные позиции заполняются нулями. Например, для многочлена пятой степени

$$\Delta(\lambda) = a_0\lambda^5 + a_1\lambda^4 + a_2\lambda^3 + a_3\lambda^2 + a_4\lambda + a_5$$

имеем:

$$D_5 = \begin{vmatrix} a_1 & a_0 & 0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 & 0 \\ a_5 & a_4 & a_3 & a_2 & a_1 \\ 0 & 0 & a_5 & a_4 & a_3 \\ 0 & 0 & 0 & 0 & a_5 \end{vmatrix}.$$

Определители D_k ($k = 1, 2, \ldots, n - 1$) образуются из определителей D_{k+1} вычеркиванием последней строки и последнего столбца:

$$D_4 = \begin{vmatrix} a_1 & a_0 & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 \\ a_5 & a_4 & a_3 & a_2 \\ 0 & 0 & a_5 & a_4 \end{vmatrix}; \quad D_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a \end{vmatrix};$$

$$D_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix}; \quad D_1 = a_1.$$

11. Нестационарные системы. Решение дифференциального уравнения первого порядка с переменным коэффициентом $\dot{x}(t)=a(t)x+v(t)$, удовлетворяющее начальному условию $x(0) = x(t_0)$, имеет вид:

$$x = \left(x_0 + \int_{t_0}^{t} e^{-b(\tau)} v(\tau)\, d\tau\right) e^{b(t)} = e^{b(t)} x_0 + e^{b(t)} \int_{t_0}^{t} e^{-b(\tau)} v(\tau)\, d\tau,$$

где

$$b(t) = \int_{t_0}^{t} a(\tau)\, d\tau.$$

По аналогии с решением однородного скалярного уравнения (*v(t)*= 0) можно попробовать найти решение матричного уравнения *x=A(t)x* в виде

$$x(t) = (\exp\int_{t_0}^{t} A(\tau)\,d\tau)\,x(t_0) = e^{B(t)}x(t_0).$$

Подставляя это выражение у уравнение, приходим к соотношению

$$\frac{dB(t)}{dt} = e^{B(t)} = A(t)e^{B(t)},$$

являющемуся условием, при котором $x(t)=e^{B(t)}x(t_0)$ — решение уравнения $\dot{x}=A(t)x$. Можно показать, что это условие равнозначно требованию, чтобы матрицы $A(t_1)$ и $A(t_2)$ для всех t_1 и t_2 в интервале определения $t_0 \le (t_1, t_2) \le t_1$ были перестановочними, т.е. $A(t_1)A(t_2) = A(t_2)A(t_1)$. При выполнении данного условия переходная матрица состояния

$$\Phi(t, t_0) = \exp\int_{t_0}^{t} A(\tau)\,d\tau.$$

Общее решение уравнений нестационарной системы для вектора состояния *x (t)* и исходного вектора *y(t)* имеет вид:

$$x(t) = \Phi(t, t_0)\,x(t_0) + \int_{t_0}^{t} \Phi(t, \tau)\,B(\tau)\,v(\tau)\,d\tau;$$

$$y(t) = C(t)\,\Phi(t, t_0)\,x(t_0) + \int_{t_0}^{t} C(t)\,\Phi(t, \tau)\,B(\tau)\,v(\tau)\,d\tau + D(\tau)\,v(\tau).$$

В общем случае переходная матрица состояния выражается рядом

$$G(A) = E + \int_{t_0}^{t} A(\tau)\,d\tau + \int_{t_0}^{t} A(\tau)\int_{t_0}^{t} A(\tau)\,d\tau d\tau + \int_{t_0}^{t} A(\tau)\int_{t_0}^{t} A(\tau)\int_{t_0}^{t} A(\tau)\,d\tau d\tau d\tau + \cdots,$$

называемым *матрициантом*. Доказывается, что для ограниченной матрицы *A* на интервале интегрирования этот ряд сходится абсолютно и равномерно. Дифференцируя обе части, получаем основное свойство матрицианта:

$$\frac{dG(A)}{dt} = A(t)\,G(A).$$

Ясно, что $G(A)$ удовлетворяет уравнению состояния и поэтому представляет собой переходную матрицу состояния нестационарной системы, т.е. $\Phi(t, t_0) = G(A)$. Практическое использование матрицианта зарудняется в связи с тем, что ряд может сходиться медленно и для получения решения потребуется большой объем вычислений. В таких случаях используют специальные методы интегрировании дифференциальных уравнений с переменными параметрами.

Для стационарной системы (при постоянной матрице A) матрициант переходит в ряд, который выражает экспоненциальную функцию:

$$E + (t - t_0)A + \frac{(t-t_0)^2}{2!}A^2 + \frac{(t-t_0)^3}{3!}A^3 + \cdots = e^{A(t-t_0)}.$$

В то время как переходная матрица состояния стационарной системы является функцией только одной переменной (разности $t - t_0$ или t при $t_0 = 0$), для нестационарной системы решение зависит от двух переменных, одной из которых является момент t_0 воздействия и другой — момент t наблюдения реакции. Соответственно интегралы свертывания обобщаются к *интегралам совмещения.*

Микромодуль 12
Идивидуальные тестовые задачи

1. Система описывается уравнениями:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= -2x_2 \\ \frac{dx_2}{dt} &= x_1 - 3x_2 + v \end{aligned}\right\}.$$

а) Запишите уравнение переменных состояния $\dot{x} = Ax + B$ и матрицы A и B.

б) Определите переходную матрицу состояния $\Phi(t) = e^{At}$ как экспоненциальную функцию от матрицы и операторным методом.

в) Найдите импульсную и передаточную характеристики для выходной переменной $y = x_1$.

2. Уравнения переменных состояния системы имеют вид:

$$\frac{dx}{dt} = \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix} x + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} v;$$

$$y = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} x.$$

а) Найдите передаточную матричную функцию $F(p) = C\Phi(p)B + D$.
б) Определите с помощью обратного преобразования Лапласа импульсную и переходную матричные характеристики системы.
3. Система с двумя входами (v_1, v_2) и двумя выходами (y_1, y_2) описывается уравнениями:

$$\left.\begin{aligned} &\frac{d^2y_1}{dt^2} + 3\frac{dy_1}{dt} + 2y_2 = v_1 \\ &\frac{d^2y_2}{dt^2} + \frac{dy_1}{dt} + y_2 = v_2 \end{aligned}\right\}$$

а) Положив $x_1 = y_1$, $x_2 = \dot{y}_1$, $x_3 = y_2$, $x_4 = \dot{y}_2$, покажите, что матрицы A, B, C и D в уравнениях

$$\dot{x} = Ax + Bv, \; y = Cx + D$$

имеют вид:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -3 & -2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

б) Определите изображение переходной матрицы состояния с помощью алгоритма Фаддеева.
в) Найдите передаточную матричную функцию системы.
4. Дана система, которая описывается в пространстве переменных состояния уравнениями:

$$\begin{aligned} &\frac{dx_1}{dt} = -\frac{7}{3}x_1 + \frac{1}{3}x_2 + \frac{4}{3}x_3 + 2v; \\ &\frac{dx_2}{dt} = \frac{2}{3}x_1 - \frac{8}{3}x_2; \\ &\frac{dx_3}{dt} = x_1 - x_2 - 2x_3 - v; \\ &y = \frac{2}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3. \end{aligned}$$

а) Запишите уравнение системы в матричной форме.
б) Найдите переходную матрицу состояния $\Phi(t)=e^{At}$ как экспоненциальную функцию от матрицы и операторным методом.
в) Выразите вектор переменных состояния $x(t)$ при $v = 0$ и начальных значениях x_{10}, x_{20}, x_{30}.
г) Преобразуйте матрицу системы A к диагональной форме заменой переменных состояния с помощью модальной матрицы и запишите решение для новых переменных в «развязанной» форме.
д) Найдите передаточную матричную функцию, а также импульсную и переходную характеристики системы.

е) Исследуйте систему на управляемость, наблюдательность и устойчивость.

5. Найдите операторным методом импульсную матричную характеристику системы, представленной уравнениями:

$$\frac{dx}{dt}=\begin{bmatrix}-1 & 0 & 0\\ 0 & -3 & 0\\ 0 & 0 & -5\end{bmatrix}x+\begin{bmatrix}1 & 0\\ 0 & 1\\ 1 & 1\end{bmatrix}v;$$

$$y=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}x.$$

6. Исследуйте управляемость, наблюдательность и устойчивость следующих систем:

а) $$\frac{dx}{dt}=\begin{bmatrix}-3 & 5\\ 2 & 0\end{bmatrix}x+\begin{bmatrix}1\\ 2\end{bmatrix}v;$$
$$y=[1\ -1]\,x.$$

б) $$\frac{dx}{dt}=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}x+\begin{bmatrix}1\\ 1\end{bmatrix}v;$$
$$y=[1\ -1]\,x+v.$$

в) $$\frac{dx}{dt}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}x+\begin{bmatrix}1\\ 1\end{bmatrix}v;$$
$$y=[1\ \ 1]\,x-v.$$

г) $$\frac{dx}{dt}=\begin{bmatrix}0 & 1 & 0\\ 5 & 0 & 2\\ -2 & 0 & -2\end{bmatrix}x+\begin{bmatrix}-1\\ 1\\ -1\end{bmatrix}v;$$
$$y=[-2\ \ 1\ \ 0]\,x.$$

7. В электрической схеме (рис. 4.6) действует два источника напряжения u_1 и u_2, выходы характеризуются напряжениями u_{k1} и u_{k2}, а переменными состояния являются напряжение на емкости u_C и токи в индуктивностях i_{L1} и i_{L2}.

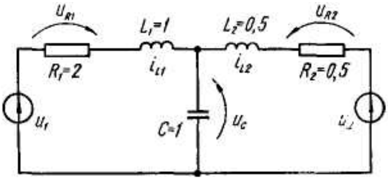


Рис. 4.6. Электрическая схема к задаче 7.

Уравнение переменные состояния $\dot{x}=Ax+Bv$ для рассматриваемой системы при заданных значениях параметров ее компонентов можно записать следующим образом:

$$\frac{d}{dt}\begin{bmatrix} u_C \\ i_{L1} \\ i_{L2} \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & -2 & 0 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} u_C \\ i_{L1} \\ i_{L2} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

а) Покажите, что выходное уравнение $y = Cx$ имеет вид:

$$\begin{bmatrix} u_{R1} \\ u_{R2} \end{bmatrix} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} u_C \\ i_{L1} \\ i_{L2} \end{bmatrix}$$

б) Определите переходную матрицу состояния $\Phi(t)=e^{At}$ как функцию от матрицы и операторным методом.

в) Найдите передаточную матричную функцию, а также импульсную и переходную характеристики схемы.

г) Представьте полученные результаты в тригонометрической форме.

Модуль 5
Общая теория линий и поверхностей второго рорядка

Микромодуль 13
Общая теория линий второго порядка

5.1. Преобразование координат на плоскости

1. Вспомним формулы преобразования декартовых прямоугольных координат на плоскости.

1) Если новые оси получаются параллельным переносом старых осей на величину a в направлении оси Ox и на величину b в направлении оси Oy, то

$$x = x' + a, \quad y = y' + b; \tag{1}$$

здесь $(x; y)$— старые координаты произвольной точки, $(x'; y')$ — новые координаты той же точки.

2) Если новые оси получаются поворотом старых осей на угол α вокруг неподвижного начала координат, то

$$\left.\begin{aligned} x &= x' \cos\alpha - y' \sin\alpha, \\ y &= x' \sin\alpha + y' \cos\alpha. \end{aligned}\right\} \tag{2}$$

Эти формулы будут применяться в ближайших параграфах для упрощения общего уравнения кривой второго порядка. Следует иметь в виду, что формулы (1) и (2) справедливы при условии, что для новой системы координат сохраняется прежний масштаб.

2. Чтобы облегчить дальнейшее употребление формул (2), запишем их более компактно:

$$\left.\begin{aligned} x &= l_1x' + l_2y', \\ y &= m_1x' + m_2y'. \end{aligned}\right\} \qquad (3)$$

Иначе говоря, мы ввели обозначения коэффициентов формул (2):

$$l_1 = \cos\alpha, \quad m_1 = \sin\alpha, \quad l_2 = -\sin\alpha, \quad m_2 = \cos\alpha.$$

Заметим, что пара коэффициентов l_1, m_1 допускает простое геометрическое истолкование. Именно, если мы отложим на новой оси Ox' отрезок длины =1, направленный из начала координат в положительную сторону этой оси, т. е. единичный вектор направления Ox', то проекции этого вектора на старые координатные оси будут:

$$l_1 = \cos\alpha, \; m_1 = \sin\alpha.$$

Таким образом

$$\{l_1;\; m_1\} = \{\cos\alpha;\; \sin\alpha\} \qquad (4)$$

есть единичный вектор, определяющий направление новой оси абсцисс. Аналогично

$$\{l_2;\; m_2\} = \{-\sin\alpha;\; \cos\alpha\} = \\ = \left\{\cos\left(\alpha + \frac{\pi}{2}\right); \quad \sin\left(\alpha + \frac{\pi}{2}\right)\right\} \qquad (5)$$

есть единичный вектор, определяющий направление новой оси ординат (рис. 1).

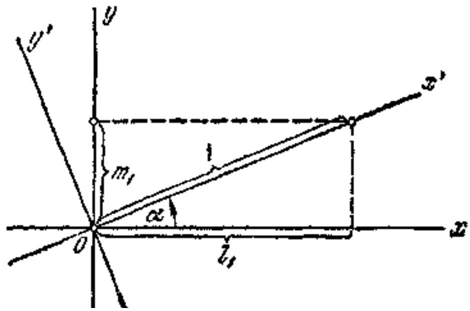


Рис. 1.

3. Формулы (2) выражают старые координаты произвольной точки через ее новые координаты. Часто требуются обратные формулы, выражающие новые координаты через старые. Чтобы найти их, заметим, что новая система осей получается поворотом старой на угол α; в свою очередь старые оси можно получить поворотом новых на угол — α. Поэтому мы можем в формулах (2) поменять ролями старые и новые координаты, одновременно заменяя α на — α, отсюда

$$x' = x\cos\alpha + y\sin\alpha,$$
$$y' = -x\sin\alpha + y\cos\alpha,$$

так как $\cos(-\alpha) = \cos\alpha$, $\sin(-\alpha) = -\sin\alpha$. Если применить введенные выше обозначения, то получим

$$\left.\begin{aligned} x' &= l_1x + m_1y, \\ y' &= l_2x + m_2y. \end{aligned}\right\} \qquad (6)$$

4. Коэффициенты l_1, m_1, l_2, m_2 формул (3) удовлетворяют следующим условиям:

$$l_1^2 + m_1^2 = 1, \quad l_2^2 + m_2^2 = 1, \qquad (7)$$

$$l_1l_2 + m_1m_2 = 0, \qquad (8)$$

$$\begin{vmatrix} l_1 & m_1 \\ l_2 & m_2 \end{vmatrix} = 1. \qquad (9)$$

Эти равенства непосредственно усматриваются из (4) и (5). Соотношения (7), (8), (9) являются не только необходимыми, но и достаточными условиями того, что формулы (3) выражают преобразования прямоугольных координат при некотором повороте системы осей (с неизменным масштабом). В самом деле, допустим, что нам дана система прямоугольных координат с осями *Ox, Oy* и даны формулы (3) с какими-то коэффициентами l_1, m_1, l_2, m_2. Если соблюдены условия (7), то найдутся угол α и угол β такие, что

$$\{l_1;\ m_1\} = \{\cos\alpha;\ \sin\alpha\}, \quad \{l_2;\ m_2\} = \{\cos\beta;\ \sin\beta\}.$$

Если соблюдено также равенство (8), то

$$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\beta - \alpha) = 0.$$

Таким образом, равенство (8) гарантирует перпендикулярность векторов $\{l_1;\ m_1\}$ и $\{l_2;\ m_2\}$. Так как $\cos(\beta - \alpha) = 0$, то либо $\beta = \alpha + \frac{\pi}{2}$, либо $\beta = \alpha + \frac{3\pi}{2}$. Соответственно, имеем:

$$\text{либо } \{l_2;\ m_2\} = \{-\sin\alpha;\ \cos\alpha\},$$
$$\text{либо } \{l_2;\ m_2\} = \{\sin\alpha;\ -\cos\alpha\}.$$

Если соблюдено (9), то второе предположение отпадает и мы получаем $\{l_2;\ m_2\}=\{-\sin\alpha;\ \ \cos\alpha\}$. Вместе с тем формулы (3) сводятся к формулам (2). Значит, при соблюдении условий (7), (8), (9) мы действительно имеем переход к новой, также прямоугольной системе координат с тем же масштабом; при этом обе оси поворачиваются на один и тот же угол в одну и ту же сторону.

5. Рассмотрим теперь преобразование прямоугольных координат, при котором ось Ox' получается поворотом оси Ox на угол α, а ось Oy' получается поворотом оси Oy на угол α+π (рис. 2).

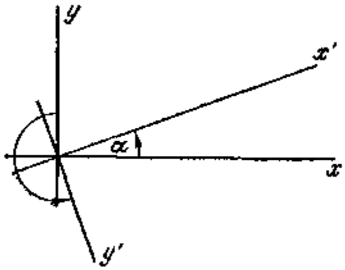


Рис. 2.

Такое преобразование нарушает ориентацию осей: если кратчайший поворот оси Ox к оси Oy совершается против часовой стрелки, то кратчайший поворот оси Ox' к оси Oy' идет по ходу часов. В этом случае старые и новые координаты связаны формулами:

$$\left.\begin{aligned} x &= x'\cos\alpha + y'\sin\alpha,\\ y &= x'\sin\alpha - y'\cos\alpha. \end{aligned}\right\} \qquad (10)$$

Чтобы убедиться в справедливости этих формул, будем рассуждать так: данное преобразование можно осуществить в два этапа, поворачивая сначала обе оси на угол α, затем меняя направление повернутой оси ординат на противоположное. Значит, нужно написать формулы (2), затем изменить знаки перед членами, которые содержат y'. Таким образом мы и получим (10). Формулы (10) можно написать в виде формул (3), если считать

$$\{l_1;\ m_1\}=\{\cos\alpha;\quad \sin\alpha\},$$
$$\{l_2;\ m_2\}=\{\sin\alpha;\ -\cos\alpha\}.$$

Легко проверить, что и в этом случае соблюдаются условия (7), (8). Однако вместо (9) будем иметь

$$\begin{vmatrix} l_1 & m_1 \\ l_2 & m_2 \end{vmatrix} = -1. \qquad (11)$$

Соотношения (7), (8), (11) являются не только необходимыми, но и достаточными условиями того, что формулы (3) выражают преобразование прямоугольных координат, нарушающее ориентацию осей (для доказательства нужно повторить рассуждения, проведенные в п° 4, до того места, где устанавливается, что либо $\beta = \alpha + \frac{\pi}{2}$, либо $\beta = \alpha + \frac{3\pi}{2}$; при условии (11) отпадает первая возможность).

Мы можем теперь сформулировать следующее утверждение: *если коэффициенты формул* (3) *удовлетворяют условиям* (7), (8), *то эти формулы выражают преобразование прямоугольных координат с неизменным масштабом* (и, конечно, с неизменным началом координат). Ориентация осей сохраняется или нарушается, в зависимости от того, будет ли соблюдено условие (9) или условие (11).

5.2. Приведение к каноническому виду уравнения линии второго порядка с центром в начале координат

6. Пусть дано уравнение

$$Ax^2 + 2Bxy + Cy^2 = H. \qquad (1)$$

Особенностью его является отсутствие членов первой степени относительно x, y. Этому соответствует известная особенность расположения линии второго порядка, которая задана уравнением (1). Именно, левая часть (1) не меняется при замене x, y на $-x$, $-y$; следовательно, если точка $M(x; y)$ лежит на линии (1), то точка $N(-x; -y)$ также лежит на этой линии. Иначе говоря, точки линии (1) расположены парами, симметрично относительно начала координат. Таким образом, если линия второго порядка задана уравнением вида (1), то а) она обладает центром (симметрии); б) начало координат помещено в центр.

7. Левая часть уравнения (1)

$$Ax^2 + 2Bxy + Cy^2 \qquad (2)$$

представляет собой однородный многочлен второй степени (т. е. многочлен, состоящий только из членов второй степени). Такой многочлен называют *квадратичной формой* от двух переменных x, y.

Мы будем сейчас заниматься задачей о приведении квадратичной формы (2) к каноническому виду. Сущность этой задачи в следующем:

требуется повернуть координатные оси так, чтобы после приведения формы (2) к новым координатам исчез член с произведением новых текущих координат. Согласно п. 5.1 дело сводится к тому, чтобы найти формулы

$$\left.\begin{aligned} x &= l_1 x' + l_2 y', \\ y &= m_1 x' + m_2 y', \end{aligned}\right\} \tag{3}$$

в силу которых имеет место тождество:

$$Ax^2 + 2Bxy + Cy^2 = \lambda_1 x'^2 + \lambda_2 y'^2. \tag{4}$$

Это значит, что после замены величин *x, у* но формулам (3) данная квадратичная форма должна принять вид, указанный в правой части (4) и называемый *каноническим видом* (в каноническом виде формы ее средний коэффициент равен нулю). Коэффициенты формул (3) надлежит подобрать так, чтобы соблюдались условия (7) — (9) п.5.1.

Если квадратичная форма (2) будет приведена к каноническому виду, то одновременно с нею получит канонический вид и уравнение (1). Вопрос о приведении к каноническому виду общего уравнения второй степени будет решен вообще в п.5.4. Здесь мы решим его для уравнения вида (1). Мы докажем, что каждую квадратичную форму (2), а вместе с тем каждое уравнение (1), можно привести к каноническому виду, и дадим простой способ искомого приведения.

8. Предположим сначала, что коэффициенты формул (3) уже найдены и тождество (4) достигнуто. Перепишем его следующим образом:

$$(Ax + By)\,x + (Bx + Cy)\,y = \lambda_1 x'x' + \lambda_2 y'y'. \tag{5}$$

Теперь каждую из скобок в левой части преобразуем по формулам (3):

$$\left.\begin{aligned} Ax + By &= (Al_1 + Bm_1)\,x' + (Al_2 + Bm_2)\,y', \\ Bx + Cy &= (Bl_1 + Cm_1)\,x' + (Bl_2 + Cm_2)\,y'. \end{aligned}\right\} \tag{6}$$

Используем, далее, обратные формулы преобразования координат:

$$\begin{aligned} x' &= l_1 x + m_1 y, \\ y' &= l_2 x + m_2 y; \end{aligned}$$

отсюда

$$\left.\begin{aligned} x'x' &= (l_1 x + m_1 y)\,x', \\ y'y' &= (l_2 x + m_2 y)\,y'. \end{aligned}\right\} \tag{7}$$

Тождество (5) вследствие (6) и (7) принимает вид:

$$\begin{aligned} (Al_1 + Bm_1)\,xx' + (Bl_1 + Cm_1)\,yx' + (Al_2 + Bm_2)\,xy' + \\ + (Bl_2 + Cm_2)\,yy' = \\ = \lambda_1 l_1 xx' + \lambda_1 m_1 yx' + \lambda_2 l_2 xy' + \lambda_2 m_2 yy'. \end{aligned} \tag{8}$$

В левой части (8) содержатся четыре различных члена; столько же аналогичных членов мы имеем в правой части (8). Тождество (8) будет соблюдено, если коэффициенты подобных членов слева и справа окажутся одинаковыми; одновременно будет соблюдено и тождество (4). Таким образом, для решения задачи достаточно подобрать коэффициенты формул (3) и числа λ_1, λ_2 так, чтобы

$$\left.\begin{aligned} Al_1+Bm_1&=\lambda_1 l_1,\\ Bl_1+Cm_1&=\lambda_1 m_1\end{aligned}\right\};\quad \left.\begin{aligned} Al_2+Bm_2&=\lambda_2 l_2,\\ Bl_2+Cm_2&=\lambda_2 m_2.\end{aligned}\right\} \qquad (9)$$

Дело свелось к системе уравнений:

$$\left.\begin{aligned} Al+Bm&=\lambda l,\\ Bl+Cm&=\lambda m.\end{aligned}\right\} \qquad (10)$$

Задача будет завершена, если мы найдем два решения l_1, m_1, λ_1 и l_2, m_2, λ_2 системы (10), удовлетворяющих условиям (7)-(9) п.5.1.

9. Займемся системой (10). Перепишем ее следующим образом:

$$\left.\begin{aligned} (A-\lambda)\,l+Bm&=0,\\ Bl+(C-\lambda)\,m&=0.\end{aligned}\right\} \qquad (11)$$

Отсюда

$$\begin{vmatrix} A-\lambda & B\\ B & C-\lambda\end{vmatrix}=0, \qquad (12)$$

так как в противном случае мы не получим из системы (11) ничего, кроме l=0, m=0. Развертывая определитель, напишем (12) в виде

$$\lambda^2-(A+C)\,\lambda+(AC-B^2)=0.$$

Из этого квадратного уравнения мы найдем нужные нам значения $\lambda=\lambda_1$, $\lambda=\lambda_2$:

$$\lambda_{1,2}=\frac{A+C\pm\sqrt{(A+C)^2-4\,(AC-B^2)}}{2}.$$

Заметим, что

$$(A+C)^2-4\,(AC-B^2)=(A-C)^2+4B^2\geqslant 0;$$

следовательно, уравнение (12) имеет только действительные корни.

Уравнение (12) называется *характеристическим уравнением* квадратичной формы (2); корни λ_1, λ_2 уравнения (12) называются *характеристическими числами* этой формы; они же получаются в качестве коэффициентов после приведения формы к каноническому виду.

Рассмотрим два возможных случая.

Первый случай. $(A-C)^2+4B^2>0$. При этом условии $\lambda_1\neq\lambda_2$

Если мы подставим $\lambda=\lambda_1$ в систему (11), то система (11) будет иметь ненулевое решение l, m. Построим вектор $\{l;\ m\}$. Направление этого вектора называется *главным направлением* данной формы, соответствующим характеристическому числу λ_1. Ясно, что вектор $\{l_1,\ m_1\}$, где $l_1=\mu l,\ m_1=\mu m$ (μ — какое-нибудь число $\neq 0$), также имеет главное направление.
Возьмем

$$\mu=\frac{\pm 1}{\sqrt{l^2+m^2}};$$

тогда

$$l_1^2+m_1^2=1.$$

Такое решение l_1, m_1 системы (11) будем называть *нормированным.* Вектор $\{l_1;\ m_1\}$ является единичным вектором главного направления, которое соответствует числу λ_1. Аналогично, решая систему (11) при $\lambda=\lambda_2$ и нормируя решение, получим единичный вектор $\{l_2;\ m_2\}$. Этот вектор определяет другое главное направление данной квадратичной формы, соответствующее характеристическому числу λ_2. Докажем, что главные направления, соответствующие различным характеристическим числам $\lambda_1,\ \lambda_2\ (\lambda_1\neq\lambda_2)$, перпендикулярны друг к другу.

Для доказательства прежде всего заметим, что числа $l_1,\ m_1,\ \lambda_1$ и $l_2,\ m_2,\ \lambda_2$ удовлетворяют равенствам (9) (поскольку эти числа найдены из (11)).

Умножим первое и второе равенства левой группы (9) соответственно на l_2, m_2 и сложим их почленно; точно так же поступим с правой группой, используя l_1, m_1. Получим:

$$(Al_1+Bm_1)\,l_2+(Bl_1+Cm_1)\,m_2=\lambda_1\,(l_1l_2+m_1m_2),$$
$$(Al_2+Bm_2)\,l_1+(Bl_2+Cm_2)\,m_1=\lambda_2\,(l_1l_2+m_1m_2).$$

Раскрывая скобки, легко обнаружить, что слева в этих двух равенствах стоит одна и та же величина. Поэтому, вычитая почленно одно равенство из другого, найдем:

$$0=(\lambda_1-\lambda_2)\,(l_1l_2+m_1m_2).$$

Но $\lambda_1\neq\lambda_2$, следовательно,

$$l_1l_2+m_1m_2=0,$$

что и требовалось (см. п° 4).

Поставленная задача в рассматриваемом случае решена. В самом деле, так как по доказанному векторы, определяющие главные направления, перпендикулярны друг к другу, то мы можем направить по ним оси новой прямоугольной системы координат. Переходя к новым координатам по формулам (3) с найденными коэффициентами l_1, m_1, l_2, m_2, мы приведем данную квадратичную форму к каноническому виду; это ясно, поскольку равенства (9) соблюдены. Следует еще учесть условие (9) п.5.1. Но и его соблюсти нетрудно, умножая, в случае надобности, l_1, m_1 (или l_2, m_2) на —1.

Второй случай. $(A-C)^2+4B^2=0.$ На этот раз $\lambda_1=\lambda_2$. Так как теперь нет надобности различать λ_1 и λ_2, будем писать просто λ. Очевидно также, что в данном случае $A=C$, $B=0$. Отсюда и из формулы решения квадратного уравнения (12) (приведена выше) находим

$$\lambda=A=C.$$

Если подставить это значение λ в систему (11), то все коэффициенты системы обратятся в нуль. Следовательно, система (11) будет состоять из тождеств и удовлетворяется любыми числами l, m. Таким образом, *если характеристические числа квадратичной формы совпадают, то любое направление является для такой формы главным.* Так как $A = C$, $B=0$, то данная форма уже имеет канонический вид

$$Ax^2+Ay^2$$

и ничего с ней делать не надо. Однако мы можем повернуть оси на любой угол, и форма при этом сохранит канонический вид.

10. Подводя итог изложенному, сформулируем следующее правило. *Для того чтобы привести данную квадратичную форму к каноническому виду, нужно решить уравнение* (12) *и найти характеристические числа* λ_1 и λ_2*; они и будут коэффициентами в каноническом виде формы. Координатные оси, следует направить по главным направлениям формы. Если ось абсцисс направлена по первому главному направлению* (которое определяется системой (11) при $\lambda=\lambda_1$)*, то* λ_1 *будет коэффициентом при квадрате абсциссы.*

Ясно, что тем самым дано также правило приведения к каноническому виду уравнения (1), поскольку его левая часть есть квадратичная форма.

Пример. Привести к каноническому виду уравнение:

$$17x^2+12xy+8y^2=20.$$

Решение. Составляем характеристическое уравнение:

$$\begin{vmatrix} 17-\lambda & 6 \\ 6 & 8-\lambda \end{vmatrix}=0.$$

Напишем его в развернутом виде:

$$\lambda^2 - 25\lambda + 100 = 0.$$

Отсюда находим характеристические числа: $\lambda_1 = 20$, $\lambda_2 = 5$. Тем самым мы уже знаем каноническое уравнение:

$$20x'^2 + 5y'^2 = 20,$$

или

$$\frac{x'^2}{1} + \frac{y'^2}{4} = 1.$$

Данная линия является эллипсом с полуосями $a = 2$, b=l. Если нас, кроме того, интересует расположение линии, то мы должны найти преобразование координат, приводящее данное уравнение к каноническому виду. Напишем систему (11):

$$\left.\begin{aligned} (17-\lambda)\,l + 6m &= 0, \\ 6l + (8-\lambda)\,m &= 0. \end{aligned}\right\} \qquad (*)$$

Полагая здесь $\lambda = \lambda_1 = 20$, получим

$$-3l + 6m = 0,$$
$$6l - 12m = 0.$$

Фактически мы имеем здесь одно уравнение; в качестве его решения можно взять, например, $l = 2$, $m = 1$. Нормируя это решение, найдем:

$$l_1 = \frac{2}{\sqrt{5}}, \quad m_1 = \frac{1}{\sqrt{5}}.$$

Полагая в системе (*) $\lambda = \lambda_2 = 5$, получим

$$12l + 6m = 0,$$
$$6l + 3m = 0.$$

Отсюда

$$l_2 = \frac{-1}{\sqrt{5}}, \quad m_2 = \frac{2}{\sqrt{5}}$$

(знаки взяты так, чтобы соблюдалось условие (9) п.5. 1). Переход к новым осям достигается поворотом на угол α, который определяется равенствами $\cos\alpha = l_1 = \frac{2}{\sqrt{5}}$, $\sin\alpha = m_1 = \frac{1}{\sqrt{5}}$.

Прощепостроить этот угол по тангенсу:

$$\operatorname{tg}\alpha = \frac{m_1}{l_1} = \frac{1}{2}$$

(рис. 3).

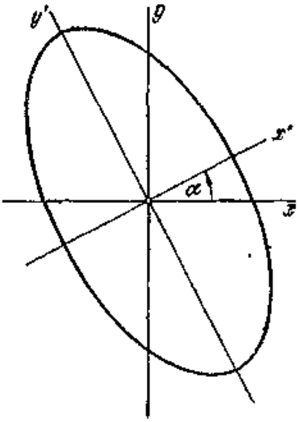


Рис. 3.

Формулы соответствующего преобразования координат напишем согласно (3):

$$x = \frac{2x' - y'}{\sqrt{5}}, \quad y = \frac{x' + 2y'}{\sqrt{5}}.$$

В том, что эти формулы действительно дают тождество

$$17x^2 + 12xy + 8y^2 = 20x'^2 + 5y'^2,$$

можно убедиться и непосредственной проверкой.

5.3. Инварианты и классификация квадратичных форм от двух аргументов

11. В предыдущем параграфе мы указали определенный способ приведения данной квадратичной формы от двух аргументов к каноническому виду. Естественно возникает вопрос: не может ли случиться, что с помощью какого-то другого преобразования прямоугольных координат данная квадратичная форма приведется к другому каноническому виду? Покажем, что этого не может быть. Иначе говоря, *каждая квадратичная форма имеет единственный канонический вид* (конечно, не считая возможности поменять ролями абсциссу и ординату). Следует оговориться, что **единственность канонического вида формы имеет место при условии, что употребляются только <u>прямоугольные координатные системы с одним и тем же масштабом.</u>**

Перейдем к доказательству нашего утверждения. Обозначим буквой Ф величину данной квадратичной формы:

$$\Phi = Ax^2 + 2Bxy + Cy^2. \quad (1)$$

В силу этого равенства каждой точке *M(x; y)* сопоставляется определенное численное значение Ф; его называют значением данной формы в точке *M (x; y).* Приведем данную форму к каноническому виду. Пусть λ_1, λ_2 — характеристические числа формы; предположим для определенности, что $\lambda_1 \leqslant \lambda_2$, и направим новую ось абсцисс по тому главному направлению, которое отвечает числу λ_1. Тогда значение Ф в точке *M* будет выражено равенством

$$\Phi = \lambda_1 x'^2 + \lambda_2 y'^2,$$

где x', y' — новые координаты точки *M.* Так как $\lambda_1 \leqslant \lambda_2$, то

$$\lambda_1 (x'^2 + y'^2) \leqslant \lambda_1 x'^2 + \lambda_2 y'^2 \leqslant \lambda_2 (x'^2 + y'^2).$$

Поскольку $x'^2 + y'^2 = x^2 + y^2$, то такие же соотношения имеем и в старых координатах:

$$\lambda_1 (x^2 + y^2) \leqslant \Phi \leqslant \lambda_2 (x^2 + y^2).$$

Ограничим возможные положения точки *M,* требуя, чтобы она находилась на единичной окружности:

$$x^2 + y^2 = 1.$$

При этом условии все возможные значения Ф (отвечающие любым положениям точки *M* на единичной окружности) заключены между λ_1 и λ_2, т. е.

$$\lambda_1 \leqslant \Phi \leqslant \lambda_2.$$

Найденные границы изменения Ф являются точными. Именно, если $x' = \pm 1$, $y' = 0$, то $\Phi = \lambda_1$; если $x' = 0$, $y' = \pm 1$, то $\Phi = \lambda_2$. Этим наше утверждение доказано. В самом деле, если данная форма в других координатах получит также канонический вид, но с другими коэффициентами, то, аналогично рассуждая, покажем, что эта же форма на той же единичной окружности имеет другие границы изменения, что невозможно.

12. Если $\lambda_1 \neq \lambda_2$, то единственной будет и та координатная система, в которой данная форма получает канонический вид (не считая возможности менять ролями положительные и отрицательные направления осей). Докажем это. Имея $\lambda_1 \neq \lambda_2$, будем считать $\lambda_1 < \lambda_2$ и вернемся к предыдущим соотношениям. Заметим, что в данном случае при $x' \neq 0$ будет

$$\Phi=\lambda_1 x'^2+\lambda_2 y'^2<\lambda_2(x'^2+y'^2)=\lambda_2,$$

т. е. при $x'\neq 0$ имеем

$$\Phi<\lambda_2.$$

Только в двух точках $x'=0,\ y'=\pm 1$ получаем

$$\Phi=\lambda_2.$$

Таким образом, на единичной окружности есть единственная пара диаметрально противоположных точек, где Ф получает наибольшее значение. Именно через эти точки и проходит ось ординат той координатной системы, в которой форма имеет канонический вид. Тем самым единственность такой системы координат доказана.

Если $\lambda_1=\lambda_2$, то форма имеет канонический вид в любых прямоугольных координатах. В самом деле, обозначая $\lambda_1=\lambda_2$ просто через λ, получим в некоторых координатах

$$\Phi=\lambda(x'^2+y'^2).$$

Но если мы еще повернем координатные оси на любой угол и перейдем к другим координатам x'', y'', то $x'^2+y'^2=x''^2+y''^2$, т. е. Ф сохраняет канонический вид. Это обстоятельство уже отмечалось в предыдущем параграфе.

13. Пусть дана произвольная квадратичная форма

$$\Phi=Ax^2+2Bxy+Cy^2,$$

и пусть совершается преобразование координат

$$x=l_1x'+l_2y',$$
$$y=m_1x'+m_2y'$$

при любом повороте осей. Подставляя эти выражения x, y приведем Ф к новым координатам:

$$Ax^2+2Bxy+Cy^2=A'x'^2+2B'x'y'+C'y'^2.$$

Поскольку квадратичные формы, написанные здесь слева и справа, преобразуются друг в друга, они приводятся к одному и тому же каноническому виду. Следовательно, обе эти формы имеют одни и те же характеристические числа λ_1, λ_2. Напишем для каждой из этих форм характеристическое уравнение (см. п.5.2):

$$\lambda^2-(A+C)\lambda+(AC-B^2)=0,$$
$$\lambda^2-(A'+C')\lambda+(A'C'-B'^2)=0.$$

Эти уравнения имеют одни и те же корни λ_1, λ_2, следовательно, не отличаются друг от друга. Таким образом,

$$A'+C'=A+C,$$
$$A'C'-B'^2=AC-B^2$$

(хотя по отдельности *A', B', C'* не обязаны совпадать с *A, B, C).*

Мы видим, что величины $A+C$ и $AC — B^2$ не меняются при любом преобразовании формы к новым прямоугольным координатам; поэтому их называют *инвариантами* формы относительно преобразования прямоугольных координат.

Инвариант $AC—B^2$ называется дискриминантом квадратичной формы и обозначается обычно через δ;

$$\delta = AC - B^2 = \begin{vmatrix} A & B \\ B & C \end{vmatrix}.$$

Согласно теореме Виета

$$\delta = \lambda_1\lambda_2,$$

где λ_1, λ_2 — характеристические числа длиной формы.

14. Квадратичную форму будем называть *эллиптической,* если ее характеристические числа λ_1, λ_2 оба отличны от нули и имеют одинаковые знаки, *гиперболической,* если числа λ_1, λ_2 разных знаков, *параболической,* если одно из чисел λ_1, λ_2 равно нулю. Так как $\delta = \lambda_1\lambda_2$, то форма будет эллиптической, если $\delta > 0$, гиперболической, если $\delta < 0$, и параболической, если $\delta = 0$.

Рассмотрим эллиптическую форму Ф; допустим, что ее характеристические числа положительны: $\lambda_1 > 0, \quad \lambda_2 > 0$. Согласно п° 11 (считая $\lambda_1 \le \lambda_2$) имеют место неравенства:

$$\lambda_1 (x^2 + y^2) \leqslant \Phi \leqslant \lambda_2 (x^2 + y^2). \qquad (2)$$

Таким образом, в любой точке *M(x; y),* не считая начала координат, величина $\Phi > 0$. Такая форма называется *положительно определенной.* Если точка *M(x; y)* обегает единичную окружность $x^2+y^2=1$, то положительно определенная форма изменяется в постоянных положительных границах:

$$\lambda_1 \leqslant \Phi \leqslant \lambda_2.$$

Допустим, что $\lambda_1 < 0, \ \lambda_2 < 0 \, (\lambda_1 \leqslant \lambda_2)$. Из соотношений (2) заключаем, что в этом случае в любой точке (кроме начала координат) будет $\Phi < 0$. Такая форма называется *отрицательно определенной.* На единичной окружности отрицательно определенная форма изменяется в постоянных отрицательных границах. Итак, каждая эллиптическая форма

$$\Phi = Ax^2 + 2Bxy + Cy^2 \ \ (\delta = AC - B^2 > 0)$$

является либо положительно определенной, либо отрицательно определенной. Чтобы узнать, к какому из этих двух классов относится данная эллиптическая форма, достаточно проверить ее знак в какой-нибудь точке. Беря $x=1$, $y=0$, заключаем: эллиптическая форма будет положительно определенной, если $A>0$, и отрицательно определенной, если $A<0$.

Рассмотрим теперь параболическую форму. Допустим, что $\lambda_1=0,\ \lambda_2>0$; тогда

$$0\leqslant\Phi\leqslant\lambda_2(x^2+y^2).$$

Таким образом, в данном случае форма Ф повсюду положительна или равна нулю. Но, в огличие от эллиптической формы, параболическая форма обращается в нуль не только в начале координат. В частности, на единичной окружности имеем

$$0\leqslant\Phi\leqslant\lambda_2,$$

т. е. Ф изменяется в границах, одна из которых есть нуль, другая — положительное число λ_2.

Аналогичным образом можно установить, что в случае

$$\lambda_1<0,\quad \lambda_2=0$$

параболическая форма Ф повсюду отрицательна или равна нулю. На единичной окружности такая форма изменяется между отрицательным числом λ_2 и нулем.

Рассмотрим, наконец, гиперболическую форму. Пусть $\lambda_1<0,\ \lambda_2>0$. Так как λ_1 и λ_2 представляют собой границы изменения Ф на единичной окружности, то гиперболическая форма является знакопеременной.

Примеры. 1) Форма $\Phi=5x^2-4xy+5y^2$ является эллиптической, так как $\delta=21>0$, и положительно определенной, так как $A=5>0$. На единичной окружности $(x^2+y^2=1)$ она заключена в границах

$$3\leqslant 5x^2-4xy+5y^2\leqslant 7.$$

2) Форма $\Phi=x^2-2xy+y^2$ является параболической, так как $\delta=0$, Эта форма положительна в каждой точке, где $x\neq y$, и равна нулю в точках, где $x=y$. На единичной окружности она заключена в границах

$$0\leqslant x^2-2xy+y^2\leqslant 2.$$

3) Форма $\Phi=x^2+12xy+y^2$ является гиперболической, следовательно, знакопеременной, так как $\delta=-35<0$. На единичной окружности она заключена между числами —5 и +7.

4) Уравнение $5x^2-4xy+5y^2=-1$ не определяет никакого действительного образа, так как данная квадратичная форма является положительно определенной, а справа в этом уравнении стоит отрицательное число.

5.4. Приведение к каноническому виду общего уравнения линии второго порядка

15. Пусть дано общее уравнение линии второго порядка

$$Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0. \qquad (1)$$

Наша цель — привести уравнение (1) к каноническому виду и тем самым установить вид данной линии. Левую часть уравнения (1) образуют следующие группы членов:

1) квадратичная форма, состоящая из старших членов $Ax^2+2Bxy+Cy^2$;

2) линейная форма членов первой степени $2Dx+2Ey$;

3) свободный член *F*.

Предположим, что координатные оси данной системы координат поворачиваются на произвольный угол; тогда координаты изменятся по формулам вида

$$\begin{aligned} x &= l_1x'+l_2y', \\ y &= m_1x'+m_2y'. \end{aligned} \qquad (2)$$

Если мы перейдем в уравнении (1) к новым координатам, т. е. заменим *x, y* их выражениями по формулам (2), то каждая из указанных групп членов левой части уравнения преобразуется автономно (не влияя друг на друга). Направим новые координатные оси по главным направлениям квадратичной формы старших членов (их называют также главными направлениями данной линии). Тогда

1) квадратичная форма старших членов примет канонический вид:

$$Ax^2+2Bxy+Cy^2=\lambda_1x'^2+\lambda_2y'^2.$$

2) Следующим образом изменится группа членов первой степени:

$$\begin{aligned} 2Dx+2Ey &= 2D(l_1x'+l_2y')+2E(m_1x'+m_2y')= \\ &= 2(Dl_1+Em_1)x'+2(Dl_2+Em_2)y'=2\mu_1x'+2\mu_2y' \end{aligned}$$

(через μ_1, μ_2 обозначены новые коэффициенты).

3) Свободный член останется без изменения.

Мы получим уравнение данной линии в новых координатах:

$$\lambda_1x'^2+\lambda_2y'^2+2\mu_1x'+2\mu_2y'+F=0. \qquad (3)$$

Дальнейшее упрощение этого уравнения производится в зависимости от типа квадратичной формы старших членов. Согласно п.5.3 будем обозначать далее через δ дискриминант квадратичной формы старших членов первоначального уравнения:

$$\delta = AC - B^2.$$

16. Предположим, что $\delta \neq 0$. Так как $\delta = \lambda_1\lambda_2$, то

$$\lambda_1 \neq 0, \ \lambda_2 \neq 0$$

и мы можем полученное выше уравнение (3) переписать так:

$$\lambda_1\left(x'^2 + 2\frac{\mu_1}{\lambda_1}x'\right) + \lambda_2\left(y'^2 + 2\frac{\mu_2}{\lambda_2}y'\right) = -F.$$

Отсюда

$$\lambda_1\left(x' + \frac{\mu_1}{\lambda_1}\right)^2 + \lambda_2\left(y' + \frac{\mu_2}{\lambda_2}\right)^2 = -F + \frac{\mu_1^2}{\lambda_1} + \frac{\mu_2^2}{\lambda_2}. \qquad (4)$$

Совершим параллельный перенос системы повернутых осей на величину $-\frac{\mu_1}{\lambda_1}$ в направлении оси *Ox'* и на величину $-\frac{\mu_2}{\lambda_2}$ в направлении оси *Oy'*; переход к новым координатам дается формулами:

$$x' = x'' - \frac{\mu_1}{\lambda_1},$$

$$y' = y'' - \frac{\mu_2}{\lambda_2}.$$

Если мы обозначим правую часть (4) одной буквой *H,* то уравнение данной линии в последних координатах примет канонический вид:

$$\lambda_1 x''^2 + \lambda_2 y''^2 = H. \qquad (5)$$

Возможны два случая.

1) Числа λ_1 и λ_2 одного знака; будем считать, что $\lambda_1 > 0, \ \lambda_2 > 0$ (в противном случае изменим знаки всех членов уравнения). Если $H \neq 0$, то уравнение (5) приводится к виду

$$\frac{x''^2}{\frac{H}{\lambda_1}} + \frac{y''^2}{\frac{H}{\lambda_2}} = 1;$$

отсюда видно, что при $H > 0$ уравнение (5) определяет эллипс с полуосями

$$a = \sqrt{\frac{H}{\lambda_1}}, \ b = \sqrt{\frac{H}{\lambda_2}}.$$

Если $H = 0$, то уравнение (5) определяет единственную вещественную точку: $x'' = 0$, $y'' = 0$. Однако в этом случае принято говорить, что уравнение (5) есть уравнение *вырожденного эллипса.*

Если $H < 0$ (при $\lambda_1 > 0,\ \lambda_2 > 0$), то **уравнение (5) никакого вещественного образа не определяет.** В этом случае говорят также, что уравнение (5) есть уравнение *мнимого эллипса.*

2) Числа λ_1 и λ_2 разных знаков; будем считать, что $\lambda_1 > 0,\ \lambda_2 < 0$. Если при этом $H > 0$, то уравнение (5) определяет *гиперболу,* которая пересекает ось абсцисс и имеет полуоси:

$$a = \sqrt{\frac{H}{\lambda_1}},\ b = \sqrt{-\frac{H}{\lambda_2}};$$

если $H<0$, то также получается гипербола, но пересекающая ось ординат. Если $H = 0$, то уравнение (5) определяет пару прямых, проходящих через начало координат. В самом деле, левая часть уравнения (5) разлагается на вещественные множители первой степени; при $H=0$ имеем

$$(\sqrt{\lambda_1}\,x'' + \sqrt{-\lambda_2}\,y'')(\sqrt{\lambda_1}\,x'' - \sqrt{-\lambda_2}\,y'') = 0;$$

следовательно, линия состоит из прямых $\sqrt{\lambda_1}x'' + \sqrt{-\lambda_2}y'' = 0$ и $\sqrt{\lambda_1}x'' - \sqrt{-\lambda_2}y'' = 0$ (например, уравнение $4x''^2 - 9y''^2 = 0$ определяет пару прямых $2x'' + 3y'' = 0,\quad 2x'' - 3y'' = 0$).

В последнем случае говорят также, что уравнение (о) определяет *вырожденную гиперболу.*

Вернемся теперь к исходному уравнению (1). По старшим коэффициентам этого уравнения мы сразу можем подсчитать $\delta = AC - B^2$. Но $\delta = \lambda_1\lambda_2$; таким образом, если $\delta > 0$, то λ_1 и λ_2 — числа одного знака; если $\delta < 0$, то $\lambda_1,\ \lambda_2$ —числа разных знаков. Отсюда заключаем: *если данное уравнение* (1) *имеет* $\delta = AC - B^2 > 0$, *то оно определяет эллипс* (вещественный, мнимый или вырожденный); *если* $\delta < 0$, *то уравнение определяет гиперболу* (настоящую или вырожденную).

17. Рассмотрим теперь уравнение (1) при условии $\delta=0$. Так как $\delta = \lambda_1\lambda_2$, то на этот раз одно из чисел $\lambda_1,\ \lambda_2$ равно нулю. Будем считать, что

$$\lambda_1 \neq 0,\ \lambda_2 = 0.$$

Уравнение (3) будет иметь вид

$$\lambda_1 x'^2 + 2\mu_1 x' + 2\mu_2 y' + F = 0,$$

или

$$\lambda_1\left(x'+\frac{\mu_1}{\lambda_1}\right)^2+2\mu_2 y'+\left(F-\frac{\mu_1^2}{\lambda_1}\right)=0.$$

Обозначая величину, стоящую в последней скобке, одной буквой *К,* получим

$$\lambda_1\left(x'+\frac{\mu_1}{\lambda_1}\right)^2+2\mu_2 y'+K=0. \qquad (6)$$

Дальнейшее преобразование уравнения (6) зависит от μ_2.

1) Если $\mu_2\neq 0$, то уравнение (6) можно переписать так:

$$\lambda_1\left(x_1+\frac{\mu_1}{\lambda_1}\right)^2+2\mu_2\left(y'+\frac{K}{2\mu_2}\right)=0.$$

Совершим параллельный перенос системы осей *Ox', Oy'* на величину

$$-\frac{\mu_1}{\lambda_1}$$

в направлении оси *Ox'* и на величину' $-\frac{K}{2\mu_2}$ в направлении оси *Oy'*; переход к новым координатам дается формулами

$$x'=x''-\frac{\mu_1}{\lambda_1}, \quad y'=y''-\frac{K}{2\mu_2}.$$

Мы получим каноническое уравнение параболы:

$$\lambda_1 x''^2+2\mu_2 y''=0.$$

2) Если $\mu_2=0$, то достаточно совершить параллельный перенос системы осей только в направлении оси *Ox'* на величину

$$-\frac{\mu_1}{\lambda_1};$$

тогда

$$x'=x''-\frac{\mu_1}{\lambda_1}, \quad y'=y''$$

и уравнение (6) приметследующий канонический вид:

$$\lambda_1 x''^2+K=0. \qquad (7)$$

Предположим $\lambda_1 > 0$. Тогда, если $K<0$, то уравнение (7) определяет пару параллельных прямых:

$$\sqrt{\lambda_1}x''+\sqrt{-K}=0, \quad \sqrt{\lambda_1}x''-\sqrt{-K}=0 \qquad (8)$$

(например, уравнение $4x''$ — $9 = 0$ определяет пару параллельных прямых: $2x''+3=0,\ 2x''-3=0$). Если $K=0$, то прямые (8) сливаются. Если $K> 0$, то уравнение (7) **никакого вещественного образа** не определяет. Однако в этом случае говорят, что уравнение (7) определяет *пару мнимых параллельных прямых* (при $K> 0$ в уравнениях

(8) число $\sqrt{-K}$ будет мнимым). Вообще уравнение (7) называют уравнением *выроженной параболы.*

Возвращаясь к исходному уравнению (1), приходим к следующему заключению: *если* $\delta = AC - B^2 = 0$, *то уравнение* (1) *является уравнением параболы* (обыкновенной или выроженной).

18. Подводя итог всему исследованию, мы приходим к следующему общему выводу: *каждое уравнение второй степени*

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

определяет либо эллипс {если $\delta = AC - B^2 > 0$), *либо гиперболу (если* $\delta < 0$), *либо параболу (если* $\delta = 0$). При этом следует учитывать мнимый и выроженный эллипсы, а также выроженные гиперболы и параболы.

5.5. Уравнения центра. Признак выроождения линии второго порядка.

19. Вернемся к п°16, где мы приводили к простейшему виду общее уравнение второй степени в случае, когда оно определяет эллипс или гиперболу (т. е. в случае $\delta \neq 0$). Мы упростили уравнение в два приема: сначала повернули оси, придав им главные направления кривой, затем совершили параллельный перенос повернутых осей. Легко понять геометрический смысл последнего преобразования: начало координат оказалось в центре данного эллипса или гиперболы. Можно провести эти действия в ином порядке: во-первых, перенести начало координат в центр линии, затем нужным образом повернуть оси. Имея в виду такой план, выведем уравнения, которые определяют центр линии второго порядка (если он есть) непосредственно по общему уравнению этой линии.

20. Прежде всего нужно уточнить, что мы имеем в виду, когда говорим о центре линии второго порядка вообще.

Некоторую точку S мы называем *центром* данной линии второго порядка, если после переноса начала координат в точку S уравнение линии примет вид

$$Ax^2 + 2Bxy + Cy^2 = H,$$

т. е. не будет содержать членов первой степени. Левая часть такого уравнения не меняется при одновременном изменении знаков текущих координат; это означает, что точки линии расположены парами, симметрично относительно точки *S,* т. е. что точка *S* является центром симметрии линии (по существу, понятие центра уже употреблялось нами раньше, в п.5. 2; сейчас мы только сформулировали это понятие более определенным образом).

Пусть дана линия второго порядка:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \qquad (1)$$

Требуется найти ее центр (если он есть). Предполагая, что центр имеется, обозначим его, как и раньше, буквой *S;* обозначим через x_0, y_0 искомые координаты центра (в данной координатной системе). Перенесем начало координат в точку *S;* при этом координаты произвольной точки изменятся поформулам

$$x = \tilde{x} + x_0, \quad y = \tilde{y} + y_0, \qquad (2)$$

где $\tilde{x}$, $\tilde{y}$ — новые координаты той же точки. Перейдем в уравнении (1) к новым координатам. С этой целью, подставив в левую часть уравнения (1) вместо *x, y* их выражения по формулам (2) и приведя подобные члены (подобные относительно новых текущих координат $\tilde{x}$, $\tilde{y}$), мы получим

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F =$$
$$= A\tilde{x}^2 + 2B\tilde{x}\tilde{y} + C\tilde{y}^2 + 2\tilde{D}\tilde{x} + 2\tilde{E}\tilde{y} + \tilde{F}, \qquad (3)$$

где

$$\left.\begin{array}{l} \tilde{D} = Ax_0 + By_0 + D, \\ \tilde{E} = Bx_0 + Cy_0 + E, \\ \tilde{F} = Ax_0^2 + 2Bx_0y_0 + Cy_0^2 + 2Dx_0 + 2Ey_0 + F; \end{array}\right\} \qquad (4)$$

коэффициенты старших членов остаются прежними.

Преобразование левой части уравнения (1) происходит согласно этим формулам в любом случае, т. е. независимо от того, что представляет собой точка *S*. Точка *S* будет центром, если $\tilde{D} = 0$, $\tilde{E} = 0$. Отсюда получаем уравнения центра:

$$\left.\begin{array}{l} Ax_0 + By_0 + D = 0, \\ Bx_0 + Cy_0 + E = 0. \end{array}\right\} \qquad (5)$$

Решая их совместно, найдем центр *S* (x_0, y_0). Система (5) может оказаться несовместной; тогда центра у данной линии нет.

Напишем определитель системы (5):

$$\delta = \begin{vmatrix} A & B \\ B & C \end{vmatrix} = AC - B^2.$$

Мы видим, что это — известный уже нам дискриминант группы старших членов.

Если $\delta \neq 0$ то система (5) совместна и имеет единственное решение. Следовательно, если $\delta \neq 0$, то данная линия имеет единственный центр. Такая линия второго порядка называется *центральной.* В случае центральной линии координаты центра выражаются формулами:

$$x_0 = \frac{\begin{vmatrix} B & D \\ C & E \end{vmatrix}}{\begin{vmatrix} A & B \\ B & C \end{vmatrix}}, \quad y_0 = \frac{\begin{vmatrix} D & A \\ E & B \end{vmatrix}}{\begin{vmatrix} A & B \\ B & C \end{vmatrix}}. \qquad (6)$$

Отсюда и из (4) найдем $\tilde{F}$. Вычисление $\tilde{F}$ можно значительно упростить, если выразить $\tilde{F}$ со следующей группировкой членов:

$$\tilde{F} = (Ax_0 + By_0 + D)\, x_0 + (Bx_0 + Cy_0 + E)\, y_0 + \\ + (Dx_0 + Ey_0 + F);$$

вследствие системы (5) имеем

$$\tilde{F} = Dx_0 + Ey_0 + F.$$

Используя (6), найдем

$$\tilde{F} = \frac{D\begin{vmatrix} B & D \\ C & E \end{vmatrix} + E\begin{vmatrix} D & A \\ E & B \end{vmatrix} + F\begin{vmatrix} A & B \\ B & C \end{vmatrix}}{\begin{vmatrix} A & B \\ B & C \end{vmatrix}}. \qquad (7)$$

Числитель полученной дроби может быть записан в виде определителя третьего порядка:

$$D\begin{vmatrix} B & D \\ C & E \end{vmatrix} + E\begin{vmatrix} D & A \\ E & B \end{vmatrix} + F\begin{vmatrix} A & B \\ B & C \end{vmatrix} = \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix}.$$

Этот определитель обозначается символом Δ:

$$\Delta = \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix}$$

и называется дискриминантом левой части общего уравнения (1). Из (7) окончательно имеем:

$$\tilde{F} = \frac{\Delta}{\delta}.$$

Итак, если линия, заданная уравнением (1), является центральной ($\delta \neq 0$), то после переноса начала координат в ее центр данное уравнение (1) приводится к виду

$$A\tilde{x}^2 + 2B\tilde{x}\tilde{y} + C\tilde{y}^2 + \frac{\Delta}{\delta} = 0 \qquad (8)$$

(коэффициенты *A, B, C*—прежние). Совершая теперь надлежащий поворот осей, можно привести уравнение (8) к каноническому виду:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \frac{\Delta}{\delta} = 0. \qquad (9)$$

Таким образом, свободный член *H* в уравнении (5) п.5. 4 может быть подсчитан по данному уравнению (1) сразу, без того, чтобы на самом деле выполнять преобразование координат $H = -\frac{\Delta}{\delta}$. Тем самым может быть написано и все уравнение, поскольку характеристические числа λ_1, λ_2 также непосредственно находятся по коэффициентам старших членов (см. п.5.2).

21. В п.5.4 установлено, что уравнение (5) п.5.4 определяет вырожденную линию при *H*=0. Отсюда заключаем: центральная линия второго порядка является вырожденной тогда и только тогда, когда $\Delta = 0$.

В следующем п° мы покажем, что точно такое же условие характеризует вырожденную параболу.

22. Предположим, что исходное уравнение (1) имеет $\delta=0$. Тогда оно приводится к виду (6) п.5.4. Мы установим сейчас формулу, которая позволяет сразу подсчитать коэффициент μ_2 в этом уравнении, зная коэффициенты исходного уравнения (1). Именно,

$$\mu_2 = \pm\sqrt{-\frac{\Delta}{\lambda_1}};$$

одновременно будет показано, что подкоренное выражение здесь ≥ 0. Так как $\delta = AC - B^2 = 0$, то оба коэффициента *A* и *C* не могут быть равными нулю (тогда все три коэффициента были бы равны нулю и уравнение не имело бы старших членов). Допустим, что $A \neq 0$; будем считать, что $A > 0$ (в противном случае изменим знаки всех членов). Положим

$$\alpha = \sqrt{A}, \quad \beta = \frac{B}{\alpha};$$

тогда $A = \alpha^2$, $B = \alpha\beta$. Отсюда и из условия $AC - B^2 = 0$ найдем $C = \beta^2$. Согласно п°17, где впервые появился коэффициент μ_2, имеем

$$\mu_2 = Dl_2 + Em_2;$$

здесь $\{l_2;\ m_2\}$—вектор, который определяет главное направление соответствующее характеристическому числу $\lambda_2 = 0$; при этом $l_2^2 + m_2^2 = 1$. Подставим $\lambda = \lambda_2 = 0$ в систему (11) п.5. 2, одновременно полагая $A = \alpha^2,\ B = \alpha\beta,\ C = \beta^2$; система примет вид:

$$\alpha^2 l + \alpha\beta m = 0,$$
$$\alpha\beta l + \beta^2 m = 0.$$

В качестве решения этой системы можно взять

$$l = -\beta,\quad m = \alpha.$$

Нормируя это решение, получим

$$l_2 = \frac{-\beta}{\pm\sqrt{\alpha^2+\beta^2}},\ m_2 = \frac{\alpha}{\pm\sqrt{\alpha^2+\beta^2}}.$$

С другой стооонны. из характеристического уравнения (12) п.5.2 найдем $\lambda_1 = A + C = \alpha^2 + \beta^2$. Следовательно,

$$l_2 = \frac{-\beta}{\pm\sqrt{\lambda_1}},\qquad m_2 = \frac{\alpha}{\pm\sqrt{\lambda_1}},$$

$$\mu_2 = Dl_2 + Em_2 = \pm\frac{E\alpha - D\beta}{\sqrt{\lambda_1}}.$$

Подсчитаем дискриминант левой части данного уравнения:

$$\Delta = \begin{vmatrix} \alpha^2 & \alpha\beta & D \\ \alpha\beta & \beta^2 & E \\ D & E & F \end{vmatrix} = D(\alpha\beta E - \beta^2 D) - E(\alpha^2 E - \alpha\beta D) =$$
$$= -E\alpha(E\alpha - D\beta) + D\beta(E\alpha - D\beta) = -(E\alpha - D\beta)^2.$$

Таким образом, $(E\alpha - D\beta)^2 = -\Delta$. Отсюда

$$\mu_2 = \pm\sqrt{-\frac{\Delta}{\lambda_1}},$$

что и утверждалось. Подкоренное выражение здесь не отрицательно, так как $\Delta \leqslant 0,\ \lambda_1 > 0$.

Рассмотрим две возможности (считая $\delta = 0$):

1) $\Delta \neq 0$; тогда $\mu_2 \neq 0$ и данное уравнение (1) определяет параболу в собственном смысле слова (см. п.5.4). Каноническое уравнение этой параболы можно написать теперь в готовом виде:

$$\lambda_1 x''^2 \pm 2\sqrt{-\frac{\Delta}{\lambda_1}}\, y'' = 0; \qquad (10)$$

выбор знака определяется в зависимости от выбора положительного направления оси ординат.

2) $\Delta = 0$; тогда $\mu_2 = 0$ и парабола будет вырожденной.

23. Теперь мы можем высказать общее утверждение: *уравнение* (1) *определяет вырожденную линию второго порядка тогда и только тогда, когда дискриминант его левой части равен нулю:* $\Delta = 0$.

24. **Пример.** Дано уравнение
$5x^2+8xy+5y^2-18x-18y+9=0;$
установить, какую именно линию определяет это уравнение, и привести его к каноническому виду.

Решение. Имеем $\delta=9>0,\ \Delta=-81$, следовательно, данная линия является невырожденным эллипсом. Составим характеристическое уравнение группы старших членов: $\lambda^2-10\lambda+9=0;$ отсюда $\lambda_1=9,\ \lambda_2=1.$ Каноническое уравнение можем написать теперь сразу; $9x^2+y^2-9=0$ (см. п° 20), или

$$\frac{x^2}{1}+\frac{y^2}{9}=1.$$

Таким образом, данная линия является действительным эллипсом с полуосями 3 и 1. Если нас, кроме того, интересует расположение этого эллипса, то придется находить его центр и главные направления (рис. 4).

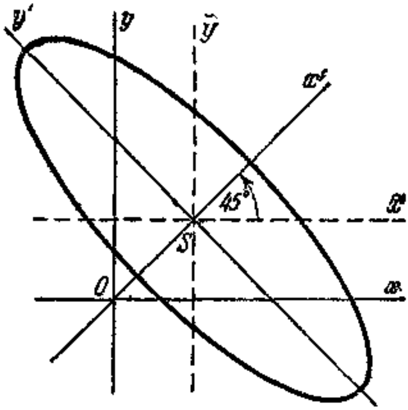


Рис. 4.

Составим уравнение центра:

$$5x_0+4y_0-9=0,$$
$$4x_0+5y_0-9=0.$$

Отсюда x_0=1, y_0=1. Перенесем начало координат в центр (без поворота осей); при этом координаты изменятся по формулам $x=\tilde{x}+1,\quad y=\tilde{y}+1,$ а уравнение примет вид

$$5\tilde{x}^2+8\tilde{x}\,\tilde{y}+5\tilde{y}^2-9=0 \qquad (*)$$

(см. п°20).

Найдем главное направление, соответствующее характеристическому числу $\lambda_1 = 9$. Для этого подставим в систему (11) п.5. 2 $A = 5,\ B = 4,\ C = 5,\ \lambda = \lambda_1 = 9$; получим

$$-4l + 4m = 0,$$
$$4l - 4m = 0.$$

В качестве решения этой системы можно взять l=1, m=1.
В данном случае легко определяется угол α, который составляет с осью $O\tilde{x}$ это главное направление:

$$\operatorname{tg}\alpha = \frac{m}{l} = 1,\ \alpha = \frac{\pi}{4}.$$

Значит, оси нужно повернуть на 45°; отсюда имеем формулы соответствующего преобразования координат:

$$\tilde{x} = \frac{x' - y'}{\sqrt{2}},\quad \tilde{y} = \frac{x' + y'}{\sqrt{2}}.$$

Преобразование уравнения (*) по этим формулам приведет нас согласно теории к каноническому уравнению $9x'^2 + y'^2 - 9 = 0$, в чем можно убедиться также и непосредственной проверкой.

Пример. Дано уравнение

$$4x^2 - 4xy + y^2 - 2x - 14y + 7 = 0;$$

установить, какую именно линию определяет это уравнение, и привести его к каноническому виду.

Решение. Имеем $\delta = 0,\ \Delta = -225$. Следовательно, данное уравнение определяет обыкновенную (невырожденную) параболу. Составим характеристическое уравнение группы старших членов $\lambda^2 - 5\lambda = 0$; отсюда $\lambda_1 = 5,\ \lambda_2 = 0$. Согласно п° 22 каноническое уравнение кривой имеет вид $5x^2 \pm 6\sqrt{5}y = 0$. Если нас, кроме того, интересует расположение нашей параболы, то придется находить ту координатную систему, в которой уравнение этой параболы имеет канонический вид.

Найдем главные направления данной параболы (относительно старых осей). В данном случае система (11) п.5. 2 при

$$\lambda = \lambda_1 = 5$$

будет

$$-l - 2m = 0,$$
$$-2l - 4m = 0.$$

В качестве решения системы возьмем $l = -2$, $m = 1$. Вектор $\{-2; 1\}$ составляет с осью Ox тупой угол. Желая повернуть оси на острый угол, поменяем ролями главные направления, будем считать $\lambda_2 = 5$, $\lambda_1 = 0$.

Тогда найденное решение дает второе главное направление; нормируя решение, получим

$$l_2 = -\frac{2}{\sqrt{5}}, \quad m_2 = \frac{1}{\sqrt{5}}.$$

В таком случае для первого главного направления (которое отвечает числу $\lambda_1 = 0$) имеем

$$l_1 = \frac{1}{\sqrt{5}}, \quad m_1 = \frac{2}{\sqrt{5}}.$$

Чтобы направить оси по этим главным направлениям, нужно повернуть их на угол α, имея $\operatorname{tg}\alpha = 2$. Соответствующие формулы преобразования координат будут:

$$x = \frac{x' - 2y'}{\sqrt{5}}, \quad y = \frac{2x' + y'}{\sqrt{5}}.$$

Переходя в данномуравнении к новым координатам, найдем

$$5y'^2 - 6\sqrt{5}x' - 2\sqrt{5}y' + 7 = 0.$$

Заметим, что квадратичный член здесь определяется сразу согласно теории так как $\lambda_1 = 0$, $\lambda_2 = 5$; свободный член переходит из данного уравнения без изменения (см. п° 15). Таким образом, фактически преобразование потребовалось выполнять лишь по отношению к членам первой степени.

Перепишем предыдущее уравнение так:

$$5\left(y' - \frac{\sqrt{5}}{5}\right)^2 - \\ -6\sqrt{5}\left(x' - \frac{\sqrt{5}}{5}\right) = 0.$$

Теперь видно, что нужен параллельный перенос системы *Ox', Oy'* на величину

$$\frac{\sqrt{5}}{5}$$

в направлении оси *Ox'* и на величину

$$\frac{\sqrt{5}}{5}$$

в направлении оси *Oy';* при этом координаты преобразуются по формулам

$$x' = x'' + \frac{\sqrt{5}}{5}, \quad y' = y'' + \frac{\sqrt{5}}{5}$$

в данное уравнение принимает канонический вид:

$$5y''^2 - 6\sqrt{5}x'' = 0.$$

Начало последней системы координат находится в вершине этой параболы, координаты вершины в системе x', y' суть $\left(\frac{\sqrt{5}}{5}; \frac{\sqrt{5}}{5}\right)$,
а в системе x, y суть

$$\left(-\frac{1}{5}; \frac{3}{5}\right).$$

Расположение параболы показано на рис. 5.

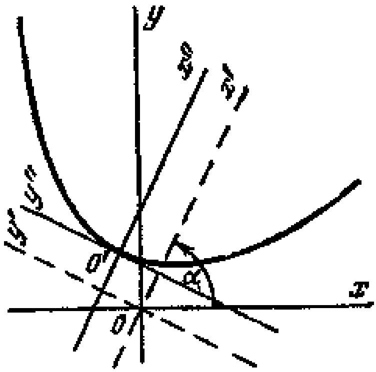


Рис. 5.

Микромодуль 14
Общая теория поверхностей второго рорядка

5.6. Преобразование декартовых прямоугольных координат в пространстве

25. Пусть дана декартова система координат с осями *Ох, Оу, Oz*. Рассмотрим новую систему координат с началом в точке *О',* оси которой $O'x'$, $O'y'$, $O'z'$ соответственно параллельны осям *Ох, Оу, Oz* старой системы и имеют те же направления (масштаб остается прежним). Пусть известны старые координаты нового начала *О'* (*a*; *b*; *с*). Тогда

$$\left.\begin{aligned} x &= x' + a, \\ y &= y' + b, \\ z &= z' + c, \end{aligned}\right\} \qquad (1)$$

где *х, у, z* — старые координаты произвольной точки, *х',у', z'* — новые координаты той же точки. Формулы (1) выражают преобразование координат при параллельном перемещении осей. Доказательство этих формул по существу очевидно. В самом деле, так как система осей параллельно перемещается на величину *а* в направлении *Ох,* на величину *b* в направлении *Оу* и на величину *с* в направлении *Oz,* то абсциссы всех точек уменьшаются на *а,* ординаты — на *b,* аппликаты — на *с* (*х'*=*х*— *а* и т. д.).

26. Теперь мы установим формулы преобразования декартовых прямоугольных координат при таком изменении координатной системы, когда изменяются ее оси (однако остаются перпендикулярными друг к другу), а начало координат и масштаб остаются неизменными.

Пусть *Ох, Оу, Oz* — старые, *Ох', Оу',* Oz' — новые координатные оси. Будем считать, что нам известны углы, которые образует каждая ось новой системы с каждой осью старой; обозначения этих углов дадим следующей таблицей:

	Ox	Oy	Oz
Ox'	α_1	β_1	γ_1
Oy'	α_2	β_2	γ_2
Oz'	α_3	β_3	γ_3

(2)

Обозначим через $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ и $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ базисные векторы старых и новых осей. Напишем разложение каждого вектора $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ по старому базису:

$$\left.\begin{aligned} \boldsymbol{i}' &= l_1\boldsymbol{i} + m_1\boldsymbol{j} + n_1\boldsymbol{k}, \\ \boldsymbol{j}' &= l_2\boldsymbol{i} + m_2\boldsymbol{j} + n_2\boldsymbol{k}, \\ \boldsymbol{k}' &= l_3\boldsymbol{i} + m_3\boldsymbol{j} + n_3\boldsymbol{k}. \end{aligned}\right\} \tag{3}$$

Так как каждый из векторов $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ является единичным, то для каждого из них коэффициентами разложения будут служить направляющие косинусы. Таким образом, вся таблица коэффициентов формул (3) дается следующим равенством:

$$\begin{pmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{pmatrix} = \begin{pmatrix} \cos\alpha_1 & \cos\beta_1 & \cos\gamma_1 \\ \cos\alpha_2 & \cos\beta_2 & \cos\gamma_2 \\ \cos\alpha_3 & \cos\beta_3 & \cos\gamma_3 \end{pmatrix}, \tag{4}$$

которое нужно понимать так: $l_1 = \cos\alpha_1$, $m_1 = \cos\beta_1$ и т. д.

Обозначим через M произвольную точку пространства, через $(x; y; z)$ —старые координаты этой точки, через $(x'; y'; z')$ — ее новые координаты. Имеет место векторное равенство:

$$x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} = x'\boldsymbol{i}' + y'\boldsymbol{j}' + z'\boldsymbol{k}', \tag{5}$$

поскольку его левая и правая части представляют собой разложение одного и того же вектора OM (по разным базисам). Заменяя в равенстве (5) векторы $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ их выражениями по формулам (3), получим

$$\begin{aligned} x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} = {} & x'(l_1\boldsymbol{i} + m_1\boldsymbol{j} + n_1\boldsymbol{k}) + \\ & + y'(l_2\boldsymbol{i} + m_2\boldsymbol{j} + n_2\boldsymbol{k}) + \\ & + z'(l_3\boldsymbol{i} + m_3\boldsymbol{j} + n_3\boldsymbol{k}), \end{aligned}$$

или

$$\begin{aligned} x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} = {} & (l_1x' + l_2y' + l_3z')\,\boldsymbol{i} + \\ & + (m_1x' + m_2y' + m_3z')\,\boldsymbol{j} + \\ & + (n_1x' + n_2y' + n_3z')\,\boldsymbol{k}. \end{aligned} \tag{6}$$

Так как коэффициенты разложения вектора по базису *i*, *j*, *k* определяются этим вектором однозначно, то из (6) вытекают три равенства:

$$\left.\begin{aligned} x &= l_1x' + l_2y' + l_3z', \\ y &= m_1x' + m_2y' + m_3z'. \\ z &= n_1x' + n_2y' + n_3z'. \end{aligned}\right\} \qquad (7)$$

Искомые формулы найдены. Если заменить в них коэффициенты l_1, l_2, ..., n_3 согласно (4), то они примут закопченный вид:

$$\left.\begin{aligned} x &= x'\cos\alpha_1 + y'\cos\alpha_2 + z'\cos\alpha_3, \\ y &= x'\cos\beta_1 + y'\cos\beta_2 + z'\cos\beta_3, \\ z &= x'\cos\gamma_1 + y'\cos\gamma_2 + z'\cos\gamma_3. \end{aligned}\right\} \qquad (8)$$

27. Формулы (7) (или (8)) выражают старые координаты произвольной точки через ее новые координаты. Часто требуются обратные формулы, выражающие новые координаты через старые. Чтобы получить их, достаточно в формулах (7) поменять ролями старые и новые координаты, одновременно транспонируя таблицы (2) и (4) (т. е. меняя ролями их строчки и столбцы). Мы получим

$$\left.\begin{aligned} x' &= l_1x + m_1y + n_1z, \\ y' &= l_2x + m_2y + n_2z, \\ z' &= l_3x + m_3y + n_3z. \end{aligned}\right\} \qquad (9)$$

28. Коэффициенты l_1, l_2, ..., n_3 в формулах (7) удовлетворяют определенным условиям. Найдем их. Прежде всею учтем, что новые базисные векторы $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ являются единичными; поэтому

$$\left.\begin{aligned} l_1^2 + m_1^2 + n_1^2 &= 1, \\ l_2^2 + m_2^2 + n_2^2 &= 1, \\ l_3^2 + m_3^2 + n_3^2 &= 1. \end{aligned}\right\} \qquad (10)$$

Далее, векторы $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ попарно перпендикулярны друг к другу; поэтому их попарно взятые скалярные произведения должны быть равны нулю:

$$\left.\begin{aligned} l_1l_2 + m_1m_2 + n_1n_2 &= 0, \\ l_1l_3 + m_1m_3 + n_1n_3 &= 0, \\ l_2l_3 + m_2m_3 + n_2n_3 &= 0. \end{aligned}\right\} \qquad (11)$$

Наконец, если тройки векторов $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ и $\boldsymbol{i}'$, $\boldsymbol{j}'$, $\boldsymbol{k}'$ обе правые (или обе левые), то смешанное произведение *i'j'k'* положительно и равно объему единичного куба, т. е. *i'j'k'*=1; следовательно, в этом случае

$$\begin{vmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{vmatrix} = 1. \qquad (12a)$$

Если же тройки векторов $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ и $\boldsymbol{i'}, \boldsymbol{j'}, \boldsymbol{k'}$ ориентированы по-разному (одна правая, другая левая), то

$$\begin{vmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{vmatrix} = -1. \qquad (12b)$$

На основании сказанного ясно, что существует два вида преобразований декартовых прямоугольных координат: сохраняющие ориентацию координатного базиса и нарушающие ее. Первые характеризуются условием (12a), вторые — условием (12b).

Заметим, что соотношения (10), (11) являются не только необходимыми, но и достаточными условиями того, что формулы (7) выражают преобразования прямоугольных координат (с неизменным масштабом). В самом деле, допустим, что нам дана система прямоугольных координат с осями *Ox, Oy, Oz* и даны формулы (7) с какими-то коэффициентами $l_1, l_2, \ldots, n_3$. По данным коэффициентам построим векторы $\boldsymbol{i'}, \boldsymbol{j'}, \boldsymbol{k'}$ согласно (3). Тогда, если соблюдены условия (10), то векторы $\boldsymbol{i'}, \boldsymbol{j'}, \boldsymbol{k'}$ — все единичные. Если соблюдены условия (11), то векторы $\boldsymbol{i'}, \boldsymbol{j'}, \boldsymbol{k'}$ попарно перпендикулярны. Значит, при соблюдении (10) и (11) мы действительно имеем переход к новой, также прямоугольной системе координат (с тем же масштабом). Сохраняется ли при этом ориентация базиса, или нарушается, легко установить проверкой условий (12a, b).

29. Если начало координат переносится в точку *O' (a; b;* c) и вместе с тем меняются направления осей, то координаты преобразуются по формулам

$$\begin{aligned} x &= l_1x' + l_2y' + l_3z' + a, \\ y &= m_1x' + m_2y' + m_3z' + b, \\ z &= n_1x' + n_2y' + n_3z' + c, \end{aligned}$$

где коэффициенты $l_1, l_2, \ldots, n_3$ определяются согласно (4) по данной таблице *(Z).*

5.7. Некоторые общие выводы, основанные на формулах преобразования координат

30. После того как получены формулы преобразования декартовых прямоугольных координат в пространстве, легко доказать следующую теорему:

Если поверхность определяется алгебраическим уравнением степени n в какой-нибудь системе декартовых прямоугольных координат, то в любой другой системе таких же координат она определяется также алгебраическим уравнением и той же степени n.

31. Из предыдущей теоремы тотчас следует, что линия пересечения поверхности второго порядка с плоскостью есть алгебраическая линия не выше второго порядка.

Доказательство. Пусть даны какая-нибудь поверхность второго порядка и некоторая плоскость α. Выберем декартову прямоугольную систему координат так, чтобы ее оси *Ox* и *Oy* расположились в плоскости α. Согласно п° 30 данная поверхность в выбранной системе координат, как и во всякой декартовой системе координат, будет определяться некоторым уравнением второй степени:

$$Ax^2+By^2+Cz^2+2Dxy+2Exz+2Fyz+2Gx+ \\ +2Hy+2Kz+L=0. \qquad (1)$$

Плоскость α имеет в этой системе координат уравнение z=0. Линия пересечения данной поверхности с плоскостью α определяется уравнением (1) совместно с уравнением z=0. Полагая в уравнении (1) z=0, мы получим

$$Ax^2+By^2+2Dxy+2Gx+2Hy+L=0. \qquad (2)$$

Это и есть уравнение линии пересечения данной поверхности с плоскостью α в системе координат, которая на плоскости α задана осями *Ox* и *Oy*. Но уравнение (2) представляет собой алгебраическое уравнение не выше второй степени. Тем самым утверждение доказано.

Замечание. Степень уравнения (2) может оказаться ниже двух, это будет именно в случае, когда *A, B* и *D* обращаются в нуль. Например, параболический цилиндр пересекается со своей плоскостью симметрии по прямой, т. е. по линии первого порядка. Но такие случаи являются исключительными; вообще говоря, поверхность второго порядка пересекается с плоскостью по линии второго порядка.

32. ***Круглый конус есть поверхность второго порядка***. Отсюда и на основании сказанного выше следует, что каждая плоскость, не проходящая через вершину круглого конуса, пересекает его по невырожденной линии второго порядка, т. е. либо по эллипсу, либо по гиперболе, либо по параболе.

5. 8. Приведение к каноническому виду уравнения поверхности второго порядка с центром в начале координат

33. Далее мы будем заниматься упрощением общего уравнения поверхности второго порядка путем перехода к новым координатам. В целях простоты последующих выкладок введем удобную запись общего уравнения второй степени относительно трех переменных *x, y, z*. Ранее мы обозначали коэффициенты такого уравнения разными буквами; нам требовалось десять букв. В последующие формулы эти буквы войдут в разных сочетаниях и формулы будет трудно запоминать. Более целесообразной оказывается другая система обозначений, когда все коэффициенты обозначаются одной буквой с двумя числовыми значками снизу (их называют индексами). По этой системе общее уравнение второй степени относительно *x, y, z* пишется так:

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + \\ + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = 0. \quad (1)$$

Здесь применяется следующий принцип: в индексах коэффициента пишется 1, 2 или 3 столько раз, сколько раз встречается сомножителем при этом коэффициенте *x*, *y* или *z;* в качестве добавочного индекса берется 4. Например, коэффициент при x^2 помечается два раза единицей, так как $x^2 = x{\cdot}x$; коэффициент при *xz* помечается одной единицей и одной тройкой; коэффициент при *x* помечается одной единицей и одним добавочным индексом — четверкой, у свободного члена оба индекса добавочные (две четверки). Порядок индексов безразличен; таким образом, $a_{12} = a_{21}$, $a_{13} = a_{31}$ и т. д. Коэффициентами уравнения (1) называются числа $a_{11}, a_{22}, \ldots, a_{12}, \ldots, a_{44}$ (хотя, строго говоря, число a_{12}, например, есть только половина коэффициента при *xy).* Причина постановки множителя 2 при некоторых коэффициентах хорошо выясняется с помощью следующего тождества;

$$\begin{aligned}a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + \\ + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = \\ = (a_{11}x + a_{12}y + a_{13}z + a_{14})\,x + \\ + (a_{21}x + a_{22}y + a_{23}z + a_{24})\,y + \\ + (a_{31}x + a_{32}y + a_{33}z + a_{34})\,z + \\ + (a_{41}x + a_{42}y + a_{43}z + a_{44}).\end{aligned} \qquad (2)$$

Отсюда видно, что члены левой части с четвертого по девятый естественным образом составляются из двух одинаковых экземпляров каждый.

34. В этом параграфе мы рассмотрим неполное уравнение второй степени следующего вида:

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz = H. \qquad (3)$$

Особенностью его является отсутствие членов первой степени. Ввиду этого левая часть (3) не меняется при замене *x, y, z* на *—x, —y, —z;* следовательно, если точка *M(x; y; z)* лежит на поверхности (3), то точка *N(—x;—y; —z)* также лежит на этой поверхности. Иначе говоря, точки поверхности (3) расположены парами, симметрично относительно начала координат. Таким образом, если поверхность второго порядка задана уравнением вида (3), то: 1) она обладает центром (симметрии); 2) начало координат помещено в центр.

35. Левая часть уравнения (3)

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz \qquad (4)$$

представляет собой однородный многочлен второй степени (т. е. многочлен, состоящий только из членов второй степени). Такой многочлен называется *квадратичной формой* от трех переменных *x, y, z*. Мы займемся сейчас задачей о приведении квадратичной формы (4) к каноническому виду. Сущность этвй задачи в следующем: требуется повернуть систему координатных осей так, чтобы после приведения формы (4) к новым (прямоугольным) координатам исчезли все члены с произведениями новых текущих координат. Согласно п.5.6 дело сводится к тому, чтобы найти формулы

$$\left.\begin{aligned} x &= l_1x' + l_2y' + l_3z', \\ y &= m_1x' + m_2y' + m_3z', \\ z &= n_1x' + n_2y' + n_3z', \end{aligned}\right\} \qquad (5)$$

в силу которых имеет место тождество

$$\begin{aligned}a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz = \\ = \lambda_1x'^2 + \lambda_2y'^2 + \lambda_3z'^2.\end{aligned} \qquad (6)$$

Это значит, что после замены величин *x, y, z* по формулам (5) данная квадратичная форма должна принять вид, указанный в правой части (6) и называемый *каноническим видом.* Коэффициенты формул (5) надлежит подобрать так, чтобы соблюдались условия (10), (11) и (12а) п.5. 6.

Если квадратичная форма (4) будет приведена к каноническому виду, то одновременно с нею получит *канонический вид* и уравнение (3). Мы докажем, что каждую квадратичную форму (4) (а вместе с тем и каждое уравнение (3)) можно привести к каноническому виду.

36. Предположим сначала, что коэффициенты формул (5) уже найдены и тождество (6) достигнуто. Перепишем его следующим образом:

$$\begin{aligned}(a_{11}x+a_{12}y+a_{13}z)\,x+(a_{21}x+a_{22}y+a_{23}z)\,y&+\\+(a_{31}x+a_{32}y+a_{33}z)\,z=\lambda_1x'x'+\lambda_2y'y'+\lambda_3z'z'&. \qquad (7)\end{aligned}$$

Теперь каждую из скобок в левой части преобразуем по формулам (5):

$$\begin{aligned}a_{11}x+a_{12}y+a_{13}z=(a_{11}l_1+a_{12}m_1+a_{13}n_1)\,x'&+\\+(a_{11}l_2+a_{12}m_2+a_{13}n_2)\,y'&+\\+(a_{11}l_3+a_{12}m_3+a_{13}n_3)\,z'&; \qquad (8)\end{aligned}$$

$$\begin{aligned}a_{21}x+a_{22}y+a_{23}z=(a_{21}l_1+a_{22}m_1+a_{23}n_1)\,x'&+\\+(a_{21}l_2+a_{22}m_2+a_{23}n_2)\,y'&+\\+(a_{21}l_3+a_{22}m_3+a_{23}n_3)\,z'&; \qquad (9)\end{aligned}$$

$$\begin{aligned}a_{31}x+a_{32}y+a_{33}z=(a_{31}l_1+a_{32}m_1+a_{33}n_1)\,x'&+\\+(a_{31}l_2+a_{32}m_2+a_{33}n_2)\,y'&+\\+(a_{31}l_3+a_{32}m_3+a_{33}n_3)\,z'&. \qquad (10)\end{aligned}$$

Используем, далее, обратные формулы преобразования координат:

$$\begin{aligned}x'&=l_1x+m_1y+n_1z,\\ y'&=l_2x+m_2y+n_2z,\\ z'&=l_3x+m_3y+n_3z.\end{aligned}$$

Отсюда

$$\left.\begin{aligned}x'x'&=l_1xx'+m_1yx'+n_1zx',\\ y'y'&=l_2xy'+m_2yy'+n_2zy',\\ z'z'&=l_3xz'+m_3yz'+n_3zz'.\end{aligned}\right\} \qquad (11)$$

Если мы подставим выражения (8), (9) и (10) в левую часть (7), то получим слева девять различных членов; если подставим выражения (11) в правую часть (7), то получим справа девять аналогичных членов. Тождество (7) будет обеспечено, если коэффициенты подобных членов слева и справа окажутся одинаковыми; одновременно будет обеспечено и тождество (6). Таким образом, для обеспечения тождества (6) достаточно подобрать коэффициенты формул (5) и числа $\lambda_1, \lambda_2, \lambda_3$ так, чтобы

$$\left.\begin{aligned} a_{11}l_1 + a_{12}m_1 + a_{13}n_1 = \lambda_1 l_1, \\ a_{21}l_1 + a_{22}m_1 + a_{23}n_1 = \lambda_1 m_1, \\ a_{31}l_1 + a_{32}m_1 + a_{33}n_1 = \lambda_1 n_1; \end{aligned}\right\} \tag{12a}$$

$$\left.\begin{aligned} a_{11}l_2 + a_{12}m_2 + a_{13}n_2 = \lambda_2 l_2, \\ a_{21}l_2 + a_{22}m_2 + a_{23}n_2 = \lambda_2 m_2, \\ a_{31}l_2 + a_{32}m_2 + a_{33}n_2 = \lambda_2 n_2; \end{aligned}\right\} \tag{12b}$$

$$\left.\begin{aligned} a_{11}l_3 + a_{12}m_3 + a_{13}n_3 = \lambda_3 l_3, \\ a_{21}l_3 + a_{22}m_3 + a_{23}n_3 = \lambda_3 m_3, \\ a_{31}l_3 + a_{32}m_3 + a_{33}n_3 = \lambda_3 n_3. \end{aligned}\right\} \tag{12c}$$

Дело свелось к системе уравнений

$$\left.\begin{aligned} a_{11}l + a_{12}m + a_{13}n = \lambda l, \\ a_{21}l + a_{22}m + a_{23}n = \lambda m, \\ a_{31}l + a_{32}m + a_{33}n = \lambda n. \end{aligned}\right\} \tag{13}$$

Задача будет завершена, если мы найдем три решения $l_1, m_1, n_1, \lambda_1, l_2, m_2, n_2, \lambda_2$ и l_3, m_3, n_3, λ_3 системы (13), удовлетворяющие условиям (10) — (12a) п.5. 6.

37. Займемся системой (13). Перепишем ее следующим образом.

$$\left.\begin{aligned} (a_{11} - \lambda)\, l + a_{12}m + a_{13}n = 0, \\ a_{21}l + (a_{22} - \lambda)\, m + a_{23}n = 0, \\ a_{31}l + a_{32}m + (a_{33} - \lambda)\, n = 0. \end{aligned}\right\} \tag{14}$$

Отсюда

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0, \tag{15}$$

так как в противном случае мы не получим из системы (14) ничего, кроме l=0, m —0, n = 0.

Уравнение (15) есть уравнение третьей степени. Из него мы должны найти нужные нам значения $\lambda=\lambda_1$, $\lambda=\lambda_2$, $\lambda=\lambda_3$, а затем, решая систему (14) при найденных значениях λ, найти l_1, m_1, n_1 и т. д. Весьма важно, что уравнение (15) имеет только вещественные корни. Это утверждение будет доказано в следующем пункте.

38. Пусть $\lambda=\lambda_1$— какой-нибудь корень уравнения (15). Если мы подставим $\lambda=\lambda_1$ в систему (14), то эта система будет иметь ненулевое решение (так как ее определитель будет равен нулю). Обозначим ненулевое решение системы (14), получаемое при $\lambda=\lambda_1$ через l_1, m_1, n_1. Если $\lambda=\lambda_2$— какой-нибудь еще корень уравнения (15), то аналогичным образом определится соответствующее ему ненулевое решение системы (14), которое мы обозначим через l_2, m_2, n_2.

Лемма. *Если, корни* λ_1, λ_2 *уравнения* (15) *различны,* $\lambda_1\neq\lambda_2$, *то соответствующие им решения системы* (14) *удовлетворяют условию перпендикулярности:*

$$l_1l_2+m_1m_2+n_1n_2=0.$$

Доказательство. Прежде всего заметим, что числа l_1, m_1, n_1 удовтетворяют равенствам (12a), поскольку они найдены из (14); аналогично, l_2, m_2, n_2 удовлетворяют (12b). Умножим первое, второе и третье равенства (12a) соответственно на l_2, m_2, n_2 и после этого сложим их почленно; мы получим справа $\lambda_1(l_1l_2+m_1m_2+n_1n_\lambda)$. Точно так же, умножая равенства (12b) на l_1, m_1, n_1 и складывая, получим справа $\lambda_2(l_1l_2+m_1m_2+n_1n_2)$. Слева и в том, и в другом случае будет получаться одна и та же величина, так как $a_{12}=a_{21}$, $a_{13}=a_{31}$ и $a_{23}=a_{32}$. Поэтому

$$\lambda_1(l_1l_2+m_1m_2+n_1n_2)=\lambda_2(l_1l_2+m_1m_2+n_1n_2).$$

Отсюда

$$(\lambda_2-\lambda_1)(l_1l_2+m_1m_2+n_1n_2)=0,$$

а так как $\lambda_2\neq\lambda_1$, то

$$l_1l_2+m_1m_2+n_1n_2=0,$$

что и требовалось.

Теорема. *Все корни уравнения* (15) *вещественны.*

Доказательство. Докажем теорему от противного. Допустим, что уравнение (15) имеет комплексный корень $\lambda_1=\alpha+\beta i;\ \beta\neq0.$ Тогда, так как коэффициенты уравнения (15) вещественны, оно должно иметь еще один комплексный корень $\lambda_2=\alpha-\beta i$ (сопряженный первому). Корню λ_1 соответствует некоторое ненулевое решение системы (14). Поскольку λ_1 — комплексное число, то при подстановке $\lambda=\lambda_1=\alpha+\beta i$ в систему (14) некоторые коэффициенты этой системы окажутся комплексными числами; поэтому ненулевое решение системы (14), соответствующее корню $\lambda=\lambda_2$, будет состоять из комплексных чисел:

$$l_1 = l + il', \quad m_1 = m + im', \quad n_1 = n + in'.$$

Так как число $\lambda_2 = \alpha - \beta i$ сопряжено числу $\lambda_1 = \alpha + \beta i$, то корню $\lambda = \lambda_2$ будет соответствовать решение l_2, m_2, n_2, сопряженное решению l_1, m_1, n_1, т. е.

$$l_2 = l - il', \quad m_2 = m - im', \quad n_2 = n - in'.$$

Поскольку $\beta \neq 0$, корни $\lambda_1 = \alpha + i\beta$ и $\lambda_2 = \alpha - i\beta$ различны; но тогда на основании леммы

$$l_1 l_2 + m_1 m_2 + n_1 n_2 = 0.$$

Отсюда

$$l^2 + l'^2 + m^2 + m'^2 + n^2 + n'^2 = 0.$$

Следовательно,

$$l = 0, \; l' = 0, \; m = 0, \; m' = 0, \; n = 0, \; n' = 0.$$

Значит, $l_1 = 0, \; m_1 = 0, \; n_1 = 0$. Но это противоречит тому, что l_1, m_1, n_1 является ненулевым решением системы (14). Ввиду полученного противоречия мы должны отвергнуть допущение, что существует комплексный корень. Теорема доказана.

Замечание. Доказательство этой теоремы основано на предыдущей лемме. Доказательство леммы существенно опиралось на условие симметрии элементов определителя (15) относительно его главной диагонали: $a_{12} = a_{21}, \; a_{13} = a_{31}, \; a_{23} = a_{32}$. Без этого условия утверждения леммы и теоремы неверны. Например, уравнение

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = 0$$

имеет корни $\lambda_1 = 1, \; \lambda_2 = 1 + i, \; \lambda_3 = 1 - i$, среди них — два комплексных. В данном случае теорема неприменима, так как

$$a_{23} = 1, \; a_{32} = -1, \; a_{23} \neq a_{32}.$$

39. Уравнение (15) называется *характеристическим уравнением* квадратичной формы (4). Корни (всегда вещественные) уравнения (15) называются $\lambda_1, \lambda_2, \lambda_3$ *характеристическими числами* формы (4); они же получаются в качестве коэффициентов после приведения формы к каноническому виду.

Если мы подставим $\lambda = \lambda_1$ в систему (14), то система (14) будет иметь ненулевое решение l, m, n. Направление вектора $\{l; \; m; \; n\}$ называется *главным направлением* данной квадратичной формы, соответствующим характеристическому числу λ_1. Ясно, что вектор $\{l_1;\ m_1;\ n_1\}$, где $l_1=\mu l,\ m_1=\mu m,\ n_1=\mu n$ (μ — какое-нибудь число $\neq 0$), также имеет главное направление. Возьмем

$$\mu=\frac{\pm 1}{\sqrt{l^2+m^2+n^2}}.$$

Тогда

$$l_1^2+m_1^2+n_1^2=1.$$

Такое решение l_1, m_1, n_1 системы (14) будем называть нормированным. Вектор $\{l_1;\ m_1;\ n_1\}$ является единичным вектором главного направления, которое соответствует числу λ_1. Аналогично, решая систему при $\lambda=\lambda_2$ и при $\lambda=\lambda_3$ и нормируя решения, получим единичные векторы $\{l_2;\ m_2;\ n_2\}$ и $\{l_3;\ m_3;\ n_3\}$ главных направлений, которые соответствуют характеристическим числам $\lambda=\lambda_2$ и $\lambda=\lambda_3$. Теперь мы имеем почти весь необходимый материал, чтобы доказать, что каждая квадратичная форма (4) может быть приведена к каноническому виду при помощи преобразования прямоугольных координат.

40. Первый случай. Характеристические числа $\lambda_1,\ \lambda_2,\ \lambda_3$ все разные. Тогда каждому из них соответствует главное направление, причем направления, соответствующие любым двум характеристическим числам, перпендикулярны между собой (см. лемму п° 38). Единичные векторы этих направлений

$$i'=\{l_1;\ m_1;\ n_1\},$$
$$j'=\{l_2;\ m_2;\ n_2\},$$
$$k'=\{l_3;\ m_3;\ n_3\}$$

находятся из системы (14) согласно указаниям п° 39. Если мы примем эти векторы в качестве базисных векторов новой системы координат, то после перехода к новым координатам квадратичная форма получит канонический вид: $\lambda_1x'^2+\lambda_2y'^2+\lambda_3z'^2$ (это ясно, так как будут соблюдены равенства (12) и, следовательно, тождество (6)).

Переход к новым координатам дается формулами (5). Условия (10) п.5.6 будут соблюдены вследствие нормировки решений системы (14); условия (11) п.5.6 соблюдены, поскольку $i',\ j',\ k'$ попарно перпендикулярны; для обеспечения условия (12а) п.5.6 нужно надлежащим образом выбрать знаки при нормировке решений системы (14) (если знаки выбраны как-нибудь и если окажется, что условие (12а) п.5.6 не соблюдено, то достаточно изменить знаки, например, чисел l_1, m_1, n_1).

41. Второй случай. Среди характеристических чисел имеются два одинаковых, а третье — отлично от них (иначе говоря, характеристическое уравнение (15) имеет один двукратный корень). Будем считать, что $\lambda_1 = \lambda_2 \neq \lambda_3$. Имеет место следующее предложение: *если в систему* (14) *подставить двукратный корень* $\lambda = \lambda_1 = \lambda_2$ *уравнения* (15), *то система* (14) *сведется к одному существенному уравнению* (остальные уравнения системы будут его следствиями). Доказательство этого утверждения мы проведем позднее, в п° 44, а пока примем его на веру. Допустим для определенности записи, что при $\lambda = \lambda_1 = \lambda_2$ существенным уравнением системы (14) будет

$$(a_{11} - \lambda)\, l + a_{12}m + a_{13}n = 0. \tag{16}$$

Когда мы говорим, что это уравнение существенно, то имеем в виду, что оно — не тождество, т. е. что среди чисел $a_{11} = \lambda$, a_{12}, a_{13} имеется не равное нулю; это означает, что вектор $\{a_{11} - \lambda;\ a_{12},\ a_{13}\}$ является ненулевым. Всякое ненулевое решение l, m, n уравнения (16) дает вектор $\{l;\ m;\ n\}$ который определяет главное направление, соответствующее характеристическому числу $\lambda = \lambda_1 = \lambda_2$. Учтем, что уравнение (16) выражает условие перпендикулярности векторов $\{l; m; n\}$ и $\{a_{11} - \lambda;\ a_{12};\ a_{13}\}$. Отсюда заключаем: характеристическому числу $\lambda = \lambda_1 = \lambda_2$ соответствует бесконечно много главных направлений, именно, любое направление, перпендикулярное вектору $\{a_{11} - \lambda;\ a_{12};\ a_{13}\}$, есть главное направление, соответствующее числу $\lambda = \lambda_1 = \lambda_2$. Отсюда в свою очередь следует, что сам вектор $\{a_{11} - \lambda;\ a_{12};\ a_{13}\}$ имеет главное направление, отвечающее числу $\lambda = \lambda_3$.

Теперь ясно, что и в случае $\lambda_1 = \lambda_2 \neq \lambda_3$ мы можем найти три единичных вектора, которые направлены по главным направлениям данной формы и попарно перпендикулярны друг другу. Нужно действовать следующим образом. Подставив в систему (14) число $\lambda = \lambda_1 = \lambda_2$, найти любые два нормированные решения l_1, m_1, n_1 и l_2, m_2, n_2 этой системы, которые взаимно удовлетворяют условию перпендикулярности $(l_1 l_2 + m_1 m_2 + n_1 n_2 = 0)$; тем самым мы получим два единичных вектора:

$$i' = \{l_1;\ m_1;\ n_1\}, \quad j' = \{l_2;\ m_2;\ n_2\}.$$

Решая систему (14) при $\lambda = \lambda_3$ и нормируя решение, найдем третий единичный вектор:

$$k' = \{l_3;\ m_3;\ n_3\}.$$

Если мы примем i', j', k' в качестве базиса новой координатной системы, то после перехода к новым координатам данная квадратичная форма получит канонический вид:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 = \lambda (x'^2 + y'^2) + \lambda_3 z'^2 \quad (\text{где} \quad \lambda = \lambda_1 = \lambda_2).$$

Указание. Практически удобно сначала найти вектор $\boldsymbol{i'}$ (т. е. какое-нибудь одно нормированное решение l_1, m_1, n_1 системы (14) при $\lambda = \lambda_1$), затем вектор $\boldsymbol{k'}$; после этого вектор $\boldsymbol{j'}$ может быть найден, как векторное произведение $\boldsymbol{k'}$ на $\boldsymbol{i'}$. Заметим еще, что для определения $\boldsymbol{k'}$ нет необходимости решать систему (14), поскольку мы уже знаем, что направление $\boldsymbol{k'}$, т. е. третье главное направление, дается вектором $\{a_{11} - \lambda;\ a_{12};\ a_{13}\}$ при $\lambda = \lambda_1 = \lambda_2$.

42. Третий случай. Все характеристические числа одинаковы: $\lambda_1 = \lambda_2 = \lambda_3$, т. е. характеристическое уравнение (15) имеет трехкратный корень.

Имеет место следующее предложение: *если в систему* (14) *подставить трехкратный кореньураенения* (15), *то все уравнения системы* $\lambda = \lambda_1 = \lambda_2 = \lambda_3$ (14) *обратятся в тождества.* Таким образом, если характеристические числа квадратичной формы совпадают, то любое направление является для такой формы главным. Доказательство этого утверждения мы проведем в п°44; пока примем его на веру.

Поскольку при $\lambda = \lambda_1 = \lambda_2 = \lambda_3$ система (14) сводится к тождествам, т. е. все коэффициенты этой системы обращаются в нуль, получаем:

$$a_{11} - \lambda = 0,\ a_{22} - \lambda = 0,\ a_{33} - \lambda = 0,\ a_{12} = a_{13} = a_{23} = 0.$$

Следовательно, данная форма уже имеет канонический вид: $\lambda x^2 + \lambda y^2 + \lambda z^2 = \lambda (x^2 + y^2 + z^2)$ и ничего с ней делать не надо. Однако мы можем как угодно повернуть систему осей, и форма при этом сохранит канонический вид.

43. Общее заключение. *Каждую квадратичную форму можно привести к каноническому виду при помощи преобразования прямоугольных координат.* Чтобы привести данную квадратичную форму к каноническому виду, нужно решить уравнение (15) и найти характеристические числа λ_1, λ_2, λ_3; они и будут коэффициентами в каноническом виде формы. Координатные оси следует направить по главным направлениям формы. Если оси абсцисс, ординат и аппликат направлены соответственно по первому, второму и третьему главным направлениям, то λ_1 будет коэффициентом при квадрате абсциссы, λ_2—при квадрате ординаты, λ_3—при квадрате аппликаты. Ясно, что тем самым дано также правило приведения к каноническому виду уравнения поверхности второго порядка (1), поскольку его левая часть есть квадратичная форма.

Пример. Привести к каноническому виду уравнение поверхности второго порядка:

$$7x^2+6y^2+5z^2-4xy-4yz-18=0. \tag{17}$$

Решение. Прежде всего составим характеристическое уравнение:

$$\begin{vmatrix} 7-\lambda & -2 & 0 \\ -2 & 6-\lambda & -2 \\ 0 & -2 & 5-\lambda \end{vmatrix}=0,$$

или $\lambda^3-18\lambda^2+99\lambda-162=0.$ Корни этого уравнения: $\lambda_1=3$, $\lambda_2=6$, $\lambda_3=9$. Тем самым мы уже знаем каноническое уравнение данной поверхности: $3x'^2+6y'^2+9z'^2-18=0$, или

$$\frac{x'^2}{6}+\frac{y'^2}{3}+\frac{z'^2}{2}=1.$$

Данная поверхность является эллипсоидом с полуосями $a=\sqrt{6},\ b=\sqrt{3},\ c=\sqrt{2}.$

Если нас, кроме того, интересует расположение поверхности, то мы должны найти преобразование координат, приводящее уравнение (17) к каноническому виду; иначе говоря, придется искать главные направления поверхности.

Составим применительно к нашей задаче систему уравнений (14):

$$\left.\begin{aligned} (7-\lambda)\,l \quad -2m \qquad\qquad &=0, \\ -2l+(6-\lambda)\quad m \quad -2\ n&=0, \\ -2m+(5-\lambda)\,n&=0. \end{aligned}\right\} \tag{18}$$

Полагая здесь $\lambda=\lambda_1=3$, получим

$$\begin{aligned} 4l-2m \qquad &=0, \\ -2l+3m-2n&=0, \\ -2m+2n&=0. \end{aligned}$$

В качестве ненулевого решения этой системы можно взять, например, $l=1$, $m=2$, $n=2$; нормируя это решение, получим единичный вектор первого главного направления:

$$i' = \{l_1;\ m_1;\ n_1\} = \left\{\frac{1}{3};\ \frac{2}{3};\ \frac{2}{3}\right\}.$$

Точно так же, полагаяв системе (18) $\lambda = \lambda_2 = 6$ и $\lambda = \lambda_3 = 9$, найдем единичные векторы двух других главных направлений:

$$j' = \{l_2;\ m_2;\ n_2\} = \left\{\frac{2}{3};\ \frac{1}{3};\ -\frac{2}{3}\right\},$$

$$k' = \{l_3;\ m_3;\ n_3\} = \left\{-\frac{2}{3};\ \frac{2}{3};\ -\frac{1}{3}\right\}.$$

Положение новых осей относительно старой системы теперь нам известно, вместе с тем известно и расположение поверхности. Формулы преобразования координат найдем согласно равенствам (5):

$$x = \frac{1}{3}x' + \frac{2}{3}y' - \frac{2}{3}z',$$

$$y = \frac{2}{3}x' + \frac{1}{3}y' + \frac{2}{3}z',$$

$$z = \frac{2}{3}x' - \frac{2}{3}y' - \frac{1}{3}z'.$$

Заметим, что определитель, составленный из координат векторов ***i, j, k,*** равен +1; таким образом, условие (12а) п.5.6 соблюдено. Это значит, что мы перешли к новым координатам, сохранив ориентацию прежних осей.

Пример. Привести к каноническому виду уравнение поверхности второго порядка:

$$x^2 - 2y^2 + z^2 + 4xy - 8xz - 4yz + 6 = 0 \qquad (19)$$

Решение. Составим характеристическое уравнение:

$$\begin{vmatrix} 1-\lambda & 2 & -4 \\ 2 & -2-\lambda & -2 \\ -4 & -2 & 1-\lambda \end{vmatrix} = 0,$$

или $\lambda^3 - 27\lambda - 54 = 0$. Корни этого уравнения:

λ_1=-3, λ_2=-3, λ_3=6.

Каноническое уравнение поверхности можем написать сразу: $-3x'^2 - 3y'^2 + 6z^2 + 6 = 0$ или

$$\frac{x'^2}{2} + \frac{y'^2}{2} - \frac{z'^2}{1} = 1.$$

Данная поверхность является однополостным гиперболоидом вращения с полуосями $a=\sqrt{2},\ b=\sqrt{2},\ c=1.$

Найдем теперь главные направления данной поверхности и формулы преобразования координат, приводящие данное уравнение к каноническому виду. Так как характеристическое уравнение имеет двукратный корень, то нужно действовать согласно указаниям п°41. Составим применительно к нашей задаче систему уравнений (14):

$$\begin{aligned}(1-\lambda)l \quad &+2m &-4n=0,\\ 2l+(-2-\lambda)m &&-2n=0,\\ -4l \quad &-2m &+(1-\lambda)n=0.\end{aligned}$$

Подставим сюда двукратный корень характеристического уравнения $\lambda=\lambda_1=\lambda_2=-3;$ мы получим

$$\begin{aligned}4l+2m-4n&=0,\\ 2l+\ m-2n&=0,\\ -4l-2m+4n&=0.\end{aligned}$$

Система свелась к одному существенному уравнению:

$$2l+m-2n=0 \qquad (20)$$

(два других ему пропорциональны; см. п°41). Беря, например, решение $l=1$, $m=2$, $n=2$ уравнения (20), получим вектор {1; 2; 2}, который определяет одно из бесчисленного множества главных направлений, соответствующих числу $\lambda=\lambda_1=\lambda_2=-3.$ С другой стороны, согласно п° 41 вектор {2; 1; —2}, определенный коэффициентами уравнения (20), дает третье главное направление (отвечающее числу ^ $\lambda=\lambda_3=6$). Умножая векторно {2; 1; —2}на {1; 2; 2}, получим вектор {6; —6; 3}, который также дает главное направление, отвечающее числу $\lambda=\lambda_1=\lambda_2=-3.$ (но отличное от ранее найденного и перпендикулярное к нему). Вместо последнего вектора удобнее взять {2; —2; 1}. Нормируя найденные векторы и располагая их в надлежащем порядке, получим

$$i'=\{l_1;\ m_1;\ n_1\}=\left\{\frac{1}{3};\ \frac{2}{3};\ \frac{2}{3}\right\},$$

$$j'=\{l_2;\ m_2;\ n_2\}=\left\{\frac{2}{3};\ -\frac{2}{3};\ \frac{1}{3}\right\},$$

$$k'=\{l_3;\ m_3;\ n_3\}=\left\{\frac{2}{3};\ \frac{1}{3};\ -\frac{2}{3}\right\}.$$

Отсюда имеем формулы искомого преобразования координат:

$$x=\frac{1}{3}x'+\frac{2}{3}y'+\frac{2}{3}z',$$
$$y=\frac{2}{3}x'-\frac{2}{3}y'+\frac{1}{3}z',$$
$$z=\frac{2}{3}x'+\frac{1}{3}y'-\frac{2}{3}z'.$$

44. Теперь мыдокажем предложения, которые в п° 41 и 42 были приняты на веру. Предварительно установим следующие леммы.

Лемма 1. *Пусть соблюдены равенства:*

$$\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}=0, \qquad (21)$$

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}+\begin{vmatrix} b_{11} & b_{13} \\ b_{31} & b_{33} \end{vmatrix}+\begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix}=0. \qquad (22)$$

Тогда, если определитель (21) *симметричен относительно главной диагонали, т. е. если* $b_{12}=b_{21}$, $b_{13}=b_{31}$, $b_{23}=b_{32}$, *то все строки определителя* (21) *пропорциональны некоторой одной его строке.*
Доказательство. Будем доказывать лемму от противного. Предположим, что определитель (21) имеет две непропорциональные строки; для определенности будем считать, что непропорциональны его первые две строки. Отсюда получим противоречие с условиями леммы.

Рассмотрим векторы $\{b_{11};\ b_{12};\ b_{13}\}$, $\{b_{21};\ b_{22},\ b_{23}\}$, $\{b_{31};\ b_{32};\ b_{33}\}$. Наше предположение означает, что из этих векторов первые два неколлинеарны. Но согласно равенству (21) все эти векторы компланарны; в таком случае третий вектор может быть разложен по двум первым:

$$\{b_{31};\ b_{32};\ b_{33}\}=\lambda\{b_{11};\ b_{12};\ b_{13}\}+\mu\{b_{21};\ b_{22};\ b_{23}\}.$$

Следовательно,

$$\left.\begin{aligned} b_{31}&=\lambda b_{11}+\mu b_{21},\\ b_{32}&=\lambda b_{12}+\mu b_{22},\\ b_{33}&=\lambda b_{13}+\mu b_{23}. \end{aligned}\right\} \qquad (23)$$

Пользуясь условием симметрии данного определителя и равенствами (23), имеем

$$\begin{vmatrix} b_{11} & b_{13} \\ b_{31} & b_{33} \end{vmatrix} = \begin{vmatrix} b_{11} & b_{13} \\ \lambda b_{11} + \mu b_{21} & \lambda b_{13} + \mu b_{23} \end{vmatrix} =$$
$$= \mu \begin{vmatrix} b_{11} & b_{13} \\ b_{21} & b_{23} \end{vmatrix} = \mu \begin{vmatrix} b_{11} & \lambda b_{11} + \mu b_{21} \\ b_{21} & \lambda b_{12} + \mu b_{22} \end{vmatrix} = \mu^2 \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}.$$

Аналогично

$$\begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} = \lambda^2 \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}.$$

Заменяя в левой частиравенства (22) второй и третий определители найденными выражениями, получим

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} (1 + \mu^2 + \lambda^2) = 0.$$

Отсюда

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = 0.$$

Далее, имеем

$$\begin{vmatrix} b_{11} & b_{13} \\ b_{21} & b_{23} \end{vmatrix} = \begin{vmatrix} b_{11} & \lambda b_{11} + \mu b_{21} \\ b_{21} & \lambda b_{12} + \mu b_{22} \end{vmatrix} = \mu \begin{vmatrix} b_{11} & b_{21} \\ b_{21} & b_{22} \end{vmatrix} = 0.$$
$$\begin{vmatrix} b_{12} & b_{13} \\ b_{22} & b_{23} \end{vmatrix} = \begin{vmatrix} b_{12} & \lambda b_{11} + \mu b_{21} \\ b_{22} & \lambda b_{12} + \mu b_{22} \end{vmatrix} = \lambda \begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{12} \end{vmatrix} = 0.$$

Итак, мы получим три равенства:

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = 0, \quad \begin{vmatrix} b_{11} & b_{13} \\ b_{31} & b_{23} \end{vmatrix} = 0, \quad \begin{vmatrix} b_{12} & b_{13} \\ b_{22} & b_{23} \end{vmatrix} = 0.$$

Но эти равенства означают пропорциональность первых двух строк определителя (21). Таким образом, при условиях леммы нельзя допустить, что первые две строчки непропорциональны. Точно так же мы получили бы противоречие, допустив непропорциональность каких-нибудь других двух строк. Лемма докатана.

Лемма 2. *Если соблюдены условия первой леммы и, кроме того,*

$$b_{11} + b_{22} + b_{33} = 0, \qquad (24)$$

то все элементы определителя (21) *равны нулю.*

Доказательство. Так как соблюдены условия первой леммы, то мы можем ее применить; согласно лемме, все строки определителя (21) пропорциональны и, следовательно, все миноры второго порядка этого определителя равны нулю. В частности,

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = 0, \quad \begin{vmatrix} b_{11} & b_{13} \\ b_{31} & b_{33} \end{vmatrix} = 0, \quad \begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} = 0.$$

Отсюда

$$b_{11}b_{22} = b_{12}^2, \quad b_{11}b_{33} = b_{13}^2, \quad b_{22}b_{33} = b_{23}^2. \tag{25}$$

Таким образом,

$$b_{11}b_{22} \geqslant 0, \quad b_{11}b_{33} \geqslant 0, \quad b_{22}b_{33} \geqslant 0;$$

поэтому числа b_{11}, b_{22}, b_{33} не могут иметь разных знаков. Но тогда из (24) следует, что b_{11} = 0, b_{22}= 0, b_{33} =0. После этого из (25) имеем: b_{12} = 0, b_{13} = 0, b_{23} =0. Лемма доказана.

Нам потребуется еще некоторые сведения из алгебры многочленов. Пусть дан многочлен третьей степени:

$$f(\lambda) = \lambda^3 + a\lambda^2 + b\lambda + c. \tag{26}$$

Известно, что он может быть записан следующим образом:

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3), \tag{27}$$

где λ_1, λ_2, λ_3 — определенные числа, называемые корнями данного многочлена, если $\lambda_1 = \lambda_2 \neq \lambda_3$, то $\lambda_1 = \lambda_2$ есть двукратный корень многочлена; в этом случае

$$f(\lambda) = (\lambda - \lambda_1)^2 (\lambda - \lambda_3). \tag{28}$$

Если λ_1=λ_2=λ_3, то имеем трехкратный корень; в этом случае

$$f(\lambda) = (\lambda - \lambda_1)^3. \tag{29}$$

Справедливы следующие утверждения:

1) Если λ_1— вообще какой-нибудь корень данного многочлена, то, полагая в выражении (26) λ= λ_1, мы должны получить

$$f(\lambda_1) = 0.$$

Это сразу усматривается из (27).

2) Если λ_1—двукратный корень, то при λ= λ_1 бращается в нуль не только саммногочлен, но и его первая производная:

$$f(\lambda_1) = 0, \quad f'(\lambda_1) = 0,$$

однако $f''(\lambda_1) \neq 0$. Чтобы доказать это утверждение, достаточно продифференцировать (28):

$$f'(\lambda) = 2(\lambda - \lambda_1)(\lambda - \lambda_3) + (\lambda - \lambda_1)^2.$$

Подставляя сюда λ= λ_1, видим, что $f'(\lambda_1)$ =0. Поскольку

$$f''(\lambda) = 2(\lambda - \lambda_3) + 4(\lambda - \lambda_1) \quad \text{и} \quad \lambda_1 \neq \lambda_3, \quad \text{то} \quad f''(\lambda_1) \neq 0.$$

3) Если λ_1—трехкратный корень, то

$$f(\lambda_1)=0, \quad f'(\lambda_1)=0, \quad f''(\lambda_1)=0$$

(однако $f'''(\lambda_1)\neq 0$. Для доказательства достаточно продифференцировать (29):

$$f'(\lambda)=3(\lambda-\lambda_1)^2, \quad f''(\lambda)=6(\lambda-\lambda_1).$$

Полагая $\lambda=\lambda_1$, находим: $f'(\lambda_1)=0, \quad f''(\lambda_1)=0 \quad (f'''(\lambda)=6\neq 0)$. Эти признаки кратности корня будут сейчас использованы.

Вернемся к предложениям, которые были высказаны в п° 41, 42 и остались недоказанными.

Обозначим через $f(\lambda)$ многочлен третьей степени, который стоит в левой части характеристического уравнения данной квадратичной формы:

$$f(\lambda)=\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}. \tag{30}$$

Найдем производную многочлена $f(\lambda)$:

$$f'(\lambda)=\begin{vmatrix} -1 & a_{12} & a_{13} \\ 0 & a_{22}-\lambda & a_{23} \\ 0 & a_{32} & a_{33}-\lambda \end{vmatrix}+\begin{vmatrix} a_{11}-\lambda & 0 & a_{13} \\ a_{21} & -1 & a_{23} \\ a_{31} & 0 & a_{33}-\lambda \end{vmatrix}+\begin{vmatrix} a_{11}-\lambda & a_{12} & 0 \\ a_{21} & a_{22}-\lambda & 0 \\ a_{31} & a_{32} & -1 \end{vmatrix}.$$

Раскрывая эти определители, получим

$$f'(\lambda)=-\left(\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix}+\begin{vmatrix} a_{11}-\lambda & a_{13} \\ a_{31} & a_{33}-\lambda \end{vmatrix}+\begin{vmatrix} a_{22}-\lambda & a_{23} \\ a_{32} & a_{33}-\lambda \end{vmatrix}\right).$$

Отсюда

$$f''(\lambda)=2\,[(a_{11}-\lambda)+(a_{22}-\lambda)+(a_{33}-\lambda)].$$

Рассмотрим теперь возможные случаи кратности корня характеристического уравнения данной формы, о которых шла речь в п° 41, 42

1) Пусть λ_1—двукратный корень характеристического уравнения. Тогда

$$f(\lambda_1)=0, \quad f'(\lambda_1)=0.$$

Таким образом, при $\lambda=\lambda_1$ для определителя (30) соблюдены условия первой леммы этого п° (считая $b_{11}=a_{11}-\lambda, \quad b_{12}=a_{12}$ и т. д). Применяя лемму, заключаем, что все строки определителя (30) пропорциональны некоторой одной строке; при этом найдется хотя бы одна строка, не состоящая сплошь из нулей (так как $f''(\lambda_1)\neq 0$, то хотя бы одно из чисел

$$a_{11}-\lambda, \quad a_{22}-\lambda, \quad a_{33}-\lambda \quad (\lambda=\lambda_1)$$

отлично от нуля). Учтем, что строчки определителя (30) составлены из коэффициентов системы (14). Следовательно, при $\lambda = \lambda_1$ в системе (14) имеется одно существенное уравнение, а другие ему пропорциональны и, значит, являются его следствиями. Именно это и утверждалось в п° 41.

2) Пусть $\lambda = \lambda_1$ — трехкратный корень характеристического і/равнения. Тогда

$$f(\lambda_1)=0, \quad f'(\lambda_1)=0, \quad f''(\lambda_1)=0.$$

Таким образом, при $\lambda = \lambda_1$ для определителя (30) соблюдены условия второй леммы этого пункта. Применяя лемму, заключаем, что

$$a_{11}-\lambda_1=0, \quad a_{22}-\lambda_1=0, \quad a_{33}-\lambda_1=0, \quad a_{12}=0, \quad a_{13}=0, \quad a_{23}=0.$$

Следовательно, при $\lambda = \lambda_1$ все коэффициенты системы (14) обращаются в нуль и система (14) сводится к тождествам. Тем самым утверждение п° 42 доказано.

5.9. Инварианты и классификация квадратичных форм от трех аргументов

45. В предыдущем параграфе мы указали определенный способ приведения данной квадратичной формы от трех аргументов к каноническому виду. Заранее не ясно, не может ли случиться, что с помощью какого-то другого преобразования прямоугольных координат данная квадратичная форма приведется к другому каноническому виду? Покажем, что этого не может быть. Иначе говоря, *каждая квадратичная форма от трех аргументов имеет единственный канонический вид* (не считая возможности поменять ролями абсциссу, ординату и аппликату). Следует оговорить, что единственность квадратичной формы имеет место при условии, что употребляются только прямоугольные координатные системы с одним и тем же масштабом.

46. Перейдем к доказательству нашего утверждения. Обозначим буквой Ф величину данной квадратичной формы:

$$\Phi = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz. \qquad (1)$$

Согласно равенству (1) каждой точке *M(x; y; z)* сопоставляется число Ф; оно называется значением формы в точке *M*. Приведем данную форму к каноническому виду, следуя указаниям п.5.8, Тогда значение Ф в точке *M* будет выражено равенством:

$$\Phi = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2, \qquad (2)$$

где x', y', z' — новые координаты точки — *M*, λ_1, λ_2, λ_3 характеристические числа формы Ф.

Предположим, что $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$; в таком случае из (2) следует

$$\lambda_1(x'^2 + y'^2 + z'^2) \leqslant \Phi \leqslant \lambda_3(x'^2 + y'^2 + z'^2). \quad (3)$$

Поскольку $x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2$, то такие же соотношения имеем и в старых координатах:

$$\lambda_1(x^2 + y^2 + z^2) \leqslant \Phi \leqslant \lambda_3(x^2 + y^2 + z^2). \quad (4)$$

Ограничим возможные положения точки *М,* требуя, чтобы она находилась на единичной сфере:

$$x^2 + y^2 + z^2 = 1.$$

При этом условии все возможные значения Ф (отвечающие любым положениям точки *М* на единичной сфере) заключены между λ_1 и λ_3, т. е.

$$\lambda_1 \leqslant \Phi \leqslant \lambda_3. \quad (5)$$

Эти границы являются точными. Именно если $x' = \pm 1$ $y' = 0$, $z' = 0$, то $\Phi = \lambda_1$; если $x' = 0$, $y' = 0$, $z' = \pm 1$ то $\Phi = \lambda_3$. Из этих рассуждений следует, что если данная форма будет приведена к каноническому виду каким-нибудь другим способом (в других координатах), то наименьший и наибольший из коэффициентов этого канонического вида снова окажутся числами λ_1 и λ_3 (в противном случае данная форма на единичной сфере имела бы другие границы изменения, что невозможно). Покажем, что данная форма не может приводиться к различным каноническим видам, которые отличаются средним по величине коэффициентом. Допустим, что в каких-то координатах данная форма получила канонический вид:

$$\Phi = \lambda_1 x'^2 + \mu y'^2 + \lambda_3 z'^2,$$

где $\lambda_1 < \mu < \lambda_3$. Тогда на единичной сфере всюду, где $x' \neq 0$ или $y' \neq 0$, будет

$$\Phi = \lambda_1 x'^2 + \mu y'^2 + \lambda_3 z'^2 < \lambda_3(x'^2 + y'^2 + z'^2) = \lambda_3,$$

т. е.

$$\Phi < \lambda_3.$$

Только в двух точках единичной сферы: $(x' = 0,\ y' = 0,\ z' = +1)$ и $(x' = 0,\ y' = 0,\ z' = -1)$ имеем;

$$\Phi = \lambda_3.$$

Таким образом, на единичной сфере есть единственная пара диаметрально противоположных точек, где Ф достигает наибольшего значения. Именно через эти точки и проходит ось аппликат той координатной системы, в которой форма имеет канонический вид.

Так же доказывается, что на единичной сфере есть единственная пара диаметрально противоположных точек, где Ф достигает наименьшего значения: Ф = λ_1. Через эти точки проходит ось абсцисс координатной системы, в которой форма имеет канонический вид. Мы видим, что положение двух осей этой координатной системы определено единственным образом; следовательно, и вся система осей однозначно определена (не считая возможности менять на осях положительные и отрицательные направления). Поскольку в случае

$$\lambda_1 < \mu < \lambda_3$$

не может быть другой координатной системы, в которой форма $\lambda_1 x'^2 + \mu y'^2 + \lambda_3 z'^2$ также имеет канонический вид, не может быть и другого канонического вида. В частности, имеем: $\mu = \lambda_2$.

Остается рассмотреть случаи $\lambda_1 = \mu < \lambda_3$ и $\lambda_1 < \mu = \lambda_3$.

Если

$$\lambda_1 = \mu < \lambda_3,$$

то на единичной сфере максимум Ф = λ_3 имеет место, как и раньше, в единственной паре точек (0, 0, ±1), а минимум Ф= λ_1 достигается во всех точках, где $z' = 0$ (т. е. на всем экваторе). Поэтому теперь только ось аппликат имеет определенное положение. Если вращать систему осей вокруг неподвижной оси аппликат, то будет $\mu = \lambda_1$. Поскольку μ, не может меняться, имеем и в этом случае $\mu=\lambda_2$. Предположение рассматривается $\lambda_1<\mu=\lambda_3$ аналогично. Итак, данная квадратичная форма приводится к единственному каноническому виду: $\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$ (не считая возможности переименовать координаты).

47. Пусть дана произвольная квадратичная форма

$$\Phi = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz$$

и пусть совершается преобразование координат:

$$\begin{aligned} x &= l_1 x' + l_2 y' + l_3 z', \\ y &= m_1 x' + m_2 y' + m_3 z', \\ z &= n_1 x' + n_2 y' + n_3 z' \end{aligned}$$

при любом повороте системы осей.

Заменяя *x, y, z* в выражении Ф по этим формулам, приведем Ф к новым (прямоугольным) координатам:

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz=$$
$$=a'_{11}x'^2+a'_{22}y'^2+a'_{33}z'^2+2a'_{12}x'y'+2a_{13}x'z'+2a'_{23}y'z'. \quad (6)$$

Поскольку квадратичные формы, написанные здесь слева и справа, преобразуются друг в друга, они приводятся к одному и тому же каноническому виду. Следовательно, обе эти формы имеют одни и те же характеристические числа $\lambda_1,\ \lambda_2,\ \lambda_3.$ Напишем характеристическое уравнение для первоначально данной формы:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}=0.$$

Развертывая определитель, получим

$$\lambda^3-\sigma\lambda^2+\varkappa\lambda-\delta=0, \quad (7)$$

где

$$\sigma=a_{11}+a_{22}+a_{33}, \quad (8)$$

$$\varkappa=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad (9)$$

$$\delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \quad (10)$$

Напишем также характеристическое уравнение для формы, стоящей в правой части равенства (6):

$$\lambda^3-\sigma'\lambda^2+\varkappa'\lambda-\delta'=0. \quad (11)$$

Здесь $\sigma',\ \varkappa',\ \delta'$ — величины, которые выражаются равенствами (8), (9), (10), если в них $a_{11},\ldots,a_{33}$ заменить через $a'_{11},\ldots,a'_{33}$. Уравнение (7) и (11) имеют одни и те же корни: $\lambda_1,\lambda_2,\lambda_3,$ следовательно, не отличаются друг от друга. Таким образом,

$$\sigma'=\sigma,\quad \varkappa'=\varkappa,\quad \delta'=\delta.$$

Мы видим, что величины $\sigma,\ \varkappa,\ \delta$ не меняются при любом преобразовании формы к новым прямоугольным координатам (с тем же масштабом). Поэтому их называют *инвариантами* формы относительно преобразования прямоугольных координат. Особенно важным из них является инвариант

$$\delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \qquad (12)$$

который называется дискриминантом данной квадратичной формы. Из выражения (27) п° 44 видно, что

$$\delta = \lambda_1\lambda_2\lambda_3. \qquad (13)$$

48. Квадратичную форму от трех аргументов будем называть *эллиптической,* если ее характеристические числа $\lambda_1, \lambda_2, \lambda_3$ все отличны от нуля и имеют одинаковые знаки; *гиперболической,* если $\lambda_1, \lambda_2, \lambda_3$ все отличны от нуля, но среди них встречаются числа разных знаков; *параболической,* если хотя бы одно из чисел $\lambda_1, \lambda_2, \lambda_3$ равно нулю.

Согласно (13) *параболические формы характеризуются условием* $\delta = 0$.

Рассмотрим эллиптическую форму Ф; допустим, что ее характеристические числа положительны: $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$. Тогда из неравенств (3) п° 46 следует, что в любой точке *M(x, y, z),* не считая начала координат, величина Ф > 0. Такая форма называется *положительно определенной.* Если точка *M(x, y, z)* обегает единичную сферу $x^2 + y^2 + z^2 = 1$ и если $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$, то положительно определенная форма изменяется в положительных границах $0 < \lambda_1 \leqslant \Phi \leqslant \lambda_3$. Допустим, что $\lambda_1 < 0$, $\lambda_2 < 0$, $\lambda_3 < 0$ $(\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3)$. Тогда из неравенств (3) следует, что в любой точке (кроме начала координат) будет Ф<0. Такая форма называется *отрицательно определенной.* На единичной сфере отрицательно определенная форма изменяется в отрицательных границах $\lambda_1 \leqslant \Phi \leqslant \lambda_3 < 0$. Итак, *каждая эллиптическая форма является либо положительно определенной, либо отрицательно определенной.*

Рассмотрим теперь гиперболическую форму: если $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$, то $\lambda_1 < 0$, $\lambda_3 > 0$.

Таким образом, на единичной сфереформа Ф изменяется в границах λ_1, λ_3, одна из которых отрицательна, другая положительна. Отсюда, в частности, следует, что *гиперболическая форма является знакопеременной.*

49. Если $\delta \neq 0$, то форма будет либо эллиптической, либо гиперболической. Чтобы установить, к какому из этих типов относится данная форма, нужно определить знаки ее характеристических чисел, т. е. знаки корней характеристического уравнения. Поскольку мы уже знаем, что характеристическое уравнение имеет только вещественные корни, к нему применима известная в алгебре теорема Декарта: *если все корни уравнения вещественны и свободный член отличен от нуля, то число положительных корней этого уравнения равно числу перемен знаков в системе его коэффициентов* (при этом кратный корень учитывается столько раз, какова его кратность). Так как общее число корней характеристического уравнения известно (равно трем), то, применяя к этому уравнению теорему Декарта, мы найдем, сколько данная форма имеет положительных и сколько отрицательных характеристических чисел, и тем самым получим полную характеристику данной формы.

Пример. Определить тип квадратичной формы

$$\Phi = 7x^2 + 6y^2 + 5z^2 - 4xy - 4yz.$$

Решение. Составим характеристическое уравнение: $\lambda^3 - 18\lambda^2 + 99\lambda - 162 = 0$. Система коэффициентов того уравнения имеет следующие знаки, которые мы выписываем отдельно от коэффициентов: + — + —. Перемена знака происходит при переходе от первого коэффициента ко второму, затем при переходе от второго к третьему и, наконец, при переходе от третьего к четвертому. Общее число перемен знака равно трем. Следовательно, все три характеристических числа положительны: форма является эллиптической и положительно определенной.

Заметим, что в первом случае легко находятся и сами корни $\lambda_1 = 3, \quad \lambda_2 = 6, \quad \lambda_3 = 9.$ Зная их, мы можем установить границы изменения данной формы на единичной сфере: если $x^2 + y^2 + z^2 = 1$, то

$$3 \leqslant 7x^2 + 6y^2 + 5z^2 - 4xy - 4yz \leqslant 9.$$

Пример. Определить тип квадратичной формы

$$\Phi = x^2 - 2y^2 + z^2 + 4xy - 8xz - 4yz.$$

Решение. Составляем характеристическое уравнение: $\lambda^3 - 27\lambda - 54 = 0$. В системе его коэффициентов есть одна перемена знака (при переходе от первого к следующему коэффициенту). Таким образом, имеется одно положительное характеристическое число и два отрицательных. Данная форма является гиперболической.

5.10. Приведение к каноническому виду общего уравнения поверхности второго порядка

50. Пусть дано общее уравнение поверхности второго порядка:

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz+ \\ +2a_{14}x+2a_{24}y+2a_{34}z+a_{44}=0. \qquad (1)$$

Левую часть уравнения (1) образуют следующие группы членов:

1) квадратичная форма, состоящая из старших членов:

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz;$$

2) линейная форма членов первой степени

$$2a_{14}x+2a_{24}y+2a_{34}z;$$

3) свободный член a_{44}.

Наша цель — путем перехода к новым прямоугольным координатам привести уравнение (1) к каноническому виду и тем самым установить вид данной поверхности. Когда мы говорим о приведении уравнения (1) к каноническому виду, то имеем в виду следующее: 1) нужно добиться, чтобы получила канонический вид квадратичная форма старших членов; 2) кроме того, чтобы число членов первой степени стало наименьшим (если возможно, совсем их уничтожить); 3) кроме того, если возможно, уничтожить свободный член.

Решение этой задачи проводится в общем так же, как решение аналогичной задачи для кривых второго порядка (чем мы занимались в 5.4).

Предположим, что система координатных осей как-нибудь поворачивается; тогда координаты изменятся по формулам вида

$$\left.\begin{aligned} x&=l_1x'+l_2y'+l_3z',\\ y&=m_1x'+m_2y'+m_3z',\\ z&=n_1x'+n_2y'+n_3z'. \end{aligned}\right\} \qquad (2)$$

Если мы заменим *x, y, z* в уравнении (1) их выражениями по формулам (2), то каждая из указанных выше групп членов левой части уравнения преобразуется автономно (не влияя на другие). Направим новые координатные оси по главным направлениям квадратичной формы старших членов (их называют также главными направлениями данной поверхности). Тогда

1) квадратичная форма старших членов примет канонический вид:

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz=$$
$$=\lambda_1 x'^2+\lambda_2 y'^2+\lambda_3 z'^2;$$

2) группа членов первой степени перейдет в аналогичную группу членов с другими коэффициентами:

$$2a_{14}x+2a_{24}y+2a_{34}z=2\mu_1 x'+2\mu_2 y'+2\mu_3 z';$$

3) свободный член останется без изменения.
Мы получим уравнение данной поверхности в новых координатах:

$$\lambda_1 x'^2+\lambda_2 y'^2+\lambda_3 z'^2+2\mu_1 x+2\mu_2 y+2\mu_3 z+a_{44}=0. \quad (3)$$

Дальнейшее упрощение этого уравнения производится в зависимости от типа квадратичной формы старших членов.

51. Предположим, что дискриминант формы старших членов отличен от нуля: δ=0. Так как $\delta=\lambda_1\lambda_2\lambda_3$, то $\lambda_1\neq 0$, $\lambda_2\neq 0$, $\lambda_3\neq 0$ и мы можем уравнение (3) переписать так;

$$\lambda_1\left(x'+\frac{\mu_1}{\lambda_1}\right)^2+\lambda_2\left(y'+\frac{\mu_2}{\lambda_2}\right)^2+\lambda_3\left(z'+\frac{\mu_3}{\lambda_3}\right)^2=$$
$$=-a_{44}+\frac{\mu_1^2}{\lambda_1}+\frac{\mu_2^2}{\lambda_2}+\frac{\mu_3^2}{\lambda_3}. \quad (4)$$

Совершим параллельный перенос системы осей *Ox', Oy', Oz'* с расчетом, чтобы координаты изменились по формулам:

$$x'=x''-\frac{\mu_1}{\lambda_1},\quad y'=y''-\frac{\mu_2}{\lambda_2},\quad z'=z''-\frac{\mu_3}{\lambda_3}.$$

Если мы обозначим правую часть (4) одной буквой *H,* то уравнение данной поверхности в последних координатах примет канонический вид:

$$\lambda_1 x''^2+\lambda_2 y''^2+\lambda_3 z''^2=H. \quad (5)$$

В случае $H\neq 0$ уравнение (5) можно переписать так:

$$\frac{x''^2}{\frac{H}{\lambda_1}}+\frac{y''^2}{\frac{H}{\lambda_2}}+\frac{z''^2}{\frac{H}{\lambda_3}}=1. \quad (6)$$

Возможны два случая.

1) Квадратичная форма старших членов данного уравнения является *эллиптической;* тогда числа λ_1, λ_2, λ_3 одного знака. Будем считать, что $\lambda_1>0,\ \lambda_2>0,\ \lambda_3>0.$

Если при этом $H>0$, то уравнение (5) приводится к виду (6) и определяется *эллипсоид* с полуосями

$$a=\sqrt{\frac{H}{\lambda_1}},\ b=\sqrt{\frac{H}{\lambda_2}},\ c=\sqrt{\frac{H}{\lambda_3}}.$$

Если H = 0, то уравнение (5) определяет единственную вещественную точку: $x''=0,\ y''=0,\ z''=0.$ Однако можно говорить, что в этом случае уравнение (5) определяет *мнимый конус.* Мнимый конус можно считать также *вырожденным эллипсоидом* (считая, что уравнение (5) при H = 0 получается в пределе при H→0).

Если H<0, то уравнение (5) никакого вещественного образа не определяет. В этом случае говорят также, что уравнение (5) есть *уравнение мнимого эллипсоида.*

2) Квадратичная форма старших членов данного уравнения является *гиперболической;* тогда среди чисел $\lambda_1,\ \lambda_2,\ \lambda_3$ два будут иметь одинаковые знаки и одно — противоположный знак. Будем считать, что $\lambda_1>0,\ \lambda_2>0,\ \lambda_3<0.$

Если при этом H>0, то уравнение (5) приводится к виду (6) и может быть написано еще таким образом:

$$\frac{x''^2}{\frac{H}{\lambda_1}}+\frac{y''^2}{\frac{H}{\lambda_2}}-\frac{z''^2}{\left|\frac{H}{\lambda_3}\right|}=1.$$

Это уравнение определяет *однополостный гиперболоид.*
Если H<0, то, деля почленно уравнение (5) на положительное число — H, приведем его к виду

$$\frac{x''^2}{\left|\frac{H}{\lambda_1}\right|}+\frac{y''^2}{\left|\frac{H}{\lambda_2}\right|}-\frac{z''^2}{\left|\frac{H}{\lambda_3}\right|}=-1.$$

Такое уравнение определяет *двухполостный гиперболоид.*
Наконец, если H=0, то, переписав уравнение (5) следующим образом:

$$\lambda_1x''^2+\lambda_2y''^2-|\lambda_3|z''^2=0, \qquad (7)$$

видим, что в этом случае оно определяет *вещественный конус.*

Заметим, что уравнение (5) при H=0 получается из уравнения того же вида при H≠0 путем предельного перехода, если H→0. Поэтому *вещественный конус второго порядка можно рассматривать как вырождение однополостного или двухполостного гиперболоида* (в зависимости от того, как стремится H к нулю, оставаясь положительным или отрицательным). На рис. 6 изображен конус и два близких к нему гиперболоида (однополостный и двухполостный). Эти три поверхности соответствуют случаям $H=0$, H>0 и H<0; когда H приближается к нулю, то оба гиперболоида приближаются к конусу, как к пределу.

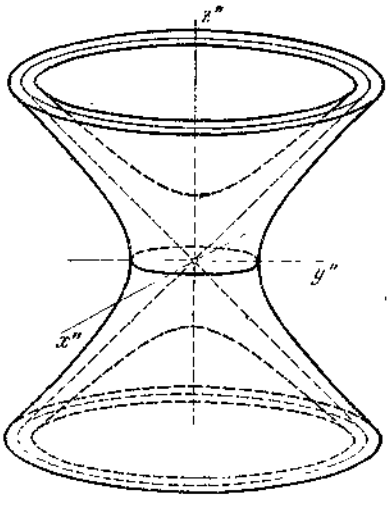


Рис. 6

52. Предположим теперь, что квадратичная форма старших членов уравнения (1) имеет дискриминант δ = 0. Так как $\delta=\lambda_1\lambda_2\lambda_3$, то на этот раз хотя бы одно из чисел λ_1, λ_2, λ_3 равно нулю (т. е. квадратичная форма является *параболической).* Будем считать, что $\lambda_3=0$, $\lambda_1\neq 0$ (все три числа λ_1, λ_1, λ_3 не могут исчезнуть, так как при $\lambda_1=\lambda_2=\lambda_3=0$ сама форма была бы тождественна нулю). Возможны два случая.

1) $\lambda_2\neq 0$; тогда из (3) имеем уравнение

$$\lambda_1 x'^2+\lambda_2 y'^2+2\mu_1 x'+2\mu_2 y'+2\mu_3 z'+a_{44}=0, \qquad (8)$$

или

$$\lambda_1\left(x'+\frac{\mu_1}{\lambda_1}\right)^2+\lambda_2\left(y'+\frac{\mu_2}{\lambda_2}\right)^2+2\mu_3 z'+H=0, \qquad (9)$$

где $H=a_{44}-\frac{\mu_1^2}{\lambda_1}-\frac{\mu_2^2}{\lambda_2}$.

Если $\mu_3\neq 0$, то, вводя еще новые координаты x'', y'',z'' по фоомулам

$$x'=x''-\frac{\mu_1}{\lambda_1},\quad y'=y''-\frac{\mu_2}{\lambda_2},\quad z'=z''-\frac{H}{2\mu_3},$$

получим из (9) каноническое уравнение данной поверхности:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + 2\mu_3 z'' = 0.$$

Это уравнение можно переписать следующим образом:

$$2z'' = \frac{x''^2}{-\frac{\mu_3}{\lambda_1}} + \frac{y''^2}{-\frac{\mu_3}{\lambda_2}}. \qquad (10)$$

Мы видим, что при $\mu_3 \neq 0$ уравнение (8) определяет *параболоид* (эллиптический или гиперболический в зависимости от знаков знаменателей в правой части (10)).

Если $\mu_3 = 0$, то уравнение (8) приводится к виду

$$\lambda_1 x''^2 + \lambda_2 y''^2 + H = 0$$

и, следовательно, определяет *эллиптический* или *гиперболический* цилиндр.

2) $\lambda_2 = 0$; тогда из (8) имеем уравнение

$$\lambda_1 x'^2 + 2\mu_1 x' + 2\mu_2 y' + 2\mu_3 z' + a_{44} = 0, \qquad (11)$$

или

$$\lambda_1 \left(x' + \frac{\mu_1}{\lambda_1}\right)^2 + 2\mu_2 y' + 2\mu_3 z' + H = 0, \qquad (12)$$

где

$$H = a_{44} - \frac{\mu_1^2}{\lambda_1}.$$

Если $\mu_2 \neq 0, \ \mu_3 \neq 0$, то мы повернем систему осей *Ox', Oy', Oz'* вокруг оси *Ox'* на угол α, беря

$$\operatorname{tg}\alpha = -\frac{\mu_2}{\mu_3}$$

и отсчитывая угола α от оси *Oy'* к оси *Oz';* систему повернутых осей перенесем параллельно на величину $-\frac{\mu_1}{\lambda_1}$ вдоль оси *Ox'*. Соответствующее преобразование координат дается формулами

$$x'' = x' + \frac{\mu_1}{\lambda_1},$$

$$y'' = \frac{\mu_2}{\mu} y' + \frac{\mu_3}{\mu} z',$$

$$z'' = -\frac{\mu_2}{\mu} y' + \frac{\mu_3}{\mu} z',$$

где $\mu = \sqrt{\mu_2^2 + \mu_3^2}$. В силу этого преобразования уравнение (12) примет вид

$$\lambda_1 x''^2 + 2\mu y'' + H = 0. \tag{13}$$

Производя еще параллельный перенос полученных осей на величину $\frac{H}{2\mu}$ вдоль оси ординат (что соответствует преобразованию $x'' = x''', \ y'' = y''' - \frac{H}{2\mu}\Big)$, найдем каноническое уравнение данной поверхности:

$$\lambda_1 x'''^2 + 2\mu y''' = 0.$$

Мы видим, что при наших предположениях $\mu_2 \neq 0, \ \mu_3 \neq 0,$ уравнение (12) определяет *параболический цилиндр.*

Если в уравнении (12) будет $\mu_2 \neq 0, \ \mu_3 \neq 0,$ то оно приводится к виду (13) одним лишь параллельным переносом системы осей *Ox', Oy', Oz'* вдоль *Ox'* на величину $-\frac{\mu_1}{\lambda_1}$ и, следовательно, также определяет параболический цилиндр. В случае $\mu_2 = 0, \ \mu_3 \neq 0$ дело сводится к предыдущему, если поменять ролями *y'* и *z'*.

Наконец, если $\mu_2 = 0, \ \mu_3 = 0,$ то уравнение (12) очевидным образом приводится к каноническому виду:

$$\lambda_1 x''^2 + H = 0. \tag{14}$$

Уравнение (14) определяет *пару параллельных плоскостей* (вещественных при $\lambda_1 > 0, \ H \leqslant 0,$ мнимых при $\lambda_1 > 0$, $H > 0$).

З а м е ч а н и е. Как установлено выше, уравнение (8) при условиии $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$, $\mu_3 \neq 0$ определяет параболоид

Если $\mu_3 \to 0$ или $\lambda_2 \to 0$, то в пределе из (8) получается уравнение цилиндра. Полому *цилиндры второго порядкам можно считать вырожденными параболоидами.*

53. Подводя итог проведенному исследованиию, мы можем сделать следующее заключение: I

Если группа старших членов уравнения (1) *является эллиптической, то это уравнение определяет эллипсоид* (обыкновенный, мнимый или вырожденный); *если группа старших членов является гиперболической, то уравнение определяет гиперболоид* (однополостный, двухполостный или вырожденный); *если группа старших членов является параболической* ($\delta = 0$), *то уравнение определяет параболоид* (эллиптический, гиперболический или вырожденный).

5.11. Уравнения центра. Признак вырождения поверхности второго порядка.

54. Если нужно привести к каноническому виду общее уравнение поверхности второго порядка, обладающей центром, то имеет смысл прежде всего перенести начало координат в центр поверхности, после чего нужным образом повернуть оси. Имея в виду такой план, выведем уравнения, которые определяют центр поверхности второго порядка (если он есть) непосредственно по общему уравнению этой поверхности.

55. Некоторую точку *S* мы называем *центром* данной поверхности второго порядка, если после переноса начала координат в точку *S* уравнение поверхности не будет содержать членов первой степени (см. п° 34).

Пусть дана поверхность второго порядка:

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + \\ + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = 0. \qquad (1)$$

Требуется найти ее центр (или убедиться, что центра нет). Предполагая, что центр имеется, обозначим его, как и раньше, буквой *S;* обозначим через x_0, y_0, z_0 искомые координаты центра (в данной координатной системе). Перенесем начало координат в точку *S;* при этом координаты произвольной точки изменятся по формулам:

$$x = \tilde{x} + x_0, \quad y = \tilde{y} + y_0, \quad z = \tilde{z} + z_0. \qquad (2)$$

Перейдем в уравнении (1) к новым координатам. Чтобы упростить запись дальнейших соотношений, обозначим всю левую часть уравнения (1) символом *F(x, y, z).* Заменяя здесь *x, y, z* их выражениями по формулам (2), получим

$$F(x, y, z) = F(\tilde{x} + x_0, \tilde{y} + y_0, \tilde{z} + z_0) = \\ = a_{11}\tilde{x}^2 + a_{22}\tilde{y}^2 + a_{33}\tilde{z}^2 + 2a_{12}\tilde{x}\tilde{y} + 2a_{13}\tilde{x}\tilde{z} + \\ + 2a_{23}\tilde{y}\tilde{z} + 2A_{14}\tilde{x} + 2A_{24}\tilde{y} + 2A_{34}\tilde{z} + A_{44}, \qquad (3)$$

где

$$\left.\begin{aligned} A_{14} &= a_{11}x_0 + a_{12}y_0 + a_{13}z_0 + a_{14}, \\ A_{24} &= a_{21}x_0 + a_{22}y_0 + a_{23}z_0 + a_{24}, \\ A_{34} &= a_{31}x_0 + a_{32}y_0 + a_{33}z_0 + a_{34}, \\ A_{44} &= F(x_0, y_0, z_0); \end{aligned}\right\} \qquad (4)$$

коэффициенты старших членов остаются прежними.

Преобразование левой части уравнения (1) происходит согласно этим формулам в любом случае, т. е. независимо от того, что представляет собой точка *S*. Точка *S* будет центром, если $A_{14}=0$, $A_{24} = 0$, $A_{34} = 0$. Отсюда получаем уравнения центра:

$$\left.\begin{aligned} a_{11}x_0 + a_{12}y_0 + a_{13}z_0 + a_{14} = 0, \\ a_{21}x_0 + a_{22}y_0 + a_{23}z_0 + a_{24} = 0, \\ a_{31}x_0 + a_{32}y_0 + a_{33}z_0 + a_{34} = 0. \end{aligned}\right\} \tag{5}$$

Решая их совместно, найдем центр $S(x_0, y_0, z_Q)$. Система (5) может оказаться несовместной; тогда центра у данной поверхности нет. Напишем определитель системы (5):

$$\delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix};$$

он совпадает с уже известным нам дискриминантом группы старших членов.

Если $\delta \neq 0$, то система (5) совместна и имеет единственное решение. Следовательно, если $\delta \neq 0$, то данная поверхность имеет единственный центр. Такая поверхность второго порядка называется *центральной.* В случае центральной поверхности координаты центра выражаются следующими формулами:

$$x_0 = \frac{\delta_x}{\delta}, \quad y_0 = \frac{\delta_y}{\delta}, \quad z_0 = \frac{\delta_z}{\delta}, \tag{6}$$

где δ_x — определитель третьего порядка, который получается из определителя δ при помощи замены элементов его первого столбца числами

$$-a_{14}, \quad -a_{24}, \quad -a_{34}; \quad \delta_y \text{ и } \delta_z$$

получаются из определителя δ путем аналогичной замены элементов его второго и соответственно третьего столбца. Из формул (4) и (6) найдем A_{41}. Мы имеем:

$$\begin{aligned} A_{44} = F(x_0, y_0, z_0) = (a_{11}x_0 + a_{12}y_0 + a_{13}z_0 + a_{14})\,x_0 + \\ + (a_{21}x_0 + a_{22}y_0 + a_{23}z_0 + a_{24})\,y_0 + (a_{31}x_0 + a_{32}y_0 + \\ + a_{33}z_0 + a_{34})\,z_0 + (a_{41}x_0 + a_{42}y_0 + a_{43}z_0 + a_{44}). \end{aligned}$$

Отсюда и вследствие системы (5)

$$A_{44} = a_{41}x_0 + a_{42}y_0 + a_{43}z_0 + a_{44}. \tag{7}$$

Используя (6), получим

$$A_{44}=\frac{a_{41}\delta_x+a_{42}\delta_y+a_{43}\delta_z+a_{44}\delta}{\delta}.$$

Числитель этой дроби может быть записан в виде определителя четвертого порядка:

$$a_{41}\delta_x+a_{42}\delta_y+a_{43}\delta_z+a_{44}\delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}. \tag{8}$$

Чтобы убедиться в справедливости равенства (8), проще всего разложить написанный справа определитель по элементам последней строки и установить, что алгебраические дополнения элементов этой строки совпадают с величинами δ_x, δ_y, δ_z, δ .

Определитель (8) называется дискриминантом левой части уравнения (1) и обозначается символом Δ:

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}.$$

Из равенства (8) окончательно имеем

$$A_{44}=\frac{\Delta}{\delta}.$$

Итак, если поверхность, заданная уравнением (1), является центральной ($\delta \neq 0$), то после переноса начала координат в ее центр данное уравнение приводится к виду

$$a_{11}\tilde{x}^2+a_{22}\tilde{y}^2+a_{33}\tilde{z}^2+2a_{12}\tilde{x}\tilde{y}+2a_{13}\tilde{x}\tilde{z}+2a_{23}\tilde{y}\tilde{z}+\frac{\Delta}{\delta}=0; \tag{9}$$

старшие коэффициенты — прежние. Совершая теперь надлежащий поворот осей, можно привести уравнение (9) к каноническому виду:

$$\lambda_1 x'^2+\lambda_2 y'^2+\lambda_3 z'^2+\frac{\Delta}{\delta}=0. \tag{10}$$

56. Сравнивая последнее уравнение с уравнением (5) п.5.10, мы видим, чго свободный член H в уравнении (5) п.5.10 может быть подсчитан по данному уравнению (1) сразу, без того, чтобы фактически выполнять преобразование координат; именно,

$$H=-\frac{\Delta}{\delta}.$$

В п.5.10 установлено, что уравнение (5) определяет вырожденную поверхность (конус) при H=0. Отсюда заключаем: центральная поверхность второго порядка является вырожденной тогда и только тогда, когда $\Delta = 0$.

Укажем (уже без доказательства), что равенство $\Delta = 0$ характеризует также и вырожденные параболоиды (т. е. цилиндры). Таким образом, справедливо общее утверждение: *уравнение* (1) *определяет вырожденную поверхность* (конус или цилиндр) *в том и только в том случае, когда* Δ=0.

Отсюда уже следует, что цилиндры, как вырожденные параболоиды, характеризуются двумя условиями: $\delta = 0$, $\Delta = 0$.

57. Пример. Привести к каноническому виду уравнение

$$7x^2+6y^2+5z^2-4xy-4yz-6x-24y+18z+18=0.$$

Решение. Прежде всего подсчитаем δ:

$$\delta=\begin{vmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{vmatrix}=162.$$

Так как $\delta \neq 0$, то данное уравнение определяет центральную поверхность. Координаты центра могут быть найдены из уравнений (5), которые в данном случае имеют следующий вид:

$$\begin{aligned} 7x_0-2y_0 \qquad\qquad\quad -3&=0, \\ -2x_0+6y_0-2z_0-12&=0, \\ -2y_0+5z_0+9&=0. \end{aligned}$$

Решая эту систему, находим $x_0=1$, $y_0 = 2$, $z_0=-1$. Перенесем начало координат в точку $S(1;\ 2;\ -1)$; при этом координаты преобразуются по формулам

$$x=\tilde{x}+1, \quad y=\tilde{y}+2, \quad z=\tilde{z}-1,$$

а данное уравнение принимает вид

$$7\tilde{x}^2+6\tilde{y}^2+5\tilde{z}^2-4\tilde{x}\tilde{y}-4\tilde{y}\tilde{z}-18=0. \qquad (*)$$

Свободный член полученного уравнения мы подсчитали по формуле (7): $A_{44}=-3x_0-12y_0+9z_0+18=-18.$ Можно поступить иначе: именно имея

$$\Delta=\begin{vmatrix} 7 & -2 & 0 & -3 \\ -2 & 6 & -2 & -12 \\ 0 & -2 & 5 & 9 \\ -3 & -12 & 9 & 18 \end{vmatrix}=-2916,$$

найдем $A_{44}=\frac{\Delta}{8}=-18.$ Производя дальнейшее преобразование (*) согласно п° 43, получим каноническое уравнение данной поверхности: $3x'^2+6y'^2+9z'^2-18=0.$

Пример. Привести к каноническому виду уравнение

$$2x^2+2y^2+3z^2+4xy+2xz+2yz-4x+6y-2z+3=0. \qquad (11)$$

Решение. Подсчитаем дискриминант старших членов:

$$\delta=\begin{vmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{vmatrix}=0.$$

Так как δ = 0, то данная поверхность не является центральной, и мы начнем с упрощения группы старших членов (см. п° 50 и 52).
Составим характеристическое уравнение:

$$\begin{vmatrix} 2-\lambda & 2 & 1 \\ 2 & 2-\lambda & 1 \\ 1 & 1 & 3-\lambda \end{vmatrix}=0$$

или

$$\lambda^3-7\lambda^2+10\lambda=0.$$

Корни этого уравнения суть $\lambda_1=2,\ \lambda_2=5,\ \lambda_3=0.$ Составим применительно к данной задаче систему уравнения (14) п.5.8:

$$\left.\begin{aligned} (2-\lambda)\, l+2m+n=0, \\ 2l+(2-\lambda)\, m+n=0, \\ l+m+(3-\lambda)\, n=0. \end{aligned}\right\} \qquad (12)$$

Полагая в уравнениях (3) $\lambda=\lambda_1=2$, получим

$$\begin{aligned} 2m+n=0, \\ 2l \qquad +n=0, \\ l+m+n=0. \end{aligned}$$

В качестве ненулевого решения этой системы можно взять,, например, l=1, m=1, n=—2; соответственно имеем вектор $\boldsymbol{a}_1$={l; 1; —2}, определяющий первое главное направление.

Полагая в уравнениях (3) $\lambda=\lambda_2=5$, получим

$$\begin{aligned} -3l+2m+n=0, \\ 2l-3m+n=0, \\ l+m-2n=0. \end{aligned}$$

В качестве ненулевого решения этой системы можно взять, например l=1, m=1, n=1; соответственно имеем вектор второго главного направления $\boldsymbol{a}_2$={l; 1; 1}. Полагая в уравнениях (3)
$\lambda=\lambda_3=0$, получим

$$\begin{aligned} 2l+2m+n&=0,\\ 2l+2m+n&=0,\\ l+m+3n&=0. \end{aligned}$$

В качестве ненулевого решения этой системы можно взять, например, l =1, m =—1, n = 0; соответственно имеем вектор $\boldsymbol{a}_3$={l; —1; 0} идущий по третьему главному направлению.

Направим новые оси по векторам $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$. Чтобы составить формулы преобразований координат, найдем базисные векторы новых осей:

$$\begin{aligned} i'&=\frac{a_1}{|a_1|}=\left\{\frac{1}{\sqrt{6}};\ \frac{1}{\sqrt{6}};\ -\frac{2}{\sqrt{6}}\right\},\\ j'&=\frac{a_2}{|a_2|}=\left\{\frac{1}{\sqrt{3}};\ \frac{1}{\sqrt{3}};\ \frac{1}{\sqrt{3}}\right\},\\ k'&=\frac{a_3}{|a_3|}=\left\{\frac{1}{\sqrt{2}};-\frac{1}{\sqrt{2}};\ 0\right\}. \end{aligned}$$

Отсюда и согласно п° 26 получаем искомые формулы:

$$\begin{aligned} x&=\frac{1}{\sqrt{6}}x'+\frac{1}{\sqrt{3}}y'+\frac{1}{\sqrt{2}}z',\\ y&=\frac{1}{\sqrt{6}}x'+\frac{1}{\sqrt{3}}y'-\frac{1}{\sqrt{2}}z',\\ z&=-\frac{2}{\sqrt{6}}x'+\frac{1}{\sqrt{3}}y'. \end{aligned}$$

Зная корни характеристического уравнения, мы можем сразу написать результат преобразования группы старших членов уравненения (2) к новым координатам; именно мы получим $2x'^2+5y'^2$. Остальные члены уравнения (2) приведем к новым координатам, пользуясь предыдущими формулами:

$$-4x+6y-2z+3=\sqrt{6}x'-5\sqrt{2}z'+3.$$

Таким образом, уравнение данной поверхности в новых координатах имеет вид

$$2x'^2+5y'^2+\sqrt{6}x'-5\sqrt{2}z'+3=0,$$

или, после некоторой перегруппировки членов,

$$2\left(x'^2+\frac{\sqrt{6}}{2}x'\right)+5y'^2-5\sqrt{2}z'+3=0.$$

Дополняя выражение внутри круглой скобки до полного квадрата, получим

$$2\left(x'+\frac{\sqrt{6}}{4}\right)^2+5y'^2-5\sqrt{2}\left(z'-\frac{9\sqrt{2}}{40}\right)=0.$$

Совершим теперь параллельный перенос осей так, чтобы координаты преобразовались по формулам:

$$x'=x''-\frac{\sqrt{6}}{4},$$

$$y'=y'',$$

$$z'=z''+\frac{9\sqrt{2}}{40}.$$

Мы получили каноническое уравнение данной поверхности

$$2x''^2+5y''^2-5\sqrt{2}z''=0,$$

или

$$5\sqrt{2}z''=2x''^2+5y''^2.$$

Данная поверхность является *эллиптическим параболоидом.*

Модуль 6
Линейные преобразования и матрицы

<u>Микромодуль 15</u>
Линейные преобразования на плоскости

6.1. Линейные преобразования на плоскости

1. Мы будем рассматривать некоторую плоскость α. Отметим на этой плоскости какую-нибудь точку O. Тогда любая точка M плоскости α определяет радиус-вектор $\boldsymbol{x}=\overline{OM}$. Если нам заранее дан произвольный вектор $\boldsymbol{x}$, лежащий в плоскости α, то мы будем считать его приложенным к точке *О.* Таким образом, каждый вектор данном плоскости будет рассматриваться как радиус-вектор какой-нибудь точки.

2. Пусть дано правило, согласно которому с любой точкой *М* плоскости α сопоставлена некоторая точка *М'* той же плоскости (иначе говоря, точка *М* перемещена в некоторую точку *М').* В таком случае мы скажем, что на плоскости α задано *преобразование точек;* точку *М'* назовем *образом* точки *М.*

Далее мы всегда будем предполагать, что при заданном преобразовании точка *О* остается на месте (совпадает со своим образом).

3. Наряду с заданным преобразованием точек мы будем рассматривать *преобразование векторов* плоскости α, при котором произвольный вектор $\boldsymbol{x}=\overline{OM}$ переводится в вектор $\boldsymbol{x}'=\overline{OM'}$. Вектор $\boldsymbol{x}'$ назовем *образом* вектора $\boldsymbol{x}$. То обстоятельство, что $\boldsymbol{x}'$ является образом вектора $\boldsymbol{x}$, будем символически записывать так: $\boldsymbol{x}' = A\boldsymbol{x}$. Этой же записью будем выражать и самый факт, что дано определенное преобразование (если нам придется рассматривать еще другие преобразования, то мы будем писать $\boldsymbol{x}'=B\boldsymbol{x}$, $\boldsymbol{x}' = C\boldsymbol{x}$ и т. п.).

4. *Преобразование* $\boldsymbol{x}'=A\boldsymbol{x}$ *называется линейным, если соблюдены следующие два условия:*

1) $A(\lambda\boldsymbol{x})=\lambda A\boldsymbol{x}$, где $\boldsymbol{x}$ — *любой вектор плоскости* α, λ — *любое число;*

2) $A(\boldsymbol{x}+\boldsymbol{y})=A\boldsymbol{x}+A\boldsymbol{y}$, *где* $\boldsymbol{x}$ *и* $\boldsymbol{y}$ — *любые два вектора плоскости* α.

Поясним эти условия наглядно. Договоримся только в дальнейшем не указывать каждый раз, что речь идет о точках и векторах плоскости α. Остановимся прежде всего на первом условии. Как известно, вектор $\lambda\boldsymbol{x}$ коллинсарен вектору $\boldsymbol{x}$ и получается растяжением его в λ раз; первое условие означает, что образ вектора $\lambda\boldsymbol{x}$ также коллинеарен образу $\boldsymbol{x}$ и также получается из него растяжением в λ раз (см. рис. 1, где $\boldsymbol{x}=\overline{OM}$, $\lambda\boldsymbol{x}=\overline{ON}$, $A\boldsymbol{x}=\overline{OM'}$, $A(\lambda\boldsymbol{x})=\overline{ON'}$; так как $A(\lambda\boldsymbol{x})=\lambda A\boldsymbol{x}$, то $\overline{ON'}$ получается из $\overline{OM'}$ таким же растяжением, каким $\overline{ON}$ получается из $\overline{OM}$).

Чтобы пояснить второе условие, положим $\boldsymbol{x}=\overline{OM}$, $\boldsymbol{y}=\overline{ON}$, $\boldsymbol{x}+\boldsymbol{y}=\overline{OP}$ (рис. 2).

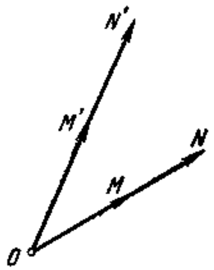


Рис. 1.

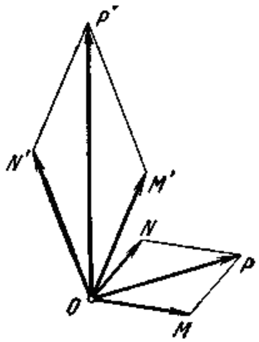


Рис. 2

Пусть M', N', P' — образы точек M, N, P при данном преобразовании. Тогда $Ax = \overline{OM'}$, $Ay = \overline{ON'}$, $A(x+y) = \overline{OP'}$ и согласно второму условию

$$\overline{OP'} = A(x+y) = Ax + Ay = \overline{OM'} + \overline{ON'}.$$

Следовательно, второе условие означает, что каждый параллелограмм *OMNP* преобразуется в четырехугольник *OM'N'P'*, который также является параллелограммом.

Пример 1. Преобразование заключается в том, что все векторы растягиваются в k раз (k — некоторое число):

$$Ax = kx.$$

Оба условия соблюдены. В самом деле,

$$A(\lambda x) = k(\lambda x) = \lambda(kx) = \lambda Ax;$$

кроме того

$$A(x+y) = k(x+y) = kx + ky = Ax + Ay.$$

Следовательно, данное преоОразование является линейным. Такое линейное преобразование называют *подобием,* а число k— *коэффициентом подобия.*

Если $x = \overline{OM}$, $Ax = \overline{OM'}$ и если точка M обегает некоторую фигуру F, то M' обегает фигуру F', которая подобна фигуре F; точка O будет центром подобия F и F' (рис. 3).

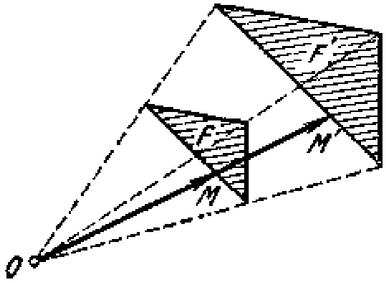


Рис. 3

Пример 2. Преобразование заключается в том, что все векторы поворачиваются вокруг точки *O* в одну и ту же сторону на один и тот же угол: $\boldsymbol{x}' = A\boldsymbol{x}$, где $\boldsymbol{x}'$ получен поворотом вектора $\boldsymbol{x}$ на данный угол α. Геометрически очевидно, что оба условия соблюдены и, следовательно, данное преобразование является линейным. Такое линейное преобразование называется *вращением* на угол α. Если $\boldsymbol{x} = \overline{OM}$. $A\boldsymbol{x} = \overline{OM'}$ и если точка *M* обегает некоторую фигуру *F*, то *M'* обегает фигуру *F* ', которая получается поворотом *F* вокруг *O* на данный угол α (рис. 4).

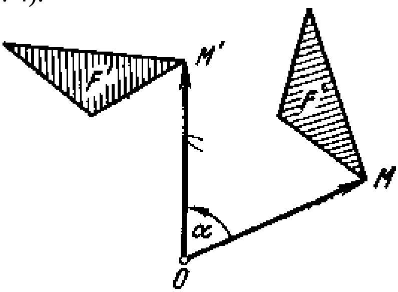


Рис. 4

Пример 3. Пусть *a* —какая-нибудь прямая, проходящая через точку *O*. Сопоставим с произвольным вектором $\boldsymbol{x}$ симметричный ему относительно прямой *a* вектор $\boldsymbol{x}' = A\boldsymbol{x}$. Преобразование $\boldsymbol{x}' = A\boldsymbol{x}$ является линейным, поскольку оба условия линейности с очевидностью соблюдены. Такое линейное преобразование называется *зеркальным отражением* относительно прямой *a*. Если $\boldsymbol{x} = \overline{OM}$. $A\boldsymbol{x} = \overline{OM'}$ и если точка *M* обегает некоторую фигуру *F*, то M' обегает фигуру F', которая является зеркальным образом F относительно прямой a (рис. 5).

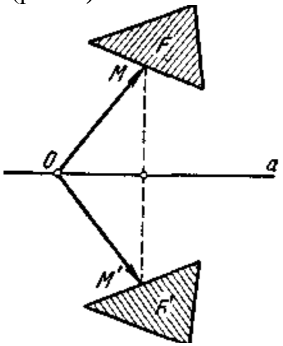


Рис. 5

П р и м е р 4. Пусть a —какая-нибудь прямая, проходящая через точку O. Сопоставим с произвольным вектором $\boldsymbol{x}=\overline{OM}$ вектор $\boldsymbol{x}'=A\boldsymbol{x}=\overline{OM'}$ при следующих условиях: если P — основание перпендикуляра, опущенного из M на a, то M' лежит на луче PM, причем отношение PM' к PM равно заранее данному положительному числу k (например, если k=1/2, то M переходит в точку M', которая вдвое ближе к a, чем M; если k = 2, то каждая точка M удаляется от прямой на расстояние, вдвое большее первоначального). Данное преобразование $\boldsymbol{x}' = A\boldsymbol{x}$ является линейным, в чем можно убедиться проверкой условий линейности; позднее это будет установлено как следствие некоторых общих положений (см. п. 8). Рассматриваемое в этом примере линейное преобразование называется *сжатием* к прямой a; число k называется *коэффициентом сжатия* (заметим, что сжатие с коэффициентом $k > 1$ по существу является растяжением). Если точка M обегает некоторую фигуру F, то M' обегает фигуру F', которая также называется сжатием F к прямой a. На рис. 6 в качестве F изображен круг; его сжатием F' оказывается эллипс.

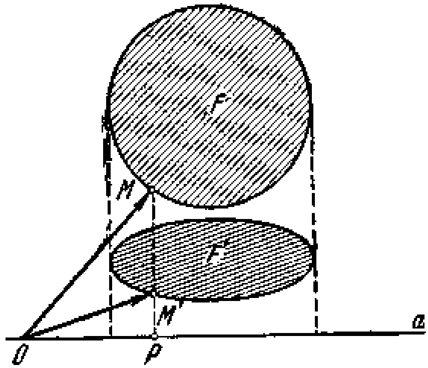


Рис. 6

5. Наша ближайшая цель — выразить линейное преобразование в координатах. Выберем любую пару векторов $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ с единственным условием, чтобы эти векторы были неколлинеарны. Считая $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ приложенными к точке *O*, рассмотрим оси Ox_1, Ox_2, по которым направлены векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ (рис. 7; естественно, мы полагаем, что положительные направления осей совпадают с направления-
ми векторов $\boldsymbol{e}_1$, $\boldsymbol{e}_2$).

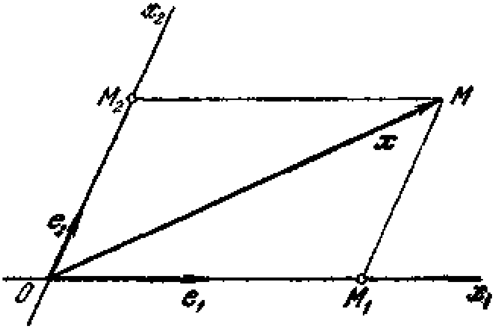


Рис. 7

Пусть $\boldsymbol{x}$ — произвольный вектор. Так как векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ неколлинеарны, то мы можем разложить вектор $\boldsymbol{x}$ по базису $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ (напомним читателю, что мы рассматриваем векторы, лежащие в одной определенной плоскости). Иначе говоря, вектор $\boldsymbol{x}$ можно представить в виде суммы

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2, \tag{1}$$

где $x_1\boldsymbol{e}_1 = \overline{OM_1}$, $x_2\boldsymbol{e}_2 = \overline{OM_2}$ — компоненты вектора $\boldsymbol{x}$ по осям Ox_1, Ox_2. Число x_1 представляет собой величину отрезка OM_1 оси Ox_1, при условии, что в качестве масштабного отрезка на этой оси принят вектор $\boldsymbol{e}_1$; аналогично x_2 — величина OM_2 в масштабе $\boldsymbol{e}_2$. Числа x_1, x_2 называются *координатами вектора* $\boldsymbol{x}$ в *базисе* $\boldsymbol{e}_1$, $\boldsymbol{e}_2$. В дальнейшем, наряду с выражением (1), мы будем писать также $\boldsymbol{x} = \{x_1; x_2\}$.

Если $\boldsymbol{x} = \overline{OM}$, то числа x_1, x_2 называются также *координатами точки* M в системе координат с данным началом O и с данными масштабными векторами $\boldsymbol{e}_1$, $\boldsymbol{e}_2$. То обстоятельство, что M имеет координаты x_1, x_2, будем записывать, как обычно: $M(x_1; x_2)$.

Рассмотрим какое-нибудь линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$. Этим преобразованием пара базисных векторов, $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ переводится в некоторую пару векторов $\boldsymbol{e}'_1=A\boldsymbol{e}_1$, $\boldsymbol{e}'_2=A\boldsymbol{e}_2$. Допустим, что векторы $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$ даны, т.е. известны их разложения по базису $\boldsymbol{e}_1$, $\boldsymbol{e}_2$:

$$\left.\begin{aligned} \boldsymbol{e}'_1 = A\boldsymbol{e}_1 = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2, \\ \boldsymbol{e}'_2 = A\boldsymbol{e}_2 = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2. \end{aligned}\right\} \qquad (2)$$

Покажем, что задание формул (2) определяет линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$.

Пусть $\boldsymbol{x} = \{x_1; x_2\}$ — произвольный вектор, $\boldsymbol{x}' = \{x'_1; x'_2\}$ —его образ (координаты этих векторов относятся к базису $\boldsymbol{e}_1$, $\boldsymbol{e}_2$). Мы имеем

$$\boldsymbol{x}' = x'_1\boldsymbol{e}_1 + x'_2\boldsymbol{e}_2 = A\boldsymbol{x} = A(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2).$$

Пользуясь двумя свойствами линейности преобразования, получим

$$\begin{aligned} A(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2) = A(x_1\boldsymbol{e}_1) + A(x_2\boldsymbol{e}_2) = \\ = x_1A\boldsymbol{e}_1 + x_2A\boldsymbol{e}_2 = x_1\boldsymbol{e}'_1 + x_2\boldsymbol{e}'_2. \end{aligned}$$

Следовательно,

$$\boldsymbol{x}' = x'_1\boldsymbol{e}_1 + x'_2\boldsymbol{e}_2 = x_1\boldsymbol{e}'_1 + x_2\boldsymbol{e}'_2.$$

Заменим в правой части последнего равенства векторы $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$ их разложениями (2):

$$\begin{aligned} \boldsymbol{x}' = x'_1\boldsymbol{e}_1 + x'_2\boldsymbol{e}_2 = x_1(a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2) + x_2(a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2) = \\ = (a_{11}x_1 + a_{12}x_2)\boldsymbol{e}_1 + (a_{21}x_1 + a_{22}x_2)\boldsymbol{e}_2. \end{aligned}$$

Отсюда, сравнивая коэффициенты при одинаковых базисных векторах, получаем

$$\left.\begin{aligned} x'_1 = a_{11}x_1 + a_{12}x_2, \\ x'_2 = a_{21}x_1 + a_{22}x_2. \end{aligned}\right\} \qquad (3)$$

Эти формулы позволяют по любому вектору $\boldsymbol{x}=\{x_1;\ x_2\}$ определить его образ $\boldsymbol{x}'=\{x'_1;\ x'_2\}$; коэффициенты этих формул даны, если даны формулы (2). Тем самым показано, что формулы (2) действительно определяют линейное преобразование $\boldsymbol{x}' = A\boldsymbol{x}$, которое переводит базисные векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ в векторы $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$. Тот же результат можно высказать так: *если мы знаем образы базисных векторов при некотором линейном преобразовании, то знаем образы всех векторов при этом преобразовании.*

Равенства (3) линейно выражают координаты вектора $\boldsymbol{x}'=A\boldsymbol{x}$ через координаты вектора x; мы будем называть их *координатным представлением* данного линейного преобразования $\boldsymbol{x}' = A\boldsymbol{x}$ в данном базисе. Таблицу коэффициентов правых частей (3) будем обозначать буквой A и называть *матрицей* данного линейного преобразования в данном базисе

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}; \qquad (4)$$

числа a_{11}, a_{12}, a_{21}, a_{22} называют *элементами* данной матрицы.

Матрицы, с которыми мы сейчас имеем дело, обладают четырьмя элементами, расположенными в двух строчках и в двух столбцах; соответственно такие матрицы называют квадратными матрицами второго порядка.

Заметим, что матрица коэффициентов формул (2) получается *транспонированием* матрицы A, т. е. перестановкой элементов a_{12} и a_{21}, симметричных относительно главной диагонали (эти элементы численно совпадать не обязаны). Результат транспонирования данной матрицы здесь и далее мы будем отмечать звездочкой. Соответственно этому матрица коэффициентов формул (2) запишется так:

$$A^*=\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}. \qquad (5)$$

6. Пусть теперь заранее даны линейные формулы, т. е. формулы вида (3), с любыми коэффициентами a_{11}, a_{12}, a_{21}, a_{22}. Тем самым задано некоторое преобразование векторов на плоскости; именно, произвольному вектору $\boldsymbol{x}=\{x_1;\ x_2\}$ сопоставлен вектор $\boldsymbol{x}'=\{x'_1;\ x'_2\}$, где x'_1 x'_2 определены формулами (3). Обозначим это преобразование символически: $\boldsymbol{x}'=A\boldsymbol{x}$. Докажем, что оно линейно. Для доказательства нужно проверить два условия линейности.

1. Пусть $\boldsymbol{x}=\{x_1;\ x_2\}$, $A\boldsymbol{x}=\boldsymbol{x}'=\{x'_1;\ x'_2\}$,

$$A(\lambda \boldsymbol{x}) = \boldsymbol{x}^* = \{x_1^*;\ x_2^*\}.$$

Мы имеем

$$\lambda \boldsymbol{x} = \{\lambda x_1;\ \lambda x_2\};$$

следовательно,

$$x_1^* = a_{11}(\lambda x_1) + a_{12}(\lambda x_2) = \lambda(a_{11}x_1 + a_{12}x_2) = \lambda x_1',$$
$$x_2^* = a_{21}(\lambda x_1) + a_{22}(\lambda x_2) = \lambda(a_{21}x_1 + a_{22}x_2) = \lambda x_2'.$$

Таким образом, $\boldsymbol{x}^* = \lambda \boldsymbol{x}'$, т. е.

$$A(\lambda \boldsymbol{x}) = \lambda A\boldsymbol{x}.$$

Мы видим, что первое условие соблюдено.

2. Пусть $\boldsymbol{x} = \{x_1;\ x_2\}$, $\boldsymbol{y} = \{y_1;\ y_2\}$— любые векторы,
$\boldsymbol{x}' = A\boldsymbol{x} = \{x_1';\ x_2'\}$, $\boldsymbol{y}' = A\boldsymbol{y} = \{y_1';\ y_2'\}$
— их образы, $A(x+y) = x^* = (x^*_1;\ \ x^*_2)$ — образ суммы векторов $\boldsymbol{x}$ и $\boldsymbol{y}$. Мы имеем

$$\boldsymbol{x} + \boldsymbol{y} = \{x_1 + y_1;\ x_2 + y_2\};$$

следовательно,

$$x_1^* = a_{11}(x_1 + y_1) + a_{12}(x_2 + y_2) =$$
$$= (a_{11}x_1 + a_{12}x_2) + (a_{11}y_1 + a_{12}y_2) = x_1' + y_1',$$
$$x_2^* = a_{21}(x_1 + y_1) + a_{22}(x_2 + y_2) =$$
$$= (a_{21}x_1 + a_{22}x_2) + (a_{21}y_1 + a_{22}y_2) = x_2' + y_2'.$$

Таким образом, $\boldsymbol{x}^* = \boldsymbol{x}' + \boldsymbol{y}'$, т. е.

$$A(\boldsymbol{x} + \boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y}.$$

Второе условие также соблюдено. Тем самым линейность рассматриваемого преобразования доказана. Найдем образы базисных векторов, считая, что линейное преобразование задано формулами (3). Отметим прежде всего очевидные равенства:

$$\boldsymbol{e}_1 = 1 \cdot \boldsymbol{e}_1 + 0 \cdot \boldsymbol{e}_2,$$
$$\boldsymbol{e}_2 = 0 \cdot \boldsymbol{e}_1 + 1 \cdot \boldsymbol{e}_2.$$

Они обозначают, что вектор $\boldsymbol{e}_1$ в базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ имеет координаты 1 и 0, а координатами вектора $\boldsymbol{e}_2$ служат 0 и 1; символически:

$$\boldsymbol{e}_1 = \{1;\ 0\}, \quad \boldsymbol{e}_2 = \{0;\ 1\}.$$

Подставляя в правую часть равенств (3) координаты вектора $\boldsymbol{e}_1$,получим

$$\boldsymbol{e}_1' = \{a_{11};\ a_{21}\}.$$

Аналогично найдем

$$e'_2=\{a_{12};\ a_{22}\}.$$

Мы нашли коэффициенты разложений векторов $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$ по базису $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, следовательно, можем написать и сами разложения:

$$e'_1=a_{11}e_1+a_{21}e_2;$$
$$e'_2=a_{12}e_1+a_{22}e_2.$$

Матрица, составленная из коэффициентов этих разложений, имеет вид (5) и получается транспонированием матрицы (4).

7. Итак, *если дано линейное преобразование* $\boldsymbol{x}'=A\boldsymbol{x}$ *и выбран базис* $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, *то данное преобразование представляется в координатах формулами вида* (3). *Обратно, если заранее даны формулы вида* (3), *то они представляют в выбранных координатах некоторое линейное преобразование* $\boldsymbol{x}'=A\boldsymbol{x}$. *Матрица A, составленная из коэффициентов формул* (3), *и матрица* A^*, *составленная из коэффициентов разложений векторов* $A\boldsymbol{e}_1$, $A\boldsymbol{e}_2$ *по базису* $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ *переводятся одна в другую транспонированием.*

8. Для иллюстрации изложенного мы найдем координатные представления линейных преобразований, которые рассматривались в примерах п° 4.

Пример 1. Пусть $\boldsymbol{x}' = A\boldsymbol{x}$ есть подобие с коэффициентом k, т. е. $A\boldsymbol{x} = k\boldsymbol{x}$.

Позьмем базис $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ произвольно. Имеем

$$\boldsymbol{x}=\{x_1;\ x_2\},\quad \boldsymbol{x}'=\{x'_1;\ x'_2\},\quad \boldsymbol{x}'=k\boldsymbol{x};$$

следовательно,

$$x'_1=kx_1,\qquad x'_2=kx_2.$$

Это и есть координатное представление данного линейного преобразования. Записав полученные формулы в виде

$$x'_1=kx_1+0\cdot x_2,$$
$$x'_2=0\cdot x_1+kx_2,$$

найдем матрицу подобия:
$$A=\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}.$$

Пример 2. Пусть $\boldsymbol{x}' = A\boldsymbol{x}$ есть вращение на угол α.

Возьмем специальный базис $\boldsymbol{i}$, $\boldsymbol{j}$ (состоящий из единичных и взаимно перпендикулярных векторов). Тогда, если $\boldsymbol{x}=\{x_1;\ x_2\}$, то $x_1=\rho\cos\theta,\ \ x_2=\rho\sin\theta,$ где ρ, θ — полярные координаты конца вектора $\boldsymbol{x}$. Так как вектор $\boldsymbol{x}'=A\boldsymbol{x}=\{x'_1;\ x'_2\}$ получается поворотом $\boldsymbol{x}$ вокруг точки O на угол α, то $x'_2=\rho\sin(\theta+\alpha).$

$x_1' = \rho\cos(\theta + \alpha)$

Отсюда

$$x_1' = \rho\cos(\theta+\alpha) = \rho\,[\cos\alpha\cos\theta - \sin\alpha\sin\theta],$$
$$x_2' = \rho\sin(\theta+\alpha) = \rho\,[\sin\alpha\cos\theta + \cos\alpha\sin\theta];$$

раскрывая скобки в правых частях этих равенств и полагая $\rho\cos\theta = x_1$, $\rho\sin\theta = x_2$, найдем

$$x_1' = x_1\cos\alpha - x_2\sin\alpha,$$
$$x_2' = x_1\sin\alpha + x_2\cos\alpha.$$

Это и есть координатное представление вращения в базисе ***i***, ***j***. Матрица вращения имеет вид

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Пример 3. Пусть ***x***′= *A**x*** есть зеркальное отражение относительно прямой *a*. Возьмем специальный базис ***i***, ***j***, причем вектор *i* направим по прямой *a*. Тогда, если $\boldsymbol{x}' = A\boldsymbol{x} = \{x_1';\ x_2'\}$, то $\boldsymbol{x} = \{x_1;\ x_2\}$,

$$x_1' = x_1, \qquad x_2' = -x_2.$$

Это и есть координатное представление в базисе ***i***, ***j*** зеркального отражения относительно оси Ox_1. Матрица этого преобразования имеет вид

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Пример 4. Пусть ***x***′ = *A**x*** есть сжатие к прямой *a* с коэффициентом *k*. Возьмем специальный базис ***i***, ***j***. располагая вектор ***i*** на прямой *a*. Тогда, если

$$\boldsymbol{x} = \{x_1;\ x_2\},\ \boldsymbol{x}' = A\boldsymbol{x} = \{x_1';\ x_2'\},$$

то

$$x_1' = x_1, \qquad x_2' = kx_2.$$

Это и есть координатное представление в базисе ***i***, ***j*** сжатия к оси Ox_1 с коэффициентом *k*. Матрица сжатия к оси Ox_1 имеет вид

$$A = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}.$$

Заметим, что в п°4, где впервые рассматривалось сжатие, линейность этого преобразования не была доказана. Теперь ясно, что сжатие есть линейное преобразование, поскольку его координатное представление имеет линейный вид.

9. В заключение параграфа рассмотрим еще один пример, где линейное преобразование заранее дается его координатным представлением. Пусть

$$\begin{aligned} x_1' &= x_1 + x_2 \\ x_2' &= \phantom{x_1 + {}} x_2, \end{aligned} \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

(базис *i*, *j*).

Если

$$\boldsymbol{x} = \{x_1; \quad x_2\} = \overline{OM}, \ \boldsymbol{x}' = \{x_1';\ x_2'\} = \overline{OM'},$$

то M' получается из M перемещением параллельно оси Ox_1 на величину x_2 (точки оси Ox_1 остаются на своих местах). Такое преобразование мы назовем перекосом плоскости (см. рис. 8, где пунктирные стрелки показывают смещение точек прямой MN).

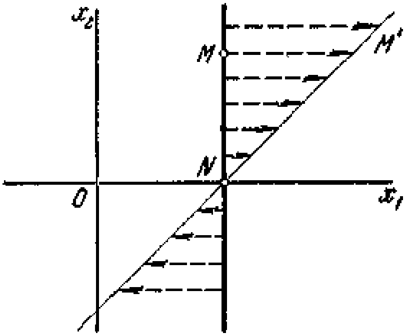


Рис. 8

6.2. Произведение линейных преобразований на плоскости и произведение квадратных матриц второго порядка.

10. Пусть даны две матрицы:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Выберем накой-нибудь базис $\boldsymbol{e}_1$, $\boldsymbol{e}_2$. Тогда, если мы напишем по данной матрице A формулы (3) п.6.1, то тем самым будет определено некоторое линейное преобразование *Ax;* аналогично по данной матрице *B* в том же базисе определим линейное преобразование ***Bx***.

По двум линейным преобразованиям *A**x*** и *B**x*** мы установим теперь некоторое новое линейное преобразование. Именно, возьмем произвольный вектор $\boldsymbol{x}=\{x_1;\ x_2\}$;преобразование с матрицей *B* переводит его в вектор, который мы обозначим через

$$\boldsymbol{y}:\ \boldsymbol{y}=B\boldsymbol{x}=\{y_1;\ y_2\};$$

полученный вектор другим данным преобразованием будет переведен в некоторый вектор $\boldsymbol{x}':\ \boldsymbol{x}'=A\boldsymbol{y}=\{x_1';\ x_2'\}$. В результате каждому вектору ***x*** сопоставлен вектор ***x***′, т. е. установлено определенное преобразование векторов. Это преобразование называется *произведением* двух данных преобразований и обозначается так: *AB**x***. Все приведенные сейчас словесные описания можно заменить одной формулой:

$$AB\boldsymbol{x}=A(B\boldsymbol{x}). \tag{1}$$

Пользуясь формулой (1), легко показать, что произведение линейных преобразований есть также линейное преобразование. Однако мы предпочтем установить тот же факт иным способом — при помощи координатных представлений данных линейных преобразований; одновременно мы найдем матрицу их произведения.

Так как ***y*** = *B**x***, то в координатах имеем

$$\left.\begin{aligned} y_1&=b_{11}x_1+b_{12}x_2,\\ y_2&=b_{21}x_1+b_{22}x_2.\end{aligned}\right\} \tag{2}$$

Так как ***x***′ =*A**y***, то

$$\left.\begin{aligned} x_1'&=a_{11}y_1+a_{12}y_2,\\ x_2'&=a_{21}y_1+a_{22}y_2.\end{aligned}\right\} \tag{3}$$

Заменяя в последних формулах y_1, y_2, правыми частями (2), найдем

$$\left.\begin{aligned} x_1'&=(a_{11}b_{11}+a_{12}b_{21})\,x_1+(a_{11}b_{12}+a_{12}b_{22})\,x_2,\\ x_2'&=(a_{21}b_{11}+a_{22}b_{21})\,x_1+(a_{21}b_{12}+a_{22}b_{22})\,x_2.\end{aligned}\right\} \tag{4}$$

Мы получили координатное представление преобразования ***x***′=*AB**x*** (в данном базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$). Поскольку формулы (4) имеют линейный вид, представляемое ими преобразование векторов является линейным (см. п.6.1). Тем самым доказано, что произведение двух линейных преобразований также есть линейное преобразование.

11. *Матрица произведения двух линейных преобразований называется произведением матриц этих преобразований.* Если данные преобразования суть *A**x*** и *B**x***, то их произведению *AB**x***

соответствует произведение матриц A и B, которое обозначает AB. Согласно (4)

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}, \qquad (5)$$

или

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}. \qquad (6)$$

Рассмотрим какой-нибудь элемент произведения данных матриц, например $a_{11}b_{12}+a_{12}b_{22}$. Этот элемент находится на пересечении первой строки и второго столбца. Заметим, что в нем участвуют элементы a_{11}, a_{12} первой строки матрицы A и элементы b_{12}, b_{22} второго столбца матрицы B, причем элементы строки умножаются на соответствующие (по порядку) элементы столбца и полученное произведение суммируется. Величина, которая составляется по указанному сейчас правилу, называется произведением строки на столбец; в частности, $a_{11}b_{12}+a_{12}b_{22}$ есть произведение первой строки матрицы A на второй столбец матрицы B. Из (6) видно, что вообще *для того, чтобы получить элемент произведения AB, занимающий место в i-й строке и в k-м столбце* (i = 1 *или* 2, k = 1 *или* 2), *нужно помножить i -ю строку матрицы A на k-й столбец матрицы B.*

12. Весьма важно учесть, что произведение линейных преобразований зависит от того, в каком порядке они последовательно производятся; порядок выполнения данных преобразований отражается в записи произведения. Если мы пишем $AB\boldsymbol{x} = A(B\boldsymbol{x})$, то имеем в виду, что сначала вектор $\boldsymbol{x}$ подвергается преобразованию B, а затем его образ преобразуется при помощи A. Если $\boldsymbol{x}$ сначала подвергнуть преобразованию A, а затем полученный вектор $A\boldsymbol{x}$ преобразовать при помощи B, то получим произведение $BA\boldsymbol{x}= B(A\boldsymbol{x})$.

Вообще говоря, $AB\boldsymbol{x} \neq BA\boldsymbol{x}$ (хотя для некоторых преобразований A и B возможно равенство $AB\boldsymbol{x} = BA\boldsymbol{x}$ при любом $\boldsymbol{x}$).

Точно так же зависит от порядка сомножителей и произведение матриц, т. е., вообще говоря, матрица AB не совпадает с матрицей BA (хотя в частных случаях совпадение AB и BA может иметь место). То обстоятельство, что произведение матриц зависит от порядка сомножителей, легко понять также, если обдумать правило перемножения матриц. Дело в том, что когда мы находим элемент произведения двух матриц, то умножаем строку левого сомножителя на столбец правого; тем самым левый и правый сомножители неравноправны.

Пример. Пусть

$$A=\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}, \qquad B=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

Тогда

$$AB=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \quad BA=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}.$$

Матрицы *AB* и *BA* оказались различными.

13. Рассмотрим еще два примера.

Пример 1. Пусть $A\mathbf{x}$—вращение на угол α, $B\mathbf{x}$ — вращение на угол β (см. п° 4 и 8). Очевидно, $AB\mathbf{x}$ есть вращение на угол α + β; в данном случае $AB\mathbf{x} = BA\mathbf{x}$.

Возьмем специальный базис ***i***, ***j***; тогда данные преобразования будут иметь соответственно матрицы

$$A=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}, \qquad B=\begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}.$$

Применяя правило умножения матриц, получим

$$AB=\begin{pmatrix}\cos\alpha\cos\beta-\sin\alpha\sin\beta & -\cos\alpha\sin\beta-\sin\alpha\cos\beta\\ \sin\alpha\cos\beta+\cos\alpha\sin\beta & -\sin\alpha\sin\beta+\cos\alpha\cos\beta\end{pmatrix}.$$

Отсюда

$$AB=\begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix}.$$

Этот результат можно было заранее предвидеть, так как *AB* есть матрица вращения на угол α + β. В данном случае матрицы *AB* и *BA* совпадают.

Пример 2. Пусть $A\mathbf{x}$ есть сжатие к оси Ox_2 (т. е. в направлении оси Ox_1) с коэффициентом k_1, $B\mathbf{x}$ — сжатие к оси Ox_2 (т. е. в направлении оси Ox_2) с коэффициенюм k_2 (см. п°4 и 8). Матрицы этих преобразований соответственно будут:

$$A=\begin{pmatrix}k_1 & 0\\ 0 & 1\end{pmatrix}, \qquad B=\begin{pmatrix}1 & 0\\ 0 & k_2\end{pmatrix};$$

умножая *A* на *B*, получим

$$AB=\begin{pmatrix}k_1 & 0\\ 0 & k_2\end{pmatrix}.$$

Матрица такого вида называется *диагональной.* Таким образом, диагональная матрица отвечает произведению двух сжатий к координатным осям.

14. Далее нам придется писать равенства матриц. Равенство двух матриц означает, что их элементы, занимающие одинаковые места, численно совпадают. Иначе говоря, если $A = B$, то A и B являются совершенно одинаковыми таблицами чисел.

15. Пусть даны три матрицы A, B, C. Умножая B на С, мы получим матрицу BC, на которую затем можем умножить A; в результате будем иметь матрицу $A(BC)$. Можно поступить иначе: произведение AB умножить на матрицу C; тогда получим матрицу $(AB)C$. Легко убедиться, что

$$A\,(BC) = (AB)\,C. \qquad (7)$$

Для доказательства используем линейные преобразования $A\mathbf{x}$, $B\mathbf{x}$ и $C\mathbf{x}$, которые имеют заданные матрицы A, B и C (в каком-нибудь базисе). Рассмотрим произведение преобразований: $A(BC)\mathbf{x}$. Это есть преобразование, которое производится так: сначала $\mathbf{x}$ подвергается преобразованию C, затем его образ преобразуется с помощью B; полученный в результате вектор дополнительно подвергается преобразованию A. Символически имеем:

$$A(BC)x = A(B(Cx)).$$

Но совершенно так же производится преобразование, котрое есть произведение $AB\mathbf{x}$ на Cx:

$$(AB)C\mathbf{x} = A(B(C\mathbf{x})).$$

Следовательно, преобразования $A(BC)\mathbf{x}$ и $(AB)\ C\mathbf{x}$ не отличаются друг от друга. Отсюда получаем равенство матриц (7). Поскольку в произведении трех матриц A, B, C сочетание сомножителей безразлично, его обозначают так: ABC, считая $ABC = A(BC)$ или, если угодно, $ABC=(AB)C$. Нельзя, однако, нарушать заданный порядок сомножителей, так как уже произведение двух матриц зависит от того, какая из них является левым сомножителем, какая правым; например, ABC, вообще говоря, не совпадает с ACB.

Произведение любого числа матриц определяется аналогично произведению трех матриц последовательного перемножения данных сомножителей; например, $ABCD = A(B(CD))$

16. Натуральные степени матрицы определяются так же, как степени чисел:

$$A^2=A\cdot A, \quad A^3 = A\cdot A\cdot A,$$

например, если

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

то

$$A^2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}.$$

17. В некоторых вопросах математики и ее приложений встречается сумма и разность матриц и произведение матрицы на число. Суммой двух матриц *A* и *B* называется матрица, элементы которой получаются путем сложения соответствующих элементов данных матриц *A* и *B:*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}.$$

Аналогично определяется разность *A — B*. Произведением матрицы *A* на число λ называется матрица, элементы которой получаются умножением элементов матрицы *A* на число λ*:*

$$\lambda A = \lambda \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix}.$$

После всего сказанного легко понимать смысл более сложных выражений, составленных из матриц сочетанием различных действий; например, если

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

то

$$A^2 + 5A = \begin{pmatrix} 12 & 20 \\ 30 & 42 \end{pmatrix}.$$

18. В заключение этого параграфа укажем правило сочетания действий умножения матриц и транспонирования. Напомним читателю, что действие транспонирования мы обозначаем звездочкой: если *A*—данная матрица, то *A** — матрица, получаемая транспонированием *A* (для матриц второго порядка транспонирование сводится к перемене местами правого верхнего и левого нижнего элементов).

Имеет место следующее тождество:

$$(AB)^* = B^*A^*, \tag{8}$$

т. е. *транспонированное произведение равно произведению транспонированных сомножителей, взятых в обратном порядке.* Для доказательства тождества (8) достаточно заметить, что строчки данной матрицы являются столбцами матрицы, получаемой ее транспонированием. Поэтому элемент произведения B^*A^*, занимающий место в i-й строчке и в k-м. столбце, получается умножением i-го столбца матрицы B на k-ю строчку матрицы A. Но точно такой же элемент получается в произведении AB на пересечении k -й строчки и i -го столбца. Следовательно, AB после транспонирования совпадает с B^*A^*.

6.3. Теорема об определителе произведения двух матриц

19. Условимся сейчас и в дальнейшем определитель данной квадратной матрицы A обозначать символом Det A (определитель называют также детерминантом; отсюда происходит обозначение определителя символом Det):

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \operatorname{Det} A=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

20. Имеет место следующая

Теорема. *Определитель произведения двух матриц равен произведению определителей этих матриц;* символически:

$$\operatorname{Det}(AB)=\operatorname{Det} A\cdot\operatorname{Det} B. \qquad (1)$$

Доказательство. Пусть

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix};$$

тогда

$$AB=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}. \qquad (2)$$

Введем обозначения: $\Delta_1=\operatorname{Det} A$, $\Delta_2=\operatorname{Det} B$, $\Delta=\operatorname{Det}(AB)$.
Из (2) имеем:

$$\Delta = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{vmatrix} + \begin{vmatrix} a_{11}b_{11} & a_{12}b_{22} \\ a_{21}b_{11} & a_{22}b_{22} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{12}b_{21} & a_{11}b_{12} \\ a_{22}b_{21} & a_{21}b_{12} \end{vmatrix} + \begin{vmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{vmatrix} =$$

$$= b_{11}b_{12} \begin{vmatrix} a_{11} & a_{11} \\ a_{21} & a_{21} \end{vmatrix} + b_{11}b_{22} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} +$$

$$+ b_{21}b_{12} \begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} + b_{21}b_{22} \begin{vmatrix} a_{12} & a_{12} \\ a_{22} & a_{22} \end{vmatrix}.$$

В полученной сумме участвуют четыре определителя, составленные из элементов матрицы *А*. Очевидно, первый и последний из них равны нулю; определитель, занимающий второе место, есть Δ_1; на третьем месте стоит определитель, равный определителю Δ_1 с обратным знаком. Таким образом,

$$\Delta = b_{11}b_{22}\Delta_1 - b_{21}b_{12}\Delta_1 = (b_{11}b_{22} - b_{21}b_{12})\,\Delta_1 = \Delta_1\Delta_2.$$

Теорема доказана. Эту теорему называют также теоремой об умножении определителей.

6.4. Вырожденные преобразования

21. Пусть ***x, y*** — два произвольных вектора, заданных своими разложениями по некоторому базису $\boldsymbol{e}_1, \boldsymbol{e}_2$:

$$\left.\begin{aligned} \boldsymbol{x} &= x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2, \\ \boldsymbol{y} &= y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2; \end{aligned}\right\} \qquad (1)$$

как и раньше, мы рассматриваем векторы, лежащие в определенной плоскости α.

Обозначим через S_0 площадь параллелограмма, построенного на базисных векторах $\boldsymbol{e}_1, \boldsymbol{e}_2$, через $\boldsymbol{n}$ — единичный вектор, нормальный к плоскости α и направленный так же, как векторное произведение $[\boldsymbol{e}_1, \boldsymbol{e}_2]$; тогда

$$[\boldsymbol{e}_1\boldsymbol{e}_2] = S_0\boldsymbol{n}. \qquad (2)$$

Заметим, что если все векторы приведены к общему началу *O* и если смотреть из конца вектора $\boldsymbol{n}$ на плоскость α, то кратчайший поворот вектора $\boldsymbol{e}_1$ к вектору $\boldsymbol{e}_2$ будет виден совершающимся против часовой стрелки.

Обозначим, далее, через S площадь параллелограмма, построенного на векторах $\boldsymbol{x}, \boldsymbol{y}$, тогда

$$[\boldsymbol{x}, \boldsymbol{y}] = \pm S\boldsymbol{n}, \qquad (3)$$

причем знак плюс будет в том случае, когда кратчайший поворот вектора $\boldsymbol{x}$ к вектору $\boldsymbol{y}$ виден из конца $\boldsymbol{n}$ совершающимся против часовой стрелки, знак минус — если этот поворот будем по часовой стрелке. В первом случае мы скажем, что пара векторов $\boldsymbol{x}, \boldsymbol{y}$ (считая $\boldsymbol{x}$ первым вектором) ориентирована так же, как пара векторов $\boldsymbol{e}_1, \boldsymbol{e}_2$, во втором — что пары $\boldsymbol{x}, \boldsymbol{y}$ и $\boldsymbol{e}_1, \boldsymbol{e}_2$ ориентированы различно. Если векторы $\boldsymbol{x}, \boldsymbol{y}$ коллинеарны, то $S = 0$ и вопрос о знаке в правой части формулы (3) отпадает.

Найдем выражение S, считая известными S_0 и коэффициенты формул (1). Мы имеем:

$$[\boldsymbol{x}, \boldsymbol{y}] = [(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2)(y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2)] =$$
$$= x_1y_1[\boldsymbol{e}_1\boldsymbol{e}_1] + x_1y_2[\boldsymbol{e}_1\boldsymbol{e}_2] + x_2y_1[\boldsymbol{e}_2\boldsymbol{e}_1] + x_2y_2[\boldsymbol{e}_2\boldsymbol{e}_2] =$$
$$= (x_1y_2 - x_2y_1)[\boldsymbol{e}_1\boldsymbol{e}_2].$$

Отсюда с учетом равенств (2) и (3) получаем

$$\pm S = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} S_0.$$

Тем самым искомое выражение S найдено.

Условимся называть ориентированной площадью параллелограмма, построенного на паре векторов $\boldsymbol{x}, \boldsymbol{y}$, число $+S$, если пары $\boldsymbol{x}, \boldsymbol{y}$ и $\boldsymbol{e}_1, \boldsymbol{e}_2$ ориентированы одинаково, число — S, если эти пары ориентированы различно (вектор $\boldsymbol{x}$ считаем первым в паре $\boldsymbol{x}, \boldsymbol{y}$); ориентированную площадь будем обозначать буквой σ. Из предыдущего равенства имеем

$$\sigma = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} S_0. \qquad (4)$$

22. Рассмотрим произвольное линейное преобразование $\boldsymbol{x}' = A\boldsymbol{x}$; пусть

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2, \\ x_2' &= a_{21}x_1 + a_{22}x_2 \end{aligned}\right\} \qquad (5)$$

— координатное представление этого преобразования в данном базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$.

Возьмем произвольную пару векторов $\boldsymbol{x} = \{x_1; \quad x_2\}$,

$\boldsymbol{y}=\{y_1;\ y_2\}$. Обозначим через σ ориентированную площадь параллелограмма, построенного на паре векторов ***x, y,*** через σ' — ориентированную площадь параллелограмма, построенного на паре их образов $\boldsymbol{x}'=A\boldsymbol{x}$, $\boldsymbol{y}'=A\boldsymbol{y}$. Найдем зависимость между σ и σ'. На основании формулы (4) имеем

$$\sigma=\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} S_0. \tag{6}$$

Если $\boldsymbol{x}'=\{x_1';\ x_2'\}$, $\boldsymbol{y}'=\{y_1';\ y_2'\}$, то

$$\sigma'=\begin{vmatrix} x_1' & y_1' \\ x_2' & y_2' \end{vmatrix} S_0. \tag{7}$$

Согласно (5)

$$x_1'=a_{11}x_1+a_{12}x_2,\quad y_1'=a_{11}y_1+a_{12}y_2,$$
$$x_2'=a_{21}x_1+a_{22}x_2,\quad y_2'=a_{21}y_1+a_{22}y_2.$$

Отсюда и по правилу умножения матриц получаем:

$$\begin{pmatrix} x_1' & y_1' \\ x_2' & y_2' \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$

Из этого матричного равенства, применяя теорему об умножении определителей, найдем соответствующее равенство для определителей:

$$\begin{vmatrix} x_1' & y_1' \\ x_2' & y_2' \end{vmatrix}=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\cdot\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

Умножим обе части последнего равенства на S_0 и учтем (6) и (7); мы получим искомое соотношение:

$$\sigma'=\sigma\,\mathrm{Det}\,A. \tag{8}$$

Таким образом, *при линейном преобразовании плоскости все параллелограммы деформируются так, что их ориентированные площади изменяются пропорционально; общим коэффициентом изменения площадей является определитель преобразования.* Из формулы (8) видно также следующее: если Det $A > 0$, то пара векторов ***x, y*** и пара их образов ориентированы одинаково (σ и σ′ имеют один и тот же знак); если Det $A<0$, то пары ***x, y*** и ***x′, y′*** ориентированы различно. Иначе говоря, если Det $A > 0$, то преобразование сохраняет ориентации всех пар векторов данной плоскости, если Det $A<0$ — меняет на противоположные.

23. Рассмотрим теперь линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ при условии Det $A = 0$. В этом случае формула (8) дает $\sigma' = 0$ при любом значении σ.

Следовательно, данное преобразование любую пару векторов $\boldsymbol{x}$, $\boldsymbol{y}$ переводит в пару коллинеарных векторов $\boldsymbol{x}'$, $\boldsymbol{y}'$. В частности, будут коллинеарны образы $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$ базисных векторов $\boldsymbol{e}_1$, $\boldsymbol{e}_2$. Обозначим через a' прямую, проходящую через точку O, на которой лежат векторы $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$. Вернемся к п°5, где показано, что образ $\boldsymbol{x}'=A\boldsymbol{x}$ любого вектора $\boldsymbol{x}=\{x_1;\ \ x_2\}$ может быть выражен в виде

$$\boldsymbol{x}'=x_1\boldsymbol{e}'_1+x_2\boldsymbol{e}'_2.$$

Отсюда заключаем: если Det A = 0, то образы всех векторов данной плоскости располагаются на одной прямой a'; можно сказать также, что образы всех точек данной плоскости лежат на одной прямой, т. е. вся плоскость отображается в некоторую прямую. Такое линейное преобразование плоскости называется *вырожденным;* согласно изложенному *вырожденное линейное преобразование характеризуется условием* Det A = 0. При этом же условии называется *вырожденной* и *матрица* A.

Пример. Линейное преобразование

$$x'_1=x_1+x_2,$$
$$x'_2=x_1+x_2$$

является вырожденным, так как образы всех точек лежат на прямой $x'_1 = x'_2$. Вместе с тем имеем

$$A=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$$

и Det A=0.

6.4. Обращение линейного преобразования на плоскости

24. Рассмотрим на данной плоскости линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$. Здесь $\boldsymbol{x}'$ — *образ* вектора $\boldsymbol{x}$; в таком случае $\boldsymbol{x}$ называется *прообразом* вектора $\boldsymbol{x}'$. Поставим задачу: по любому вектору $\boldsymbol{x}'$ найти его прообраз $\boldsymbol{x}$. Если будет указано правило, согласно которому по каждому вектору $\boldsymbol{x}'$ находится его прообраз $\boldsymbol{x}$, то тем самым на плоскости будет установлено некоторое новое преобразование, которое с вектором $\boldsymbol{x}'$ сопоставляет вектор $\boldsymbol{x}$. Такое преобразование называют *обратным данному* и обозначают через $\boldsymbol{x}= A^{-1}\boldsymbol{x}'$.

25. Предположим, что данное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ является вырожденным; тогда образы всех векторов плоскости лежат на некоторой прямой a'. Если мы возьмем вектор $\boldsymbol{x}'$, не лежащий на этой прямой, то для него нет прообраза. Поэтому *для вырожденного преобразования нет обратного.*

26. Пусть теперь $\boldsymbol{x}'=A\boldsymbol{x}$ — невырожденное линейное преобразование,

$$x_1' = a_{11}x_1 + a_{12}x_2, \quad x_2' = a_{21}x_1 + a_{22}x_2 \qquad (1)$$

—его координатное представление в некотором базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad (2)$$

—матрица данного преобразования. Так как преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ предполагается невырожденным, то

$$\Delta = \operatorname{Det} A \neq 0.$$

В равенствах (1) величины x'_1, x'_2 являются координатами образа, величины x_1, x_2 — координатами его прообраза. Чтобы получить преобразование, обратное данному, нужно выразить x_1, x_2 через x'_1, x'_2 т. е. решить систему (1) относительно x_1, x_2, считая x'_1, x'_2 известными. Поскольку $\Delta\neq 0$, мы получаем

$$x_1 = \frac{\begin{vmatrix} x_1' & a_{12} \\ x_2' & a_{22} \end{vmatrix}}{\Delta}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & x_1' \\ a_{21} & x_2' \end{vmatrix}}{\Delta}.$$

Отсюда

$$x_1 = \frac{a_{22}}{\Delta} x_1' - \frac{a_{12}}{\Delta} x_2', \quad x_2 = -\frac{a_{21}}{\Delta} x_1' + \frac{a_{11}}{\Delta} x_2'. \qquad (3)$$

Мы получили координатное представление преобразования $\boldsymbol{x} = A^{-1}\boldsymbol{x}'$, обратного данному линейному преобразованию $\boldsymbol{x}'=A\boldsymbol{x}$ (в том же базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$). Так как выражения (3) имеют линейный вид, заключаем, что *обратное преобразование также является линейным.* Матрицу обратного преобразования $A^{-1}\boldsymbol{x}'$ обозначают через A^{-1} и называют *обратной матрицей* для данной матрицы A. Из (3) имеем

$$A^{-1} = \begin{pmatrix} \frac{a_{22}}{\Delta} & -\frac{a_{12}}{\Delta} \\ -\frac{a_{21}}{\Delta} & \frac{a_{11}}{\Delta} \end{pmatrix}. \qquad (4)$$

Итак, *для каждого невырожденного линейного преобразования имеется обратное преобразование, которое также является линейным;* если матрица A данного преобразования известна, то матрица A^{-1} обратного преобразования (обратная матрица) дается равенством (4).

27. К числу линейных относится следующее весьма специальное преобразование: $x'=x$; в данном случае образ каждого вектора $\boldsymbol{x}$ совпадает с самим вектором $\boldsymbol{x}$. Такое преобразование называется *тождественным.* Координатное представление тождественного преобразования в любом базисе имеет вид

$$x_1' = x_1, \quad x_2' = x_2.$$

Матрица тождественного преобразования обозначается буквой E и называется *единичной;* очевидно,

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

28. Пусть $\boldsymbol{x}' = A\boldsymbol{x}$ — любое невырожденное линейное преобразование, $\boldsymbol{x} = A^{-1}\boldsymbol{x}'$ — обратное ему. Рассмотрим произведение этих двух преобразований: имеем $A^{-1}A\boldsymbol{x}=\boldsymbol{x}$ (сначала преобразование A переводит вектор $\boldsymbol{x}$ в его образ $\boldsymbol{x}'$, затем обратное преобразование A^{-1} по вектору $\boldsymbol{x}'$ снова дает его прообраз $\boldsymbol{x}$). Следовательно, *произведение данного и обратного преобразования есть тождественное преобразование.* Отсюда

$$A^{-1}A = E, \tag{5}$$

т. е. произведение данной и обратной матрицы есть единичная матрица.

Если взять произведение данного и обратного преобразования в другом порядке, то мы также получим тождественное преобразование $AA^{-1}\boldsymbol{x}' = \boldsymbol{x}'$; поэтому наряду с равенством (5) имеет место аналогичное ему равенство

$$AA^{-1}=E. \tag{6}$$

29. Заметим еще, что единичная матрица играет в теории матриц такую же роль, какую в обычной арифметике играет единица. Именно, умножение любой данной матрицы на единичную не изменяет данной матрицы:

$$AE = A, \quad EA = A; \tag{7}$$

в справедливости этих равенств легко убедиться, вычисляя AE (или EA) по правилу умножения матриц.

30. **Пример.** Дано координатное представление линейного преобразования:

$$x_1' = 3x_1 + 2x_2,$$
$$x_2' = 7x_1 + 5x_2,$$

найти обратное преобразование.

Решение. Имеем

$$A=\begin{pmatrix}3 & 2\\ 7 & 5\end{pmatrix},$$

Det A = 1≠0. Матрица A не вырождена, следовательно, имеет обратную. Согласно (4) находим

$$A^{-1}=\begin{pmatrix}5 & -2\\ -7 & 3\end{pmatrix}.$$

Отсюда получаем координатное представление обратного преобразования:

$$x_1=5x_1'-2x_2',$$
$$x_2=-7x_1'+3x_2'.$$

6.5. Преобразование координат векторов при переходе к новому базису

31. Выведем формулы, по которым преобразуются координаты произвольного вектора плоскости при переходе к новому базису.

Пусть $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ — первоначально выбранный базис, $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$ — новый базис. Будем считать, что нам известны коэффициенты разложений векторов $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$ по старому базису $\boldsymbol{e}_1$, $\boldsymbol{e}_2$:

$$\left.\begin{aligned}\tilde{\boldsymbol{e}}_1&=l_{11}\boldsymbol{e}_1+l_{21}\boldsymbol{e}_2,\\ \tilde{\boldsymbol{e}}_2&=l_{12}\boldsymbol{e}_1+l_{22}\boldsymbol{e}_2.\end{aligned}\right\} \tag{1}$$

Матрицу коэффициентов правых частей (1) обозначим через L^*:

$$L^*=\begin{pmatrix}l_{11} & l_{21}\\ l_{12} & l_{22}\end{pmatrix}. \tag{2}$$

Рассмотрим в данной плоскости произвольный вектор $\boldsymbol{x}$; координаты x_1, x_2 вектора $\boldsymbol{x}$ по старому базису определяются как коэффициенты разложения $\boldsymbol{x}$ по этому базису:

$$\boldsymbol{x}=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2.$$

Аналогично, координаты $\tilde{x}_1$, $\tilde{x}_2$ того же вектора $\boldsymbol{x}$ по новому базису даются разложением

$$\boldsymbol{x}=\tilde{x}_1\tilde{\boldsymbol{e}}_1+\tilde{x}_2\tilde{\boldsymbol{e}}_2.$$

Отсюда

$$x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2=\tilde{x}_1\tilde{\boldsymbol{e}}_1+\tilde{x}_2\tilde{\boldsymbol{e}}_2. \tag{3}$$

Заменим в правой части равенства (3) векторы $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2$ по формулам (1); после приведения подобных членов мы получим

$$x_1 e_1 + x_2 e_2 = (l_{11}\tilde{x}_1 + l_{12}\tilde{x}_2)\, e_1 + (l_{21}\tilde{x}_1 + l_{22}\tilde{x}_2)\, e_2.$$

Сравнивая коэффициенты, стоящие в этом равенстве слева и справа при одинаковых базисных векторах, найдем

$$\left.\begin{aligned} x_1 &= l_{11}\tilde{x}_1 + l_{12}\tilde{x}_2, \\ x_2 &= l_{21}\tilde{x}_1 + l_{22}\tilde{x}_2. \end{aligned}\right\} \tag{4}$$

Это и есть искомые формулы преобразования координат векторов. Матрицу коэффициентов формул (3) мы обозначим буквой L:

$$L = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix}. \tag{5}$$

Очевидно, L и L^* преобразуются друг в друга транспонированием (что и выражено обозначением *).

Учтем, что базисные векторы $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2$ (как и $\boldsymbol{e}_1, \boldsymbol{e}_2$) неколлинеарны; отсюда следует, что определитель матрицы L^* не равен нулю (см. п° 21). Но

$$\operatorname{Det} L^* = \operatorname{Det} L.$$

Следовательно,

$$\operatorname{Det} L \neq 0.$$

Итак, если новый базис задан формулами (1), то старые координаты любого вектора выражаются через его новые координаты по формулам (4); матрицы коэффициентов формул (1) и (4) переходят друг в друга при помощи транспонирования и обе являются невырожденными.

32. Если $\boldsymbol{x} = \overline{OM}$, то координаты вектора $\boldsymbol{x}$ совпадают с координатами точки M. Поэтому формулы (4) являются также формулами преобразования координат точек при переходе к новому базису (без перемены начала координат).

33. Рассмотрим частный случай преобразования координат, когда: 1) старый базис состоит из единичных и перпендикулярных друг к другу векторов $\boldsymbol{i}, \boldsymbol{j}$; 2) новый базис $\tilde{\boldsymbol{i}}, \tilde{\boldsymbol{j}}$ имеет аналогичный вид. В этом случае мы изменим обозначения коэффициентов формул (1) и напишем эти формулы так:

$$\begin{aligned} \tilde{\boldsymbol{i}} &= l_1 \boldsymbol{i} + m_1 \boldsymbol{j}, \\ \tilde{\boldsymbol{j}} &= l_2 \boldsymbol{i} + m_2 \boldsymbol{j}; \end{aligned}$$

соответственно

$$L^* = \begin{pmatrix} l_1 & m_1 \\ l_2 & m_2 \end{pmatrix}, \qquad L = \begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix}.$$

Так как $\tilde{i}^2 = 1, \ \tilde{j}^2 = 1, \ \tilde{i}\tilde{j} = 0$, то

$$l_1^2 + m_1^2 = 1, \quad l_2^2 + m_2^2 = 1, \quad l_1 l_2 + m_1 m_2 = 0. \qquad (6)$$

Вследствие равенств (6) имеем по правилу умножения матриц:

$$\begin{pmatrix} l_1 & m_1 \\ l_2 & m_2 \end{pmatrix} \begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \qquad (7)$$

обратно, из (7) следует (6). Равенство (7) допускает краткую запись:

$$L^* L = E; \qquad (8)$$

отсюда

$$L^* = L^{-1}. \qquad (9)$$

Матрица *L*, удовлетворяющая условию (8), называется *ортогональной.* Согласно (9) можно сказать также, что *ортогональная матрица L характеризуется тем, что для нее обратная матрица совпадает с транспонированной.*

На основании изложенного в этом пункте заключаем: при переходе от базиса ***i, j*** к аналогичному базису $\tilde{i}, \tilde{j}$ получаются формулы преобразования координат, матрица которых является ортогональной. Заметим еще, что из (8) следует числовое равенство"

$$\operatorname{Det} L^* \cdot \operatorname{Det} L = 1,$$

откуда $(\operatorname{Det} L)^2 = 1$. Следовательно,

$$\operatorname{Det} L = \pm 1, \qquad (10)$$

т. е. определитель ортогональной матрицы может быть равен только - + 1 или — 1.

В первом случае базисы i, j и $\tilde{i}, \tilde{j}$ ориентированы одинаково, во втором — различно (см. п. 21, формула (4))

34. Вернемся к общему случаю преобразования координат.

Если сравнить формулы (4) с формулами (3) п.6.1, то легко усмотреть их полную алгебраическую аналогию. Между тем геометрическое содержание этих формул разное.

Действительно, в то время как формулы (3) п.6.1 выражают координаты образа $\boldsymbol{x}'$ через координаты прообраза $\boldsymbol{x}$ (в одном и том же базисе), формулы (4) выражают старые координаты вектора $\boldsymbol{x}$ через новые координаты того же самого вектора $\boldsymbol{x}$ (при изменении базиса).

Представляется целесообразным установить единые правила краткой записи такого рода формул, не считаясь с их геометрическим смыслом. С этой целью рассмотрим любые соотношения вида

$$\left.\begin{aligned} y_1 = b_{11}x_1 + b_{12}x_2, \\ y_2 = b_{21}x_1 + b_{22}x_2, \end{aligned}\right\} \qquad (11)$$

полагая, что матрица их коэффициентов задана:

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Формулы (11) можно заменить одним символическим равенством

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad (12)$$

понимая его правую часть как произведение квадратной матрицы B на матрицу, состоящую из одного столбца; само такое произведение (левая часть равенства (12)) также есть матрица, состоящая из одного столбца. Равенство (12) расшифровывается так: чтобы получить y_1, нужно первую строчку матрицы B умножить на столбец из пары чисел x_1, x_2; чтобы получить y_2, нужно умножить на этот столбец вторую строчку матрицы B (см. формулы (11)). Условимся теперь пару чисел x_1, x_2 обозначать одной буквой X, а пару чисел y_1, y_2 одной буквой Y. После этого равенство (12) принимает совсем компактный вид:

$$Y = BX. \qquad (13)$$

Можно сказать, что пара чисел Y получается путем умножения матрицы B на пару чисел X (или пары чисел X на матрицу B). Допустим, что пара чисел Y в свою очередь умножается на матрицу A и дает новую пару чисел X':

$$X' = AY.$$

Тогда

$$X' = ABX,$$

т. е. X' можно получить непосредственно, умножая произведение матриц AB на X. Чтобы убедиться в этом, достаточно вернуться к произведению линейных преобразований; то, что мы сейчас рассматривали, совершенно не отличается от произведения линейных преобразований, заданных координатными представлениями (см. формулы (2), (3), (4) в п° 10).

Применяя введенную символику, мы запишем формулы (4) так:

$$X = L\tilde{X}, \qquad (14)$$

где X — пара старых координат вектора, $\tilde{X}$—пара его новых координат. Отсюда

$$\tilde{X} = L^{-1}X; \tag{15}$$

это значит, что пара новых координат вектора получается из пары его старых координат умножением на матрицу L^{-1}. В частном случае, когда L — ортогональная матрица, имеем

$$X = L\tilde{X}, \quad \tilde{X} = L^{*}X. \tag{16}$$

6.6. Изменение матрицы линейного преобразования на плоскости при переходе к новому базису

35. Координатное представление данного линейного преобразования векторов и матрица этого преобразования зависят от выбора базиса. Мы установим сейчас закон изменения матрицы линейного преобразования, если выбранный базис заменяется другим.

36. Пусть дано линейное преобразование векторов на плоскости: $\boldsymbol{x}' = A\boldsymbol{x}$; пусть

$$\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2, \\ x'_2 &= a_{21}x_1 + a_{22}x_2, \end{aligned} \qquad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{1}$$

— координатное представление и матрица этого преобразования в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$. Перейдем к новому базису $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2$. Тогда координаты всех векторов преобразуются по формулам:

$$\left.\begin{aligned} x_1 &= l_{11}\tilde{x}_1 + l_{12}\tilde{x}_2, \\ x_2 &= l_{21}\tilde{x}_1 + l_{22}\tilde{x}_2 \end{aligned}\right\} \tag{2}$$

с матрицей

$$L = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix}.$$

Обозначим через X пару старых координат вектора x, через $\tilde{X}$ — пару новых координат этого же вектора: $X = \{x_1;\ x_2\}$, $\tilde{X} = \{\tilde{x}_1;\ \tilde{x}_2\}$.

Аналогично, через X' и $\tilde{X}'$ обозначим пару старых и пару новых координат вектора $\boldsymbol{x}'$ $(X' = \{x'_1;\ x'_2\},\ \tilde{X}' = \{\tilde{x}'_1;\ \tilde{x}'_2\})$.

Считая, что формулы (2) относятся к координатам вектора $\boldsymbol{x}$, запишем их кратко:

$$X = L\tilde{X}.$$

Но по тем же формулам изменяются координаты вектора $\boldsymbol{x}'$, поэтому

$$X' = L\tilde{X}'.$$

Координатное представление (1) данного преобразования в базисе $\boldsymbol{e_1}, \boldsymbol{e_2}$ может быть записано в аналогичном виде

$$X' = AX.$$

Из последних трех равенств получаем

$$L\tilde{X}' = AL\tilde{X}.$$

Здесь слева и справа записана некоторая пара чисел. Умножим ее на матрицу L^{-1}:

$$L^{-1}L\tilde{X}' = L^{-1}AL\tilde{X}. \tag{3}$$

Так как $L^{-1}L=E$ (единичная матрица), то умножение на $L^{-1}L$ не изменяет пары чисел. Следовательно, равенство (3) может быть переписано так:

$$\tilde{X}' = L^{-1}AL\tilde{X}. \tag{4}$$

Это равенство выражает пару новых координат вектора $\boldsymbol{x}'$ через пару новых координат вектора $\boldsymbol{x}$; значит, равенство (4) есть координатное представление данного линейного преобразования $\boldsymbol{x}'=A\boldsymbol{x}$ в новом базисе. Если мы обозначим матрицу данного линейного преобразования в новом базисе через $\tilde{A}$, то из (4) получим

$$\tilde{A} = L^{-1}AL. \tag{5}$$

Итак, *если мы переходим к новому базису*

$$\tilde{\boldsymbol{e}}_1 = l_{11}\boldsymbol{e}_1 + l_{21}\boldsymbol{e}_2,$$
$$\tilde{\boldsymbol{e}}_2 = l_{12}\boldsymbol{e}_1 + l_{22}\boldsymbol{e}_2,$$

то матрица A любого линейного преобразования векторов плоскости заменяется матрицей $\tilde{A}$, которая выражается равенством (5); *в этом равенстве*

$$L = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix}.$$

Укажем одно важное следствие формулы (5). Именно, по теореме об определителе произведения матриц имеем

$$\operatorname{Det}\tilde{A} = \operatorname{Det} L^{-1}\cdot\operatorname{Det} A\cdot\operatorname{Det} L =$$
$$= \operatorname{Det} A\cdot\operatorname{Det} L^{-1}\cdot\operatorname{Det} L.$$

Но $L^{-1}L=E$; следовательно,

$$\operatorname{Det} L^{-1}\operatorname{Det} L = \operatorname{Det} E = 1.$$

Таким образом,

$$\operatorname{Det}\tilde{A} = \operatorname{Det}A,$$

т. е. *определитель матрицы линейного преобразования не зависит от выбора базиса.* По существу, это обстоятельство уже было установлено в п° 22, где показано, что определитель преобразования есть коэффициент изменения площадей. Поскольку площадь фигуры не зависит от выбора координат, то и рассматриваемый определитель не может зависеть от этого выбора.

Пример. Дано координатное представление некоторого линейного преобразования в базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$:

$$x'_1 = x_1 + 2x_2,$$

$$x_2 = 3x_1 + 4x_2.$$

Найти координатное представление того же преобразования в базисе $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$, где

$$\tilde{\boldsymbol{e}}_1 = 11\boldsymbol{e}_1 + 7\boldsymbol{e}_2,$$
$$\tilde{\boldsymbol{e}}_2 = 3\boldsymbol{e}_1 + 2\boldsymbol{e}_2.$$

Решение. Имеем:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad L = \begin{pmatrix} 11 & 3 \\ 7 & 2 \end{pmatrix}.$$

Так как $\operatorname{Det}L = 1 \neq 0$, то существует

$$L^{-1} = \begin{pmatrix} 2 & -3 \\ -7 & 11 \end{pmatrix}.$$

По формуле (5)

$$\tilde{A} = L^{-1}AL = \begin{pmatrix} -133 & -38 \\ 496 & 138 \end{pmatrix}.$$

Таким образом, данное линейное преобразование в новом базисе имеет координатное представление

$$\tilde{x}_1 = -133\tilde{x}_1 - 38\tilde{x}_2,$$
$$\tilde{x}'_2 = 496\tilde{x}_1 + 138\tilde{x}_2.$$

6.7. Матричная запись системы двух линейных уравнений

37. Изложенные в предыдущих параграфах матричные операции удобно применять к решению систем линейных уравнений.

Пусть дана система уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 &= h_1, \\ a_{21}x_1 + a_{22}x_2 &= h_2. \end{aligned}\right\} \qquad (1)$$

Составим матрицу из коэффициентов при неизвестных:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Согласно п° 34 мы можем систему (1) записать в следующем виде:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \qquad (2)$$

Обозначим пару неизвестных одной буквой *X*, а пару чисел h_1, h_2 — одной буквой H $(X = \{x_1;\ x_2\},\ H = \{h_1;\ h_2\})$. Тогда система (1) запишется в компактном виде:

$$AX = H. \qquad (3)$$

Соотношение (3) называется *матричной записью данной системы.*

Если Det *A*, то матрица *A* имеет обратную; в этом случае

$$X = A^{-1}H. \qquad (4)$$

Формула (4) выражает пару неизвестных через данные величины. Эту формулу называют *матричной записью решения.* Чтобы вычислить решение по формуле (4), нужно; 1) составить матрицу A^{-1} (см. п° 26), 2) умножить квадратную матрицу A^{-1} на матрицу *H,* состоящую из одного столбца (см. п° 34).

38. Матричная запись решения имеет практическое значение в тех задачах, где приходится многократно решать систему при одних и тех же коэффициентах в левых частях, но с различными заданиями правых частей. В таких задачах достаточно один раз затратить труд на составление обратной матрицы, чтобы затем уже по готовым формулам находить неизвестные в зависимости от правых частей системы. Особое значение этот метод получает в задачах, сводящихся к линейным системам с большим числом неизвестных.

39. Пример. Дана система

$$\begin{aligned} 5x_1 + 3x_2 &= h_1, \\ 3x_1 + 2x_2 &= h_2. \end{aligned}$$

Найти выражение неизвестных x_1, x_2 через h_1, h_2.

Решение. Имеем

$$A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix},$$

Det A=1≠0. Матрица A является невырожденной, следовательно, имеет обратную. Согласно п° 26 находим

$$A^{-1}=\begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix};$$

отсюда

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix},$$

или

$$x_1=2h_1-3h_2,$$
$$x_2=-3h_1+5h_2.$$

Микромодуль 16
Линейные преобразования в пространстве

6.8. Линейное преобразование в пространстве и квадратные матрицы третьего порядка

В этом параграфе кратко излагаются основные сведения о линейных преобразованиях в пространстве. Пункты этого параграфа даются с подзаголовками. По ним будет легко установить аналогичные вопросы о линейных преобразованиях на плоскости, которые излагались нами с достаточной подробностью.

40. Понятие линейного преобразования в пространстве. Координатное представление. Отметим в пространстве какую-нибудь точку O; всякий вектор $\boldsymbol{x}$ будем считать приложенным к точке O и рассматривать как радиус-вектор некоторой точки M: $\boldsymbol{x}=\overline{OM}$. Говорят, что в пространстве задано *преобразование точек,* если указано правило, согласно которому с любой точкой M сопоставлена некоторая точка M' (иначе говоря, M перемещена в M'). Далее всегда предполагается, что точка O остается на месте. Одновременно с преобразованием точек устанавливается *преобразование векторов,* при котором произвольный вектор $\boldsymbol{x}=\overline{OM}$ переводится в вектор $\boldsymbol{x}'=\overline{OM'}$. Вектор $\boldsymbol{x}'$ мы назовем *образом* вектора $\boldsymbol{x}$ и будем писать: $\boldsymbol{x}'=A\boldsymbol{x}$.

Преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ *называется линейным, если* 1) $A(\lambda \boldsymbol{x}) = \lambda A\boldsymbol{x}$, *где* $\boldsymbol{x}$ *— любой вектор пространства,* λ *—любое число;*
2) $A(\boldsymbol{x}+\boldsymbol{y}) = A\boldsymbol{x}+A\boldsymbol{y}$, *где* $\boldsymbol{x}, \boldsymbol{y}$ *—любые два вектора.*

Выберем в пространстве базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ (в качестве $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ разрешается взять любые векторы, лишь бы они были некомпланарны). Тогда каждый вектор $\boldsymbol{x}$ можно разложить по выбранному базису:

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3.$$

Коэффициенты этого разложения назовем *координатами вектора* $\boldsymbol{x}$ и будем писать: $\boldsymbol{x} = \{x_1;\ x_2;\ x_3\}$. Эти же коэффициенты будем называть *координатами точки М,* если $\boldsymbol{x} = \overline{OM}$ (тем самым точка *O* принята в качестве начала координат).

Рассмотрим какое-нибудь линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$. Предположим, что нам даны образы базисных векторов $\boldsymbol{e}_1' = A\boldsymbol{e}_1,\ \boldsymbol{e}_2' = A\boldsymbol{e}_2,\ \boldsymbol{e}_3' = A\boldsymbol{e}_3$, т. е. известны их разложения по базису $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$:

$$\left.\begin{aligned} \boldsymbol{e}_1' &= A\boldsymbol{e}_1 = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2 + a_{31}\boldsymbol{e}_3, \\ \boldsymbol{e}_2' &= A\boldsymbol{e}_2 = a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2 + a_{32}\boldsymbol{e}_3, \\ \boldsymbol{e}_3' &= A\boldsymbol{e}_3 = a_{13}\boldsymbol{e}_1 + a_{23}\boldsymbol{e}_2 + a_{33}\boldsymbol{e}_3. \end{aligned}\right\} \tag{1}$$

Легко показать, что задание формул (1) определяет линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$. Пусть $\boldsymbol{x} = \{x_1;\ x_2;\ x_3\}$—произвольный вектор, $\boldsymbol{x}' = \{x_1';\ x_2';\ x_3'\}$—его образ (координаты этих векторов относятся к базису $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$). Мы имеем

$$\begin{aligned} \boldsymbol{x}' &= x_1'\boldsymbol{e}_1 + x_2'\boldsymbol{e}_2 + x_3'\boldsymbol{e}_3 = A\boldsymbol{x} = A(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3) = \\ &= A(x_1\boldsymbol{e}_1) + A(x_2\boldsymbol{e}_2) + A(x_3\boldsymbol{e}_3) = \\ &= x_1A\boldsymbol{e}_1 + x_2A\boldsymbol{e}_2 + x_3A\boldsymbol{e}_3 = \\ &= x_1\boldsymbol{e}_1' + x_2\boldsymbol{e}_2' + x_3\boldsymbol{e}_3'. \end{aligned}$$

Следовательно,

$$\boldsymbol{x}' = x_1'\boldsymbol{e}_1 + x_2'\boldsymbol{e}_2 + x_3'\boldsymbol{e}_3 = x_1\boldsymbol{e}_1' + x_2\boldsymbol{e}_2' + x_3\boldsymbol{e}_3'. \tag{2}$$

Заменим в правой части равенства (2) векторы $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ их разложениями (1) и сгруппируем подобные члены; мы получим:

$$\begin{aligned} \boldsymbol{x}' = x_1'\boldsymbol{e}_1 &+ x_2'\boldsymbol{e}_2 + x_3'\boldsymbol{e}_3 = \\ &= (a_{11}x_1 + a_{12}x_2 + a_{13}x_3)\,\boldsymbol{e}_1 + \\ &+ (a_{21}x_1 + a_{22}x_2 + a_{23}x_3)\,\boldsymbol{e}_2 + \\ &+ (a_{31}x_1 + a_{32}x_2 + a_{33}x_3)\,\boldsymbol{e}_3. \end{aligned}$$

Отсюда, сравнивая коэффициенты при одинаковых базисных векторах, найдем

$$\left.\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x'_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{aligned}\right\} \quad (3)$$

Формулы (3) позволяют по любому вектору $\boldsymbol{x}=\{x_1;\ x_2;\ x_3\}$ определить его образ $\boldsymbol{x}' = \{\ x'_1;\ x'_2;\ x'_3\}$; коэффициенты этих формул даны, если даны формулы (1). Тем самым показано, что формулы (1) действительно определяют линейное преобразование $\boldsymbol{x}'= A\boldsymbol{x}$, которое переводит базисные векторы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ в векторы $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$, $\boldsymbol{e}'_3$. Иначе говоря, *если мы знаем образы базисных векторов при некотором линейном преобразовании, то знаем образы всех векторов.* Равенства (3) линейно выражают координаты вектора $\boldsymbol{x}'= A\boldsymbol{x}$ через координаты вектора $\boldsymbol{x}$; мы будем их называть *координатным представлением* данного линейного преобразования $\boldsymbol{x}'= A\boldsymbol{x}$ в данном базисе.

Таблицу коэффициентов правых частей (3) обозначим буквой A и назовем *матрицей* данного преобразования в данном базисе:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}; \quad (4)$$

числа $a_{11}, ..., a_{33}$ называют *элементами* данной матрицы. Матрицы вида (4) имеют три строчки и три столбца; такие матрицы называются квадратными матрицами третьего порядка. Матрица коэффициентов формул (1) получается *транспонированием* матрицы A, т. е. перестановкой элементов, симметричных относительно главной диагонали. Результат транспонирования данной матрицы мы будем обозначать звездочкой. Соответственно этому матрица коэффициентов формул (1) запишется так:

$$A^* = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}. \quad (5)$$

Укажем еще, что *любые формулы вида* (3) (т. е. с любыми коэффициентами) *дают координатное представление некоторого линейного преобразования.* Это утверждение доказывается в полной аналогии с тем, как было доказано соответствующее утверждение о линейных преобразованиях на плоскости, и останавливаться на его доказательстве мы не будем.

Пример. Пусть α—какая-нибудь плоскость, проходящая через точку O. Сопоставим с произвольным вектором $\boldsymbol{x}=\overline{OM}$

вектор $\boldsymbol{x}' = A\boldsymbol{x} = \overline{OM'}$ при следующих условиях: если P— основание перпендикуляра, опущенного из M на α, то M' лежит на луче PM, причем отношение PM' к PM равно заранее данному положительному числу k. Доказать, что рассматриваемое преобразование $\boldsymbol{x}' = A\boldsymbol{x}$ является линейным.

Решение. Возьмем базис, состоящий из единичных и попарно перпендикулярных векторов $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, векторы $\boldsymbol{i}$, $\boldsymbol{j}$ расположим в плоскости α. Тогда, если $\boldsymbol{x} = \{x_1;\ x_2;\ x_3\}$, $\boldsymbol{x}' = \{x'_1;\ x'_2;\ x'_3\}$, то

$$x'_1 = x_1, \quad x'_2 = x_2, \quad x'_3 = kx_3. \tag{6}$$

Полученные формулы имеют вид (3) и, следовательно, являются координатным представлением линейного преобразования. В данном случае

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{pmatrix}. \tag{7}$$

Линейное преобразование (6) называется *сжатием* к плоскости α (содержащей первые две координатные оси), число k называется *коэффициентом сжатия,* матрица (7) называется *матрицей сжатия* (к той же плоскости).

41. **Произведение линейных преобразований в пространстве. Произведение квадратных матриц третьего порядка.** Пусть даны две матрицы:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Возьмем какой-нибудь базис $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ и рассмотрим линейные преобразования $A\boldsymbol{x}$ и $B\boldsymbol{x}$, которые в этом базисе имеют данные матрицы A и B.

Пусть $\boldsymbol{x} = \{x_1;\ x_2;\ x_3\}$ — произвольный вектор. Преобразование с матрицей B переводит его в некоторый вектор $\boldsymbol{y} = B\boldsymbol{x} = \{y_1,\ y_2,\ y_3\}$; полученный вектор другим данным преобразованием будет переведен в вектор $\boldsymbol{x}' = A\boldsymbol{y} = = \{\ x'_1;\ x'_2;\ x'_3\}$. В результате с каждым вектором $\boldsymbol{x}$ сопоставлен вектор $\boldsymbol{x}'$, т. е. установлено определенное преобразование векторов. Это преобразование обозначается $AB\boldsymbol{x}$ и называется *произведением* двух данных преобразований:

$$AB\boldsymbol{x} = A(B\boldsymbol{x}).$$

Проводя выкладки аналогично тому, как делалось в п.5.3, можно показать, что произведение двух линейных преобразований есть также линейное преобразование, и найти его матрицу.

Матрица преобразования $AB\mathbf{x}$ называется *произведением матриц A, B* и обозначается символом *AB*. Произведение *AB* может быть непосредственно найдено по данным матрицам *A* м *B,* согласно следующему правилу:

$$AB=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31} & a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32} & a_{31}b_{13}+a_{32}b_{23}+a_{33}b_{33} \end{pmatrix}. \quad (8)$$

Чтобы пояснить правило умножения матриц, выраженное равенством (8), рассмотрим элемент произведения, занимающий место в i-й строке и в k-м столбце; этот элемент согласно (8) записывается так:

$$a_{i1}b_{1k}+a_{i2}b_{2k}+a_{i3}b_{3k}.$$

Здесь участвуют элементы a_{i1}, a_{i2}, a_{i3}, составляющие i -ю строку матрицы *A,* и элементы b_{1k}, b_{2k}, b_{3k}, составляющие k -й столбец матрицы *B*. Сама *сумма* $a_{i1}b_{1k}+a_{i2}b_{2k}+a_{i3}b_{3k}$ *называется произведением i-й строки матрицы A на k-й столбец матрицы B*. Таким образом, *чтобы получить элемент произведения AB, занимающий место в i-й строке и в k-м столбце* (i = 1, 2, 3, k=1, 2, 3), *нужно помножить i-ю строку матрицы A на k-й столбец матрицы B*.

Свойства произведений матриц третьего порядка аналогичны свойствам произведений матриц второго порядка; мы не будем задерживаться на их изложении. Точно так же, в полной аналогии с матрицами второго порядка, определяется умножение матрицы третьего порядка на число и сумма таких матриц.

Пример. Пусть $A\mathbf{x}$ есть сжатие к плоскости Ox_2x_3 с коэффициентом k_1, $B\mathbf{x}$—сжатие к плоскости Ox_1x_3 с коэффициентом k_2, $C\mathbf{x}$—сжатие к плоскости Ox_1x_2 с коэффициентом k_3 (см. пример в п° 40). Матрицы этих преобразований соответственно будут;

$$A=\begin{pmatrix} k_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B=\begin{pmatrix} 1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k_3 \end{pmatrix}.$$

Перемножая *A, B, C,* получим

$$ABC=\begin{pmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{pmatrix}.$$

Матрица такого вида называется *диагональной.* Таким образом, *диагональная матрица третьего порядка отвечает произведению трех сжатий к координатным плоскостям.*

43 Теорема об определителе произведения двух матриц. Для матриц третьего порядка, как и для матриц второго порядка, справедливо равенство

$$\mathrm{Det}\, AB = \mathrm{Det}\, A \cdot \mathrm{Det}\, B. \qquad (9)$$

Более того, это равенство естественным образом обобщается на матрицы любого порядка.

44. Геометрический смысл определителя линейного преобразования в пространстве. Вырожденные преобразования в пространстве. Прежде всего мы найдем выражение смешанного произведения трех векторов ***x, y, z,*** заданных своими разложениями по произвольному базису $\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}$:

$$\left.\begin{aligned} \boldsymbol{x} &= x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3, \\ \boldsymbol{y} &= y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + y_3\boldsymbol{e}_3, \\ \boldsymbol{z} &= z_1\boldsymbol{e}_1 + z_2\boldsymbol{e}_2 + z_3\boldsymbol{e}_3. \end{aligned}\right\} \qquad (10)$$

Перемножим сначала *x* и *y* векторно:

$$[\boldsymbol{xy}] = (x_2y_3 - x_3y_2)[\boldsymbol{e}_2\boldsymbol{e}_3] + \\ + (x_3y_1 - x_1y_3)[\boldsymbol{e}_3\boldsymbol{e}_1] + (x_1y_2 - x_2y_1)[\boldsymbol{e}_1\boldsymbol{e}_2]. \qquad (11)$$

Теперь умножим скалярно обе части равенства (11) на обе части последнего равенства (10); мы получим

$$(\boldsymbol{xyz}) = \{(x_2y_3 - x_3y_2)z_1 + (x_3y_1 - x_1y_3)z_2 + \\ + (x_1y_2 - x_2y_1)z_3\}(\boldsymbol{e}_1\boldsymbol{e}_2\boldsymbol{e}_3), \qquad (12)$$

где ***xyz*** — смешанное произведение векторов ***x, y, z;*** $\boldsymbol{e_1e_2e_3}$ — смешанное произведение базисных векторов. Предположим для определенности, что тройка векторов $\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}$ — правая и что векторное произведение ориентировано по правилу правой руки. Тогда смешанное произведение $\boldsymbol{e_1e_2e_3}$ есть положительная величина, равная объему параллелепипеда, построенного на векторах $\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}$; обозначим этот объем через V_0. Имеем: $\boldsymbol{e_1e_2e_3} = V_0$.

Смешанное произведение ***xyz*** равно объему параллелепипеда, построенного на векторах ***x, y, z,*** взятому со знаком плюс, если тройка ***x, y, z*** правая, и со знаком минус, если эта тройка левая. Условимся смешанное произведение ***xyz*** называть ориентированным объемом параллелепипеда, построенного на векторах ***x, y, z,*** и обозначать буквой $\boldsymbol{\tau}$; имеем: $\boldsymbol{xyz} = \boldsymbol{\tau}$. Заметим, наконец, что в фигурных скобках равенства (12) записан в развернутом виде определитель, составленный из координат данных векторов ***x, y, z***. Учтя все сказанное, получаем формулу:

$$\boldsymbol{\tau} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} V_0. \qquad (13)$$

Рассмотрим какое-нибудь линейное преобразование пространства ***x'***=*A**x***. Пусть ***x', y', z'*** — образы векторов ***x, y, z*** в данном преобразовании, $\boldsymbol{\tau}'$ — ориентированный объем параллелепипеда, построенного на векторах ***x', y', z'***. Тогда аналогично (13) будет иметь место равенство:

$$\boldsymbol{\tau}' = \begin{vmatrix} x_1' & y_1' & z_1' \\ x_2' & y_2' & z_2' \\ x_3' & y_3' & z_3' \end{vmatrix} V_0. \qquad (14)$$

Здесь определитель составлен по координатам векторов ***x', y', z'***. Далее, если матрица преобразования обозначена стандартным образом (как в равенстве (4)), то

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{aligned}$$

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + a_{13}y_3, \\ y_2' &= a_{21}y_1 + a_{22}y_2 + a_{23}y_3, \\ y_3' &= a_{31}y_1 + a_{32}y_2 + a_{33}y_3, \end{aligned}$$

$$\begin{aligned} z_1' &= a_{11}z_1 + a_{12}z_2 + a_{13}z_3, \\ z_2' &= a_{21}z_1 + a_{22}z_2 + a_{23}z_3, \\ z_3' &= a_{31}z_1 + a_{32}z_2 + a_{33}z_3. \end{aligned}$$

Согласно правилу умножения матриц мы можем все эти девять равенств записать в виде одного матричного равенства:

$$\begin{pmatrix} x_1' & y_1' & z_1' \\ x_2' & y_2' & z_2' \\ x_3' & y_3' & z_3' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix}.$$

Отсюда по теореме об определителе произведения матриц находим:

$$\begin{vmatrix} x_1' & y_1' & z_1' \\ x_2' & y_2' & z_2' \\ x_3' & y_3' & z_3' \end{vmatrix} = \operatorname{Det} A \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

Умножим обе части этого равенства на V_0; тогда в силу (13) и (14) получится соотношение, которое и было нашей целью:

$$\boldsymbol{\tau}' = \boldsymbol{\tau} \cdot \mathrm{Det}\, A. \qquad (15)$$

На основании соотношения (15) заключаем: *при линейном преобразовании пространства все параллелепипеды деформируются так, что их ориентированные объемы изменяются пропорционально; общим коэффициентом изменения объемов является определитель преобразования.* Из формулы (15) видно также следующее: если Det $A>0$, то тройка векторов $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ и тройка их образов ориентированы одинаково ($\boldsymbol{\tau}$ и $\boldsymbol{\tau}'$ имеют один и тот же знак); если Det $A<0$, то тройки $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ и $\boldsymbol{x}', \boldsymbol{y}', \boldsymbol{z}'$ ориентированы различно.

Рассмотрим теперь линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ при условии Det $A = 0$. В этом случае формула (15) дает $\boldsymbol{\tau}' = 0$ при любом значении $\boldsymbol{\tau}$. Следовательно, данное преобразование любую тройку векторов $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ переводит в три компланарных вектора $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. В частности, будут компланарны образы $\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3$ базисных векторов $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. Обозначим через α' плоскость, проходящую через точку O, на которой лежат векторы $\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3$. Вернемся к формуле (2), согласно которой образ $\boldsymbol{x}'=A\boldsymbol{x}$ любого вектора $x = \{x_1; x_2; x_3\}$ может быть выражен в виде

$$\boldsymbol{x}' = x_1\boldsymbol{e}'_1 + x_2\boldsymbol{e}'_2 + x_3\boldsymbol{e}'_3.$$

Отсюда заключаем: если Det $A=0$, то образы всех векторов пространства располагаются на одной плоскости α'. Можно сказать также, что образы всех точек пространства лежат на одной плоскости, т. е. все пространство отображается в некоторую плоскость. Такое линейное преобразование пространства называется *вырожденным.* Согласно изложенному, *вырожденное линейное преобразование характеризуется условием:* Det $A= 0$. При этом же условии называется *вырожденной и матрица A,*

Пример. Линейное преобразование

$$x'_1 = x_1 + 2x_2 + 3x_3,$$
$$x'_2 = 3x_1 + 2x_2 + x_3,$$
$$x'_3 = 4x_1 + 4x_2 + 4x_3$$

является вырожденным, так как образы всех точек пространства лежат на плоскости $x'_1 + x'_2 - x'_3 = 0$. Вместе с тем

$$\mathrm{Det}\, A = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 4 & 4 & 4 \end{vmatrix} = 0.$$

44. Обращение линейного преобразования в пространстве. Рассмотрим в пространстве линейное преобразование $\mathbf{x}' = A\mathbf{x}$; здесь $\mathbf{x}'$—*образ* вектора $\mathbf{x}$, вектор $\mathbf{x}$ — *прообраз* вектора $\mathbf{x}'$. Преобразование, которое с произвольным вектором $\mathbf{x}'$ сопоставляет его прообраз $\mathbf{x}$, в данном преобразовании, называется обратным данному и обозначается $x = A^{-1}\mathbf{x}'$. Займемся задачей: по данному линейному преобразованию найти обратное ему.

Предположим, что данное преобразование $\mathbf{x}' = A\mathbf{x}$ является вырожденным; тогда образы всех векторов пространства лежат на одной плоскости α'. Если мы возьмем вектор $\mathbf{x}'$, не лежащий на этой плоскости, то для него нет прообраза. Поэтому *для вырожденного преобразования нет обратного.*

Пусть теперь $\mathbf{x}' = A\mathbf{x}$ — невырожденное линейное преобразование,

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned}\right\} \qquad (16)$$

— его координатное представление в некотором базисе $\mathbf{e}_1$,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

—матрица данного преобразования. Так как преобразование не вырождено, то Δ= Det A≠0. Чтобы получить преобразование, обратное данному, нужно выразить координаты прообраза $\{x_1;\ x_2;\ x_3\}$ через координаты образа $\{x_1';\ x_2';\ x_3'\}$, т. е. решить систему (16) относительно x_1, x_2, x_3, считая x'_1, x'_2, x'_3 известными. Поскольку Δ≠0, мы находим

$$x_1 = \frac{\begin{vmatrix} x_1' & a_{12} & a_{13} \\ x_2' & a_{22} & a_{23} \\ x_3' & a_{32} & a_{33} \end{vmatrix}}{\Delta}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & x_1' & a_{13} \\ a_{21} & x_2' & a_{23} \\ a_{31} & x_3' & a_{33} \end{vmatrix}}{\Delta}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & x_1' \\ a_{21} & a_{22} & x_2' \\ a_{31} & a_{32} & x_3' \end{vmatrix}}{\Delta}. \qquad (17)$$

Будем обозначать через A_{ik} алгебраическое дополнение элемента a_{ik} определителя матрицы A. Рассмотрим первый из написанных выше определителей. Разлагая этот определитель по элементам первого столбца, получим

$$\begin{vmatrix} x'_1 & a_{12} & a_{13} \\ x'_2 & a_{22} & a_{23} \\ x'_3 & a_{32} & a_{33} \end{vmatrix} = A_{11}x'_1 + A_{21}x'_2 + A_{31}x'_3. \qquad (18)$$

В самом деле, x'_1, x'_2, x'_3 занимают места элементов a_{11}, a_{21}, a_{31} определителя матрицы A, поэтому в разложении определителя (18) по элементам первого столбца числа x'_1, x'_2, x'_3 должны помножаться на алгебраические дополнения A_{11}, A_{21}, A_{31}. Производя аналогичные разложения двух других определителей в формулах (17), приведем эти формулы к следующему виду:

$$\left.\begin{aligned} x_1 &= \frac{A_{11}}{\Delta}x'_1 + \frac{A_{21}}{\Delta}x'_2 + \frac{A_{31}}{\Delta}x'_3, \\ x_2 &= \frac{A_{12}}{\Delta}x'_1 + \frac{A_{22}}{\Delta}x'_2 + \frac{A_{32}}{\Delta}x'_3, \\ x_3 &= \frac{A_{13}}{\Delta}x'_1 + \frac{A_{23}}{\Delta}x'_2 + \frac{A_{33}}{\Delta}x'_3. \end{aligned}\right\} \qquad (19)$$

Мы получили координатное представление преобразования $\boldsymbol{x} = A^{-1}\boldsymbol{x}'$, обратного данному линейному преобразованию $\boldsymbol{x}'=A\boldsymbol{x}$ (в том же базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$). Так как выражения (19) имеют линейный вид, заключаем, что *обратное преобразование также является линейным.* Матрицу преобразования $\boldsymbol{x} = A^{-1}\boldsymbol{x}'$обозначают через A^{-1} и называют *обратной матрицей* для данной матрицы A. Из (19) находим:

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \frac{A_{31}}{\Delta} \\ \frac{A_{12}}{\Delta} & \frac{A_{22}}{\Delta} & \frac{A_{32}}{\Delta} \\ \frac{A_{13}}{\Delta} & \frac{A_{23}}{\Delta} & \frac{A_{33}}{\Delta} \end{pmatrix}. \qquad (20)$$

В пространстве, как и на плоскости, рассматривается *тождественное преобразование,* которое в любом базисе дается формулами: $x'_1 = x_1$, $x'_2 = x_2$, $x'_3 = x_3$. Матрица тождественного преобразования обозначается буквой E и называется *единичной:*

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Имеют место равенства:

$$A^{-1}A = E, \quad AA^{-1} = E. \qquad (21)$$

Кроме того, для любой матрицы A

$$AE = A, \quad EA = A. \qquad (22)$$

Чтобы доказать равенства (21), (22), достаточно повторить соображения, приведенные в п° 28 и 29.

45. Преобразование координат векторов пространства при переходе к новому базису. Пусть базис $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\boldsymbol{e_3}$ заменяется новым базисом:

$$\left.\begin{aligned}\tilde{e}_1 &= l_{11}e_1 + l_{21}e_2 + l_{31}e_3,\\ \tilde{e}_2 &= l_{12}e_1 + l_{22}e_2 + l_{32}e_3,\\ \tilde{e}_3 &= l_{13}e_1 + l_{23}e_2 + l_{33}e_3.\end{aligned}\right\} \qquad (23)$$

Матрицу коэффициентов этих разложений мы будем считать известной и обозначим через L^*. Так как векторы, составляющие базис (как старый, так и новый), не компланарны, то Det $L^* \neq 0$. (Это следует из того, что объем параллелепипеда, построенного на векторах $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_3$ отличен от нуля, см. формулу (13) в п° 43.) Рассмотрим произвольный вектор $\boldsymbol{x}$ с координатами x_1, x_2, x_3 в базисе $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\boldsymbol{e_3}$; этот же вектор $\boldsymbol{x}$ в новом базисе $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_3$ будет иметь другие координаты $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$. Связь между старыми и новыми координатами вектора $\boldsymbol{x}$ дается формулами:

$$\left.\begin{aligned}x_1 &= l_{11}\tilde{x}_1 + l_{12}\tilde{x}_2 + l_{13}\tilde{x}_3,\\ x_2 &= l_{21}\tilde{x}_1 + l_{22}\tilde{x}_2 + l_{23}\tilde{x}_3,\\ x_3 &= l_{31}\tilde{x}_1 + l_{32}\tilde{x}_2 + l_{33}\tilde{x}_3.\end{aligned}\right\} \qquad (24)$$

Формулы (24) аналогичны формулам (4) п° 31 и доказываются вполне аналогичным образом. Матрицу коэффициентов формул (24) мы обозначим через L (L и L^* переводятся друг в друга транспонированием). Будем употреблять умножение квадратной матрицы третьего порядка на матрицу из одного столбца (составленного тремя числами); эта операция определяется в полной аналогии с тем, что делалось в п°34. Тогда три равенства (24) заменятся одним матричным равенством:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix}.$$

Условимся тройку чисел $\{x_1, x_2, x_3\}$ обозначать одной буквой X, а тройку чисел $\{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3\}$ — одной буквой $\tilde{X}$. После этого равенства (24) получат следующий вид:

$$X = L\tilde{X}. \qquad (25)$$

Обратные формулы (выражающие новые координаты вектора через его старые координаты) сведутся к матричному равенству

$$\tilde{X} = L^{-1}X. \tag{26}$$

Матрица L^{-1} существует, так как $\operatorname{Det} L = \operatorname{Det} L^* \neq 0.$

46. Преобразование прямоугольных координат. Ортогональная матрица третьего порядка. Рассмотрим частный случай преобразования координат, когда: 1) старый базис состоит из единичных и перпендикулярных друг к другу векторов ***i, j, k****;* 2) новый базис $\tilde{i}, \tilde{j}, \tilde{k}$ имеет аналогичный вид. В этом случае мы изменим обозначения коэффициентов формул (23) и напишем эти формулы так:

$$\begin{aligned} \tilde{i} &= l_1 i + m_1 j + n_1 k, \\ \tilde{j} &= l_2 i + m_2 j + n_2 k, \\ \tilde{k} &= l_3 i + m_3 j + n_3 k; \end{aligned}$$

соответственно

$$L^* = \begin{pmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{pmatrix}, \quad L = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{pmatrix}.$$

Так как $\tilde{i}^2 = 1,\ \tilde{j}^2 = 1,\ \tilde{k}^2 = 1,\ \tilde{i}\,\tilde{j} = 0,\ \tilde{i}\,\tilde{k} = 0,\ \tilde{j}\,\tilde{k} = 0,$ то

$$\left.\begin{aligned} l_1^2 + m_1^2 + n_1^2 = 1,\ \ l_1 l_2 + m_1 m_2 + n_1 n_2 = 0, \\ l_2^2 + m_2^2 + n_2^2 = 1,\ \ l_1 l_3 + m_1 m_3 + n_1 n_3 = 0, \\ l_3^2 + m_3^2 + n_3^2 = 1,\ \ l_2 l_3 + m_2 m_3 + n_2 n_3 = 0. \end{aligned}\right\} \tag{27}$$

Вследствие этих равенств имеем

$$L^* L = E; \tag{28}$$

обратно, из формулы (28) следует (27). Матрица *L,* удовлетворяющая условию (28), называется *ортогональной.* Из (28) следует также, что

$$L^* = L^{-1}, \tag{29}$$

т. е. *для ортогональной матрицы обратная матрица совпадает с транспонированной.* На основании изложенного заключаем: при переходе от базиса ***i, j, k*** к аналогичному базису $\tilde{i}, \tilde{j}, \tilde{k}$ получаются формулы преобразования координат с ортогональной матрицей. В этом случае формулы прямого и обратного преобразования координат имеют матричную запись:

$$X = L\tilde{X},\ \ \tilde{X} = L^* X \tag{30}$$

(см. формулы (25), (26) п° 45 и формулу (29)). Заметим еще, что вследствие (28)

$$\operatorname{Det} L^* \cdot \operatorname{Det} L = 1;$$

но $\operatorname{Det} L^* = \operatorname{Det} L$. Следовательно, $(\operatorname{Det} L)^2 = 1$ и

$$\operatorname{Det} L = \pm 1.$$

Таким образом, определитель ортогональной матрицы может быть равен только +1 или —1. В первом случае базисы $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ и $\tilde{\boldsymbol{i}}, \tilde{\boldsymbol{j}}, \tilde{\boldsymbol{k}}$ ориентированы одинаково, во втором — различно (см. фирмулу (13) п° 43).

47. Изменение матрицы линейного преобразования в пространстве при переходе к новому базису. Пусть в пространстве дано линейное преобразование $\boldsymbol{x}' = A\boldsymbol{x}$. Если выбран некоторый базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, то данное линейное преобразование получит определенное координатное представление с определенной матрицей A. Предположим, что базис $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ заменяется новым базисом $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3$ по формулам (23) п° 45. Тогда то же самое линейное преобразование в новом базисе будет иметь другое координатное представление с другой матрицей A. При этом

$$\tilde{A} = L^{-1}AL. \tag{31}$$

Вывод формулы (31) производится точно так же, как вывод аналогичной формулы для линейных преобразований на плоскости и воспроизводить его мы не будем.

Отметим, что из матричной формулы (31) следует числовое равенство:

$$\operatorname{Det} \tilde{A} = \operatorname{Det} A,$$

т. е. определитель матрицы линейного преобразования не зависит от выбора базиса. По существу, этот факт уже установлен в п° 43, где показано, что определитель линейного преобразования выражает отношение объемов; поскольку объемы тел не зависят от выбора координат, то и рассматриваемый определитель не может зависеть от этого выбора.

Пример. Дано координатное представление линейного преобразования в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$:

$$\begin{aligned} x'_1 &= x_1 + x_2, \\ x'_2 &= x_1 \quad\;\; + x_3, \\ x'_3 &= \quad\;\; x_2 + x_3. \end{aligned}$$

Найти координатное представление этого преобразования в базисе

$$\begin{aligned} \tilde{e}_1 &= e_1 + 2e_2 + e_3, \\ \tilde{e}_2 &= 2e_1 + e_2 + 3e_3, \\ \tilde{e}_3 &= e_1 + e_2 + e_3. \end{aligned}$$

Ре ш е н и е. Имеем

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix}.$$

Применяя формулу (20) к матрице L, найдем

$$L^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix}.$$

Отсюда

$$L^{-1}A = \begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 2 \\ -1 & 0 & 1 \\ 4 & 2 & -4 \end{pmatrix};$$

далее,

$$L^{-1}AL = \begin{pmatrix} -1 & -1 & 2 \\ -1 & 0 & 1 \\ 4 & 2 & -4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 3 & 0 \\ 0 & 1 & 0 \\ 4 & -2 & 2 \end{pmatrix}.$$

Следовательно, в новом базисе данное линейное преобразование представляется формулами

$$\begin{aligned} \tilde{x}'_1 &= -\tilde{x}_1 + 3\tilde{x}_2, \\ \tilde{x}'_2 &= \qquad\quad\;\; \tilde{x}_2, \\ \tilde{x}'_3 &= 4\tilde{x}_1 - 2\tilde{x}_2 + 2\tilde{x}_3. \end{aligned}$$

48. **Матричная запись системы трех линейных уравнений**. В заключение параграфа дадим приложение матричных методов к системам трех линейных уравнений с тремя неизвестными. То, что мы сейчас будем делать, аналогично изложенному ранее. Пусть дана система уравнений:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = h_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = h_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = h_3. \end{aligned}\right\} \qquad (32)$$

Составим матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

и заменим уравнения (32) одним матричным равенством;

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}.$$

Обозначим тройку неизвестных одной буквой X, а тройку чисел h_1, h_2, h_3 — буквой H. Тогда предыдущее соотношение примет такой вид:

$$AX = H. \qquad (33)$$

Равенство (33) называется *матричной записью данной системы.*
Если $\operatorname{Det} A \neq 0$, то матрица A имеет обратную; в этом случае

$$X = A^{-1}H. \qquad (34)$$

Формула (34) выражает тройку неизвестных через данные величины. Эту формулу называют *матричной записью решения.*

Пример. Дана система

$$\begin{aligned} x_1 + 2x_2 + x_3 = h_1, \\ 2x_1 + x_2 + x_3 = h_2, \\ x_1 + 3x_2 + x_3 = h_3. \end{aligned}$$

Выразить неизвестные x_1, x_2, x_3 через h_1, h_2, h_3.

Решение. Имеем

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix}.$$

$\operatorname{Det} A = 1 \neq 0.$ Матрица A является невырожденной, следовательно, имеет обратную. Согласно п° 44 находим:

$$A^{-1}=\begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix}.$$

Отсюда

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ 5 & -1 & -3 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix},$$

или

$$\begin{aligned} x_1 &= -2h_1+h_2+h_3, \\ x_2 &= -h_1 \qquad\quad +h_3, \\ x_3 &= 5h_1-h_2-3h_3. \end{aligned}$$

6.9. Собственные векторы линейного преобразования

49. Пусть дано линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$. Вектор $\boldsymbol{x}$ (не равный нулю) называется *собственным вектором* данного линейного преобразования, если для этого вектора

$$A\boldsymbol{x} = \lambda\boldsymbol{x}, \tag{1}$$

где λ— некоторое число; число λ называется *собственным числом* вектора $\boldsymbol{x}$. Согласно высказанному определению собственный вектор $\boldsymbol{x}$ характеризуется тем, что при данном преобразовании переходит в коллинеарный ему вектор $\boldsymbol{x}'$; собственное число есть отношение вектора $\boldsymbol{x}'$ к вектору $\boldsymbol{x}$ (т. е. коэффициент в равенстве $\boldsymbol{x}'=\lambda\boldsymbol{x}$).

Отметим следующее важное обстоятельство: *если $\boldsymbol{x}$ — собственный вектор данного преобразования, то всякий не равный нулю коллинеарный ему сектор будет также собственным вектором данного преобразования с тем же собственным числом.* В самом деле, пусть $\boldsymbol{x}^*$ коллинеарен $\boldsymbol{x}$; тогда, поскольку $\boldsymbol{x}\neq 0$, найдется число α такое, что $\boldsymbol{x}^*= \alpha\boldsymbol{x}$. На основании известного свойства линейных преобразований имеем $A\boldsymbol{x}^* = A(\alpha\boldsymbol{x}) = \alpha A\boldsymbol{x} = \alpha(\lambda\boldsymbol{x}) = \lambda(\alpha\boldsymbol{x}) = \lambda\boldsymbol{x}^*$.

Итак, $A\boldsymbol{x}^* = \lambda\boldsymbol{x}^*$; а это и значит, что $\boldsymbol{x}^*$ есть собственный вектор с собственным числом λ.

50. Для иллюстрации понятия собственного вектора и собственного числа рассмотрим линейные преобразования на плоскости, которые приводились в примерах п°3 (здесь они даются в той же последовательности).

Пример 1. *Ах*—подобие с коэффициентом *k*. Для такого преобразования каждый вектор является собственным вектором; все векторы имеют одно и то же собственное число, равное коэффициенту подобия.

Пример 2. *А**x*** — вращение на угол α. Если $0 < \alpha < \pi$, то *А**x*** не имеет собственных векторов (все векторы поворачиваются так, что ни один образ не лежит на оси своего прообраза). Если α=0, то *А**x*** является тождественным преобразованием; в этом случае каждый вектор является собственным вектором с собственным числом =1. Если α=π, то ***x'***=*А**x***= —***x***; для такого преобразования каждый вектор является собственным вектором с собственным числом = —1.

Пример 3. *А**x*** — зеркальное отражение относительно прямой *а.* В этом случае собственными будут: 1) векторы, лежащие на прямой *а;* для них собственное число =1; 2) векторы, перпендикулярные к прямой *а;* собственное число этих векторов =—1.

Пример 4. *А**x*** — сжатие к прямой *а* с коэффициентом *k*. Собственными векторами являются: 1) векторы, лежащие на прямой *а* (с собственным числом =1); 2) векторы, перпендикулярные к прямой *а* (с собственным числом = *k*).

51. Пусть ***x'***=*А**x*** — линейное преобразование на плоскости. Предположим, что имеются два собственных вектора этого преобразования, которые не коллинеарны друг другу. Обозначим их через $\boldsymbol{e}_1, \boldsymbol{e}_2$; обозначим через λ_1, λ_2 собственные числа этих векторов. Так как $\boldsymbol{e}_1, \boldsymbol{e}_2$ не коллинеарны, мы можем принять их в качестве базиса. Найдем координатное представление данного линейного преобразования в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$. Пусть $\boldsymbol{e}'_1, \boldsymbol{e}'_2$ — образы векторов $\boldsymbol{e}_1, \boldsymbol{e}_2$; мы имеем

$$\boldsymbol{e}'_1 = A\boldsymbol{e}_1 = \lambda_1 \boldsymbol{e}_1,$$
$$\boldsymbol{e}'_2 = A\boldsymbol{e}_2 = \lambda_2 \boldsymbol{e}_2.$$

Отсюда

$$A^* = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Очевидно, что *А*=А;* следовательно,

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{2}$$

Итак, *если в качестве базиса приняты собственные векторы* $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ *данного линейного преобразования, то матрица А этого преобразования получает диагональный вид;* на диагонали матрицы *А* располагаются собственные числа векторов $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ в порядке нумерации. В базисе $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ данное линейное преобразование представляется формулами

$$\left.\begin{aligned} x_1' &= \lambda_1 x_1, \\ x_2' &= \qquad \lambda_2 x_2. \end{aligned}\right\} \tag{3}$$

В частном случае может быть $\lambda_1 = \lambda_2$. Не имея необходимости различать в этом случае числа λ_1 и λ_2, обозначим их одной буквой λ. Формулы (3) будут иметь вид

$$\begin{aligned} x_1' &= \lambda x_1, \\ x_2' &= \qquad \lambda x_2. \end{aligned}$$

Такие формулы определяют *подобие с коэффициентом* λ. *В этом случае каждый вектор плоскости оказывается собственным вектором с собственным числом* $= \lambda$.

57. Пусть теперь $\boldsymbol{x}'=A\boldsymbol{x}$—линейное преобразование в пространстве. Предположим, что имеются три некомпланарных собственных вектора этого преобразования: $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\boldsymbol{e_3}$; пусть λ_1, λ_2, λ_3—их собственные числа.

В полной аналогии с предыдущим можно показать, что если собственные векторы $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\boldsymbol{e_3}$ приняты в качестве базиса, то матрица данного линейного преобразования будет иметь диагональный вид

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}; \tag{4}$$

в базисе $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\boldsymbol{e_3}$ данное преобразование получит координатное представление

$$\left.\begin{aligned} x_1' &= \lambda_1 x_1, \\ x_2' &= \qquad \lambda_2 x_2, \\ x_3' &= \qquad\qquad \lambda_3 x_3. \end{aligned}\right\} \tag{5}$$

В частности, если $\lambda_1 = \lambda_2 = \lambda_3$ и если мы обозначим равные числа λ_1, λ_2, λ_3 одной буквой λ, то формулы (5) будут

$$\begin{aligned} x_1' &= \lambda x_1, \\ x_2' &= \qquad \lambda x_2, \\ x_3' &= \qquad\qquad \lambda x_3. \end{aligned}$$

Такие формулы определяют в пространстве *подобие с коэффициентом* λ*. В этом случае каждый вектор пространства является собственным вектором с собственным числом* λ*.*

53. На основании изложенного в последних двух пунктах легко понять важную роль собственных векторов в теории линейных преобразований. Как мы видели, если существует базис из собственных векторов, то в этом базисе преобразование имеет особенно простое координатное представление и вполне определяется собственными числами базисных векторов.

Методы разыскания собственных векторов и собственных чисел излагаются в ближайших параграфах.

6.10. Характеристическое уравнение матрицы линейного преобразования

54. Рассмотрим сначала линейное преобразование на плоскости $\boldsymbol{x}'=A\boldsymbol{x}$. Поставим задачу: найти все собственные векторы этого преобразования.

Выберем произвольный базис $\boldsymbol{e}_1, \boldsymbol{e}_2$; пусть

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2, \\ x_2' &= a_{21}x_1 + a_{22}x_2 \end{aligned}\right\} \tag{1}$$

— координатное представление данного преобразования в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2}$$

— матрица преобразования в этом базисе.

Предположим, что вектор $\boldsymbol{e} = \{l;\ m\}$ является собственным вектором данного линейного преобразования, λ— его собственное число. Тогда

$$A\boldsymbol{e} = \lambda\boldsymbol{e}. \tag{3}$$

В координатах это векторное равенство заменяется двумя числовыми равенствами:

$$\left.\begin{aligned} a_{11}l + a_{12}m &= \lambda l, \\ a_{21}l + a_{22}m &= \lambda m. \end{aligned}\right\} \tag{4}$$

Перепишем равенства (4) так:

$$\left.\begin{aligned} (a_{11} - \lambda)\, l + a_{12}m &= 0, \\ a_{21}l + (a_{22} - \lambda)\, m &= 0. \end{aligned}\right\} \tag{5}$$

Отсюда следует, что

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0, \qquad (6)$$

поскольку в равенствах (5) хотя бы одно из чисел l, m отлично от нуля (так как собственный вектор, по определению, не равен нулю: $\boldsymbol{e} = \{l;\ m\} \neq 0$). Тем самым показано, что *всякое собственное число является корнем уравнения* (6). Обратно, допустим, что некоторое число λ есть корень уравнения (6). Тогда, если в системе (5) в качестве λ взять именно это число, система (5) будет иметь ненулевое решение l, m (так как определитель системы (5) равен нулю). Рассмотрим вектор $\boldsymbol{e}=\{l;m\}$; для координат этого вектора соблюдаются равенства (5), следовательно, для самого вектора $\boldsymbol{e}$ и числа λ имеет место формула (3). Значит, вектор $\boldsymbol{e}$ есть собственный вектор с собственным числом λ.

Мы можем теперь сформулировать следующее правило: *чтобы найти все собственные векторы данного линейного преобразования, нужно прежде всего решить уравнение* (6); *каждый действительный корень* λ *уравнения* (6) *является собственным числом; соответствующие числу* λ *собственные векторы находятся из системы* (5).

Уравнение (6) называется *характеристическим уравнением* матрицы данного линейного преобразования.

Пример. Найти собственные векторы преобразования

$$\left.\begin{aligned} x_1' &= x_1\cos\alpha - x_2\sin\alpha, \\ x_2' &= x_1\sin\alpha + x_2\cos\alpha, \end{aligned}\right\} \qquad (*)$$

где $0 < \alpha < \pi$.

Решение. Составим характеристическое уравнение

$$\begin{vmatrix} \cos\alpha-\lambda & -\sin\alpha \\ \sin\alpha & \cos\alpha-\lambda \end{vmatrix} = 0.$$

Развертывая определитель, получаем

$$\lambda^2 - 2\lambda\cos\alpha + 1 = 0.$$

Отсюда $\lambda_{1,2} = \cos\alpha \pm \sqrt{\cos^2\alpha - 1}$. Так как $\cos^2\alpha < 1$ (при $0 < \alpha < \pi$), то корни характеристического уравнения оказываются комплексными числами. Поскольку характеристическое уравнение не имеет действительных корней, данное преобразование не имеет собственных векторов (этот результат легко понять, если учесть, что формулы (*) в прямоугольных координатах представляют вращение на угол α; см. пример 2 в п° 47).

Пример. Найти собственные векторы преобразования

$$x_1' = x_1 + x_2,$$
$$x_2' = \qquad x_2.$$

Решение. Составим характеристическое уравнение

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = 0.$$

Отсюда $(1-\lambda)^2 = 0$ и $\lambda = 1$. Система (5) в данном случае имеет вид

$$(1-\lambda)\, l + m = 0,$$
$$(1-\lambda)\, m = 0.$$

Подставляя сюда λ=1, получим: m=0 (l — любое). Таким образом, собственными векторами данного преобразования являются только векторы {l; 0}, лежащие на оси абсцисс.

55. Пусть теперь $\boldsymbol{x}'=A\boldsymbol{x}$ — линейное преобразование в пространстве. Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

— матрица этого преобразования в некотором базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$. Аналогично предыдущему можно показать, что собственные числа данного линейного преобразования определяются характеристическим уравнением:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} = 0. \tag{7}$$

Именно, если λ — любой вещественный корень уравнения (7), то система

$$\left.\begin{aligned} (a_{11}-\lambda)\, l + a_{12}m + a_{13}n = 0, \\ a_{21}l + (a_{22}-\lambda)\, m + a_{23}n = 0, \\ a_{31}l + a_{32}m + (a_{33}-\lambda)\, n = 0 \end{aligned}\right\} \tag{8}$$

имеет ненулевое решение l, m, n; вектор $\boldsymbol{e}=\{l;m;n\}$ есть собственный вектор данного преобразованияс собственным числом λ.

56. Если развернуть определитель, стоящий в левой части характеристического уравнения, то получится алгебраический (степенной) многочлен относительно λ. Этот многочлен называется *характеристическим многочленом* матрицы линейного преобразования. В случае линейных преобразований на плоскости характеристический многочлен есть многочлен второй степени, для линейных преобразований в пространстве — третьей степени.

Заметим, что характеристический многочлен является определителем матрицы $A=\lambda E$. В самом деле, для линейных преобразований на плоскости матрица $A=\lambda E$ напишется так:

$$A-\lambda E=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}-\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{pmatrix};$$

отсюда получаем характеристический многочлен

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix}=\mathrm{Det}\,(A-\lambda E).$$

Ясно, что таким же образом выражается и характеристический многочлен для преобразований в пространстве.

57. Имеет место следующая важная

Теорема. *Характеристический многочлен матрицы линейного преобразования не зависит от выбора базиса.*

Доказательство. Пусть A — матрица данного линейного преобразования в некотором базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$, $\tilde{A}$—матрица того же преобразования в новом базисе $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2$; пусть L— матрица преобразования координат при переходе от старого базиса к новому. Согласно п.6.8 (п° 44)

$$\tilde{A}=L^{-1}AL.$$

Кроме того, для единичной матрицы имеем соотношение

$$E=L^{-1}EL.$$

Из последних двух равенств получаем

$$\tilde{A}-\lambda E=L^{-1}(A-\lambda E)\,L.$$

Из этого матричного равенства и на основании теоремы об определителе произведения матриц вытекает числовое равенство

$$\mathrm{Det}\,(\tilde{A}-\lambda E)=\mathrm{Det}\,L^{-1}\cdot\mathrm{Det}\,(A-\lambda E)\cdot\mathrm{Det}\,L.$$

Но

$$\mathrm{Det}\,L^{-1}\cdot\mathrm{Det}\,L=\mathrm{Det}\,L^{-1}L=\mathrm{Det}\,E=1.$$

Следовательно,

$$\operatorname{Det}(\tilde{A}-\lambda E)=\operatorname{Det}(A-\lambda E)$$

и теорема доказана.

58. Последнее тождество, наличие которого утверждает теорема, мы напишем для случая линейных преобразований на плоскости подробно:

$$\begin{vmatrix} \tilde{a}_{11}-\lambda & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22}-\lambda \end{vmatrix} = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix}.$$

Здесь $\tilde{a}_{ik}$—элементы матрицы $\tilde{A}$, a_{ik}—элементы матрицы *A*, λ—любое число. Развертывая определители, получим

$$\lambda^2-(\tilde{a}_{11}+\tilde{a}_{22})\lambda+(\tilde{a}_{11}\tilde{a}_{22}-\tilde{a}_{12}\tilde{a}_{21})= \\ =\lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-a_{12}a_{21}).$$

Так как это — тождество, т. е. равенство, справедливое при любом λ, то

$$\tilde{a}_{11}+\tilde{a}_{22}=a_{11}+a_{22},$$
$$\tilde{a}_{11}\tilde{a}_{22}-\tilde{a}_{12}\tilde{a}_{21}=a_{11}a_{22}-a_{12}a_{21}.$$

Таким образом, для линейных преобразований на плоскости доказанная теорема означает, что *величины* $a_{11}+a_{22}$ *и* $a_{11}a_{22}-a_{12}a_{21}$ *являются инвариантами матрицы линейного преобразования,* т. е. что эти величины не меняются при переходе к новому базису. Величина $a_{11}a_{22}-a_{12}a_{21}$ представляет собой не что иное, как определитель матрицы *A;* инвариантность этой величины доказана еще в п.6.6.

Матрица линейного преобразования в пространстве имеет три инварианта; они находятся аналогичным способом. Выписывать их мы не будем.

6.11. Симметрические линейные преобразования.

59. В п.6.9 мы показали, что матрица линейного преобразования принимает диагональный вид, если в качестве базиса приняты собственные векторы. Координатное представление линейного преобразования в таком базисе оказывается особенно простым.

Однако не для всякого линейного преобразования возможно такое упрощение матрицы и координатного представления. В самом деле, например, на плоскости встречаются линейные преобразования, которые совсем не имеют собственных векторов; возможны преобразования, которые имеют только коллинеарные собственные векторы (см. примеры в п° 54). Для таких преобразований выбрать базис, состоящий из собственных векторов, нельзя.

Далее мы будем изучать некоторый специальный класс линейных преобразований, называемых *симметрическими*. Они существенно применяются в алгебре, в геометрии, в механике. Симметрические преобразования всегда имеют собственные векторы, из которых можно составить базис. Вместе с тем матрица всякого симметрического преобразования приводится к диагональному виду.

60. Пусть дано линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ (на плоскости или в пространстве). Рассмотрим два произвольных вектора $\boldsymbol{x}$, $\boldsymbol{y}$; $\boldsymbol{x}'=A\boldsymbol{x}$, $\boldsymbol{y}'=A\boldsymbol{y}$ — их образы. Данное линейное преобразование называется *симметрическим,* если скалярное произведение $\boldsymbol{xy}'$ равно скалярному произведению $\boldsymbol{yx}'$:

$$\boldsymbol{xy}'=\boldsymbol{yx}'. \tag{1}$$

Иначе говоря, *линейное преобразование* $\boldsymbol{x}'=A\boldsymbol{x}$ *является симметрическим, если*

$$\boldsymbol{x}A\boldsymbol{y}=\boldsymbol{y}A\boldsymbol{x} \tag{2}$$

для любых векторов $\boldsymbol{x}$ *и* $\boldsymbol{y}$.

61. В дальнейшем, изучая симметрические преобразования, мы будем употреблять только такие базисы, которые состоят из единичных и перпендикулярных (ортогональных) друг к другу векторов:

1) на плоскости: $\boldsymbol{e}_1=\boldsymbol{i}$, $\boldsymbol{e}_2=\boldsymbol{j}$,

2) в пространстве: $\boldsymbol{e}_1=\boldsymbol{i}$, $\boldsymbol{e}_2=\boldsymbol{j}$, $\boldsymbol{e}_3=\boldsymbol{k}$.

Такого рода базисы будем называть *ортогональными.*

62. Докажем, что *всякое симметрическое линейное преобразование в любом ортогональном базисе имеет симметрическую матрицу* (т. е. элементы a_{ik} и a_{ki} этой матрицы одинаковы: $a_{ik}=a_{ki}$).

Чтобы не писать слишком длинных выражений, проведем доказательство для случая симметрических преобразований на плоскости.

Напишем разложения образов базисных векторов:

$$\left.\begin{aligned} A\boldsymbol{e}_1&=a_{11}\boldsymbol{e}_1+a_{21}\boldsymbol{e}_2,\\ A\boldsymbol{e}_2&=a_{12}\boldsymbol{e}_1+a_{22}\boldsymbol{e}_2.\end{aligned}\right\} \tag{3}$$

Умножим скалярно обе части первого из этих равенств на вектор $\boldsymbol{e}_2$, а обе части второго — на вектор $\boldsymbol{e}_1$; принимая во внимание, что $\boldsymbol{e}_1\boldsymbol{e}_2 = 0,\ \boldsymbol{e}_1\boldsymbol{e}_1 = 1,\ \boldsymbol{e}_2\boldsymbol{e}_2 = 1$, получим

$$\boldsymbol{e}_2A\boldsymbol{e}_1 = a_{21},\ \ \boldsymbol{e}_1A\boldsymbol{e}_2 = a_{12}.$$

По условию симметрии преобразования имеем

$$\boldsymbol{e}_2A\boldsymbol{e}_1 = \boldsymbol{e}_1A\boldsymbol{e}_2.$$

Отсюда

$$a_{21} = a_{12},$$

что и означает симметрию матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Справедливо также обратное утверждение: *если в некотором ортогональном базисе матрица линейного преобразования симметрична, то симметрично и само преобразование.*

Доказательство этого утверждения мы также проведем для случая линейных преобразований на плоскости.

Предположим, что для разложений (3) соблюдается равенство $a_{21}= a_{12}$. Пусть $\boldsymbol{x}, \boldsymbol{y}$ — два произвольных вектора

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2, \quad \boldsymbol{y} = y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2.$$

Используя разложения (3), получим

$$\begin{aligned} A\boldsymbol{x} &= x_1A\boldsymbol{e}_1 + x_2A\boldsymbol{e}_2 = \\ &= (a_{11}x_1 + a_{12}x_2)\,\boldsymbol{e}_1 + (a_{21}x_1 + a_{22}x_2)\,\boldsymbol{e}_2, \\ A\boldsymbol{y} &= y_1A\boldsymbol{e}_1 + y_2A\boldsymbol{e}_2 = \\ &= (a_{11}y_1 + a_{12}y_2)\,\boldsymbol{e}_1 + (a_{21}y_1 + a_{22}y_2)\,\boldsymbol{e}_2. \end{aligned}$$

Составим скалярные произведения $\boldsymbol{x}$ на $A\boldsymbol{y}$ и $\boldsymbol{y}$ на $A\boldsymbol{x}$:

$$\begin{aligned} \boldsymbol{x}A\boldsymbol{y} &= x_1\,(a_{11}y_1 + a_{12}y_2) + x_2\,(a_{21}y_1 + a_{22}y_2) = \\ &= a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2, \\ \boldsymbol{y}A\boldsymbol{x} &= y_1\,(a_{11}x_1 + a_{12}x_2) + y_2\,(a_{21}x_1 + a_{22}x_2) = \\ &= a_{11}x_1y_1 + a_{12}x_2y_1 + a_{21}x_1y_2 + a_{22}x_2y_2. \end{aligned}$$

Так как $a_{21} = a_{12}$, то из предыдущих равенств находим

$$\boldsymbol{x}A\boldsymbol{y} = \boldsymbol{y}A\boldsymbol{x},$$

что и доказывает симметрию преобразования.

63. Наша главная цель — доказать, что: 1) *каждое симметрическое линейное преобразование на плоскости имеет по* *крайней мере одну пару собственных векторов, которые перпендикулярны друг к другу;* 2) *каждое симметрическое линейное преобразование в пространстве имеет по крайней мере одну тройку попарно перпендикулярных собственных векторов.* Вместе с тем мы имеем в виду показать, как практически находить такие пары или тройки собственных векторов. В остальных пунктах этого параграфа поставленный вопрос будет решен для симметрических линейных преобразований на плоскости.

64. Пусть на плоскости задано симметрическое линейное преобразование $\boldsymbol{x}' = A\boldsymbol{x}$;

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

— его матрица в некотором ортогональном базисе $\boldsymbol{e_1}$, $\boldsymbol{e_2}$. Составим характеристическое уравнение матрицы A:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0. \quad \textbf{(4)}$$

Ввиду симметрии матрицы A имеем $a_{21} = a_{12}$. Поэтому уравнение (4) запишется так:

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}^2) = 0.$$

Отсюда получаем корни уравнения (4):

$$\lambda_{1,2} = \frac{a_{11} + a_{22} \pm \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2)}}{2}. \qquad (5)$$

Заметим, что

$$(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2) = (a_{11} - a_{22})^2 + 4a_{12}^2 \geqslant 0;$$

следовательно, числа λ_1, λ_2 вещественны. Итак, *если матрица симметрична, то характеристическое уравнение этой матрицы имеет только вещественные корни.*

Согласно п°57 для каждого корня ($\lambda=\lambda_1$ или $\lambda=\lambda_2$) характеристического уравнения (4) найдется собственный вектор $\boldsymbol{a} = \{l;\ m\}$, имеющий корень λ своим собственным числом; координаты вектора $\boldsymbol{a} = \{l;\ m\}$ даются системой уравнений

$$\left. \begin{aligned} (a_{11} - \lambda)\, l + a_{12} m = 0, \\ a_{21} l + (a_{22} - \lambda)\, m = 0. \end{aligned} \right\} \qquad (6)$$

Условимся решение l, m системы (6) называть *нормированным,* если $l^2 + m^2 = 1$; если l, m — нормированное решение, то $\boldsymbol{a} = \{l;\ m\}$ — единичный вектор. Мы будем обозначать дальше через l_1, m_1

нормированное решение системы (6), получаемое при $\lambda=\lambda_1$; аналогично через l_2, m_2 будет обозначаться нормированное решение, получаемое при $\lambda=\lambda_2$. Тогда

$$a_1=\{l_1;\ \ m_1\} \quad \text{и} \quad a_2=\{l_2;\ \ m_2\}$$

— единичные собственные векторы с собственными числами λ_1 и λ_2.

65 Рассмотрим случай, когда корни характеристического уравнения различны: $\lambda_1\neq\lambda_2$. Докажем, что *в этом случае векторы* $\boldsymbol{a}_1$ *и* $\boldsymbol{a}_2$ *перпендикулярны друг к другу.* Для доказательства напишем соотношения

$$Aa_1=\lambda_1 a_1, \quad Aa_2=\lambda_2 a_2, \tag{7}$$

которые выражают тот факт, что $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ являются собственными векторами с собственными числами λ_1 и λ_2. Умножим скалярно обе части первого равенства (7) на вектор $\boldsymbol{a}_2$, обе части второго — на вектор $\boldsymbol{a}_1$.

Мы получим

$$a_2Aa_1=\lambda_1 a_1a_2, \quad a_1Aa_2=\lambda_2 a_1a_2.$$

Вследствие симметрии преобразования имеем

$$a_2Aa_1=a_1Aa_2.$$

Отсюда и на основании предыдущих равенств

$$(\lambda_1-\lambda_2)\,a_1a_2=0. \tag{8}$$

Так как $\lambda_1\neq\lambda_2$, то из (8) следует $\boldsymbol{a}_1\boldsymbol{a}_2=0$, т. е. перпендикулярность векторов $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$. Тем самым наше утверждение доказано.

66. Предположим теперь, что характеристическое уравнение (4) имеет равные корни: $\lambda_1=\lambda_2$, будем обозначать их одной буквой λ. Из формулы (5) следует, что в этом случае $(a_{11}-a_{22})^2+4a_{12}^2=0$ и

$$\lambda=\frac{a_{11}+a_{22}}{2}.$$

Но тогда $a_{11}-a_{22}=0$, $a_{12}=0$; значит, $a_{11}=a_{22}=\lambda$, и преобразование представляется формулами $x'_1=\lambda x_1$, $x'_2=\lambda x_2$. Такое преобразование, как мы знаем, есть подобие с коэффициентом λ. Если $\boldsymbol{x}'=A\boldsymbol{x}$ — подобие, то каждый вектор является собственным вектором с собственным числом λ. В качестве пары взаимно перпендикулярных собственных векторов можно взять любую пару ненулевых векторов $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, лишь бы они были перпендикулярны друг к другу.

67. Итак, *всякое симметрическое линейное преобразование на плоскости имеет пару взаимно перпендикулярных собственных векторов* $\boldsymbol{a}_1$, $\boldsymbol{a}_2$.

Если эти векторы принять в качестве базисных: $\bar{e}_1 = a_1$, $\bar{e}_2 = a_2$, то матрица преобразования получит диагональный вид

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

где λ_1, λ_2 — собственные числа векторов $\boldsymbol{a}_1$, $\boldsymbol{a}_2$. Отсюда следует, что *каждое симметрическое линейное преобразование на плоскости является произведением двух сжатий по двум перпендикулярным направлениям*.

68. Сформулируем теперь в окончательном виде план решения задачи о выборе ортогонального базиса из собственных векторов симметрического линейного преобразования $\boldsymbol{x}' = A\boldsymbol{x}$.

Пусть A — матрица данного преобразования в любом ортогональном базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$. Нужно прежде всего составить характеристическое уравнение матрицы A и найти его корни $\lambda = \lambda_1$, $\lambda = \lambda_2$ (λ_1, λ_2 будут вещественными числами). Возможны два случая:

1) $\lambda_1 \neq \lambda_2$. В этом случае следует подставить в систему (6) $\lambda = \lambda_1$ затем $\lambda = \lambda_2$ и каждый раз найти нормированное решение этой системы l_1 m_1 и l_2, m_2. Векторы

$$\boldsymbol{a}_1 = \{l_1;\ m_1\}, \quad \boldsymbol{a}_2 = \{l_2;\ m_2\}$$

будут единичными собственными векторами с собственными числами λ_1 и λ_2; $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ обязательно окажутся перпендикулярными друг к другу. Искомый базис определяется единственным образом: $\tilde{e}_1 = \boldsymbol{a}_1$, $\tilde{e}_2 = \boldsymbol{a}_2$ (если не говорить о возможности изменить нумерацию векторов $\bar{e}_1$, $\bar{e}_2$ или заменить их противоположными по направлению).

2) $\lambda_1 = \lambda_2$. В этом случае преобразование является подобием и в качестве $\tilde{e}_1$, $\bar{e}_2$ можно взять любую пару единичных и перпендикулярных друг к другу векторов.

При переходе к новому ортогональному базису $\tilde{e}_1$, $\bar{e}_2$ координаты всех векторов данной плоскости преобразуются по формулам п.6.5, где

$$L = \begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix};$$

матрица L — ортогональная.

Пример. В ортогональном базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ дано линейное преобразование $\boldsymbol{x}' = A\boldsymbol{x}$:

$$x_1' = 6x_1 + 2x_2,$$
$$x_2' = 2x_1 + 3x_2$$

Привести матрицу этого преобразования к диагональному виду путем перехода к новому ортогональному базису.

Решение. Матрица данного преобразования симметрична:

$$A = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix},$$

следовательно, поставленная задача может быть решена по изложенному выше плану. Составим характеристическое уравнение;

$$\begin{vmatrix} 6-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = 0,$$

или

$$\lambda^2 - 9\lambda + 14 = 0.$$

Корни этого уравнения суть $\lambda_1=2$, $\lambda_2=7$. Соответствующие собственные векторы найдем из системы (6) п° 64; в данном случае эта система имеет вид

$$(6-\lambda)\,l + 2m = 0,$$
$$2l + (3-\lambda)\,m = 0.$$

Подставляя сюда $\lambda = \lambda_1 = 2$, получим

$$4l + 2m = 0,$$
$$2l + m = 0.$$

В качестве решения можно взять $l = 1$, $m = -2$. Нормируя это решение, найдем единичный собственный вектор с собственным числом $\lambda_1=2$:

$$\tilde{e}_1 = \left\{ \frac{1}{\sqrt{5}} ;\ \frac{-2}{\sqrt{5}} \right\}.$$

Аналогично найдем единичный собственный вектор с собственным числом $\lambda_2=7$:

$$\tilde{e}_2 = \left\{ \frac{2}{\sqrt{5}} ;\ \frac{1}{\sqrt{5}} \right\}.$$

При переходе к базису $\tilde{e}_1, \tilde{e}_2$ координаты всех векторов (и точек) преобразуются по формулам п.6.5 с матрицей

$$L=\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix};$$

именно

$$\left.\begin{aligned} x_1 &= \frac{1}{\sqrt{5}}\tilde{x}_1+\frac{2}{\sqrt{5}}\tilde{x}_2, \\ x_2 &= -\frac{2}{\sqrt{5}}\tilde{x}_1+\frac{1}{\sqrt{5}}\tilde{x}_2, \end{aligned}\right\} \qquad \left.\begin{aligned} \tilde{x}_1 &= \frac{1}{\sqrt{5}}x_1-\frac{2}{\sqrt{5}}x_2, \\ \tilde{x}_2 &= \frac{2}{\sqrt{5}}x_1+\frac{1}{\sqrt{5}}x_2. \end{aligned}\right\}.$$

В базисе $\tilde{e}_1,\ \tilde{e}_2$ матрица данного линейного преобразования имеет вид

$$\tilde{A}=\begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix};$$

само преобразование представляется формулами

$$\begin{aligned} \tilde{x}'_1 &= 2\tilde{x}_1, \\ \tilde{x}'_2 &= \qquad 7\tilde{x}_2. \end{aligned}$$

Заметим, что если расположение векторов $\tilde{e}_1,\ \tilde{e}_2$ нас не интересует, то задача решается короче: достаточно знать, что $\lambda_1=2$, $\lambda_2=7$, чтобы написать $\tilde{A}$ сразу.

6.12. Приведение к диагональному виду матрицы симметрического линейного преобразования в пространстве

69. Пусть $\boldsymbol{x}'=A\boldsymbol{x}$ — симметрическое линейное преобразование в пространстве. Требуется доказать, что оно имеет три попарно перпендикулярных собственных вектора.

Выберем какой-нибудь ортогональный базис $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$. В этом базисе данное преобразование получит координатное представление с симметричной матрицей A. Составим характеристическое уравнение матрицы A. Для симметрических преобразований в пространстве, как и на плоскости, характеристическое уравнение имеет только вещественные корни. Однако в пространственном случае этот факт доказывается не очень просто и мы пока не будем на него опираться. С другой стороны, совсем легко доказать, что характеристическое уравнение матрицы A имеет хотя бы один вещественный корень. Дело в том, что характеристическое уравнение матрицы линейного преобразования в пространстве есть уравнение третьей степени, а **всякое уравнение <u>третьей степени</u> с вещественными коэффициентами имеет вещественный корень**.

Чтобы убедиться в этом, напишем произвольное уравнение третьей степени $\lambda^3+a\lambda^2+b\lambda+c=0$. Ясно, что при достаточно больших по абсолютной величине значениях λ первый член этого уравнения будет подавляюще велик по сравнению с остальными членами. Поэтому, если $Q>0$ — достаточно большое число, то при $\lambda=-Q$ левая часть уравнения будет отрицательна, а при $\lambda=+Q$—положительна. Если λ, возрастает от $-Q$ до $+Q$, то левая часть уравнения, непрерывно изменяясь, сменит знак. Следовательно, при некотором значении $\lambda=\lambda_1$ левая часть уравнения станет равной нулю. Тем самым доказано наличие вещественного корня λ.

Согласно п° 55 по любому вещественному корню λ характеристического уравнения находится собственный вектор линейного преобразования с собственным числом λ (путем решения системы (8) п° 55).

Отсюда приходим к важному выводу: *всякое линейное преобразование в пространстве имеет хотя бы один собственный вектор.*

В дальнейших рассуждениях мы уже существенно воспользуемся условием, что данное нам линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ является симметрическим.

По доказанному один собственный вектор данного преобразования обязательно найдется; обозначим его через $\boldsymbol{a}_1$. Пусть λ_1— собственное число этого вектора. Проведем через начало координат плоскость α_1, перпендикулярную к вектору $\boldsymbol{a}_1$. Возьмем в плоскости α_1 произвольный вектор $\boldsymbol{x}$; убедимся, что его образ $\boldsymbol{x}'=A\boldsymbol{x}$ также лежит в плоскости α_1. Чтобы установить этот факт, заметим, что скалярное произведение $\boldsymbol{a}_1\boldsymbol{x}=0$ (так как вектор $\boldsymbol{x}$ мы взяли в плоскости α_1). Подсчитаем скалярное произведение $\boldsymbol{a}_1\boldsymbol{x}'$. Имеем

$$\boldsymbol{a}_1\boldsymbol{x}'=\boldsymbol{a}_1A\boldsymbol{x}=\boldsymbol{x}A\boldsymbol{a}_1.$$

Последнее равенство справедливо, поскольку преобразование является симметрическим. Так как $\boldsymbol{a}_1$ — собственный вектор с собственным числом λ_1 то $A\boldsymbol{a}_1=\lambda_1\boldsymbol{a}_1$; поэтому

$$\boldsymbol{a}_1\boldsymbol{x}'=\boldsymbol{x}A\boldsymbol{a}_1=\boldsymbol{x}(\lambda_1\boldsymbol{a}_1)=\lambda_1(\boldsymbol{a}_1\boldsymbol{x})=0.$$

Отсюда следует, что $\boldsymbol{x}'$ перпендикулярен к $\boldsymbol{a}_1$, т. е. лежит в плоскости α_1, что и утверждалось.

На основании только что доказанного мы можем рассматривать преобразование $\boldsymbol{x}' = A\boldsymbol{x}$ как преобразование векторов плоскости α_1, если в качестве $\boldsymbol{x}$ брать только векторы этой плоскости (поскольку образы таких векторов остаются в той же плоскости). Но согласно п.6.11 всякое линейное симметрическое преобразование на плоскости имеет пару перпендикулярных собственных векторов. Следовательно, и данное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ имеет в плоскости α_1 пару перпендикулярных собственных векторов; обозначим их через $\boldsymbol{a}_2$, $\boldsymbol{a}_3$. Вместе с тем мы получаем в пространстве три вектора $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, каждый из которых является собственным вектором и перпендикулярен к двум другим.

Итак, мы доказали, что *всякое симметрическое линейное преобразование в пространстве имеет хотя бы одну тройку попарно перпендикулярных собственных векторов.*

70. Предполагая векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ единичными, примем их в качестве нового ортогонального базиса: $\tilde{\boldsymbol{e}}_1 = \boldsymbol{a}_1$, $\tilde{\boldsymbol{e}}_2 = \boldsymbol{a}_2$, $\tilde{\boldsymbol{e}}_3 = \boldsymbol{a}_3$. Матрица данного преобразования получит в этом базисе диагональный вид

$$\tilde{A} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

где λ_1, λ_2, λ_3 —собственные числа векторов $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$.

Отсюда заключаем: *каждое симметрическое линейное преобразование в пространстве есть произведение трех сжатий по трем попарно перпендикулярным направлениям* (см. п° 38). В частности, если $\lambda_1 = \lambda_2 = \lambda_3$, то преобразование является подобием (см. п° 52).

71. Составим характеристический многочлен матрицы *А:*

$$\mathrm{Det}\,(\tilde{A} - \lambda E) = \begin{vmatrix} \lambda_1 - \lambda & 0 & 0 \\ 0 & \lambda_2 - \lambda & 0 \\ 0 & 0 & \lambda_3 - \lambda \end{vmatrix} = \\ = -(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3).$$

Очевидно, числа λ_1, λ_2, λ_3 являются корнями этого многочлена. Вспомним теперь теорему п° 57, которая утверждает, что характеристический многочлен матрицы линейного преобразования не зависит от выбора базиса. Из этой теоремы следует, что в любом базисе характеристический многочлен матрицы данного симметрического преобразования имеет своими корнями числа λ_1, λ_2, λ_3. Тем самым мы доказали, что *характеристическое* *уравнение матрицы симметрического линейного преобразования имеет только действительные корни.*

72. Пусть $\lambda_1, \lambda_2, \lambda_3$ — корни характеристического уравнения. Рассмотрим следующие случаи.

1) Числа $\lambda_1, \lambda_2, \lambda_3$ — все разные. Согласно п° 52 найдутся собственные векторы $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}$ с собственными числами $\lambda_1, \lambda_2, \lambda_3$. Эти векторы обязательно будут перпен дикулярными друг к другу (см. п° 65). Мы можем считать векторы $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}$ единичными. При этом условии тройка векторов $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}$ единственна с точностью до знаков. В самом деле, никакой вектор, кроме $\pm \boldsymbol{a_1}$, не может быть единичным собственным вектором с собственным числом λ_1 так как он не будет перпендикулярен к векторам $\boldsymbol{a_2}, \boldsymbol{a_3}$. По той же причине, кроме $\pm \boldsymbol{a_2}$ и $\pm \boldsymbol{a_3}$, нет других единичных собственных векторов с собственными числами λ_2 и λ_3.

2) Среди чисел $\lambda_1, \lambda_2, \lambda_3$ — два одинаковых; пусть, например, $\lambda_1 \neq \lambda_2 = \lambda_3$. Обозначим $\lambda_2 = \lambda_3$ одной буквой λ. В данном случае λ — двукратный корень характеристического уравнения.

Докажем, что *имеется определенный с точностью до знака единичный собственный вектор* $\boldsymbol{a_1}$ *с собственным числом* λ_1. Наоборот, *двукратному корню* λ *соответствует бесконечное множество единичных собственных векторов с одним и тем же собственным числом* λ. Именно, каждый вектор, перпендикулярный к вектору $\boldsymbol{a_1}$, будет собственным вектором данного линейного преобразования с собственным числом λ.

Для доказательства этого утверждения вернемся к п° 69, согласно которому данное симметрическое преобразование $\boldsymbol{x}' = A\boldsymbol{x}$ имеет тройку попарно перпендикулярных собственных векторов; обозначим их через $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}$ (считая для удобства изложения единичными). На основании п° 70 и 71 заключаем, что в рассматриваемом случае два из этих трех векторов должны иметь одно и то же собственное число $\lambda\,(\lambda = \lambda_2 = \lambda_3)$; пусть это будут векторы $\boldsymbol{a_2}$ и $\boldsymbol{a_3}$. Тогда $\boldsymbol{a_1}$ имеет собственное число $\lambda_1\,(\lambda_1 \neq \lambda)$. Обозначим через α_1 плоскость, проходящую через начало координат и перпендикулярную к вектору $\boldsymbol{a_1}$. Плоскость α_1 содержит векторы $\boldsymbol{a_2}$ и $\boldsymbol{a_3}$. В п° 69 показано, что данное преобразование можно рассматривать как линейное преобразование векторов плоскости α_1. Но такое преобразование является подобием (на плоскости α_1), поскольку имеет два взаимно перпендикулярных собственных вектора ($\boldsymbol{a_2}$ и $\boldsymbol{a_3}$) с одинаковыми собственными числами (см. п° 66). Следовательно, все векторы, лежащие в плоскости α_1, являются собственными векторами данного преобразования с одним и тем же собственным числом λ.

Остается показать, что никакой вектор, кроме $\pm\boldsymbol{a}_1$, не может быть единичным собственным вектором с собственным числом λ (λ_1— простой корень характеристического уравнения). Но это ясно, так как всякий собственный вектор, имеющий собственное число λ_1, должен быть перпендикулярным к плоскости α_1 (вспомним, что в α_1 лежат собственные векторы с собственным числом λ и что $\lambda_1 \neq \lambda$). Ясно также, что единичный вектор, перпендикулярный к α_1 есть либо $\boldsymbol{a}_1$, либо $-\boldsymbol{a}_1$. Утверждение доказано полностью.

3) Все числа λ_1, λ_2, λ_3 одинаковы, т. е. $\lambda=\lambda_1=\lambda_2=\lambda_3$ есть трехкратный корень характеристического уравнения.

В этом случае данное преобразование является подобием в пространстве с коэффициентом λ (см. п° 70). Следовательно, каждый вектор пространства является собственным вектором данного преобразования с собственным числом λ (см. п° 52).

73. После всего изложенного можно указать следующий план решения задачи о нахождении ортогонального базиса из собственных векторов симметрического линейного преобразования в пространстве.

Пусть $\boldsymbol{x}'=A\boldsymbol{x}$ — симметрическое преобразование в пространстве,

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

— его матрица в каком-нибудь ортогональном базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ (A симметрична, т. е. $a_{21}=a_{12}$, $a_{31}=a_{13}$, $a_{32}=a_{23}$).

Для решения нашей задачи нужно прежде всего составить характеристическое уравнение матрицы A:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}=0, \tag{1}$$

затем найти корни этого уравнения

$$\lambda=\lambda_1,\ \lambda=\lambda_2,\ \lambda=\lambda_3$$

(числа λ_1, λ_2, λ_3 вещественны).

Согласно п° **55** для каждого корня $\lambda\ (=\lambda_1,\ \lambda_2,\ \lambda_3)$ характеристического уравнения найдется собственный вектор $\boldsymbol{a}=\{l;\ m;\ n\}$, имеющий корень λ своим собственным числом;

координаты вектора $\boldsymbol{a}=\{l;\ m;\ n\}$ даются системой уравнений

$$\left.\begin{aligned}(a_{11}-\lambda)\,l+a_{12}m+a_{13}n=0,\\ a_{21}l+(a_{22}-\lambda)\,m+a_{23}n=0,\\ a_{31}l+a_{32}m+(a_{33}-\lambda)\,n=0\end{aligned}\right\}\qquad(2)$$

(причем система (2) заведомо имеет ненулевое решение l, m, n, если λ — корень уравнения (1)).

Дальнейшие действия сводятся к нахождению нужных решений системы (2) при $\lambda=\lambda_1$, $\lambda=\lambda_2$, $\lambda=\lambda_3$.

Условимся решение l, m, n системы (2) называть *нормированным,* если $l^2+m^2+n^2=1$. Если l, m, n — нормированное решение, то $\boldsymbol{a}=\{l;\ m;\ n\}$ — единичный вектор. Мы будем обозначать дальше через l_1, m_1, n_1 нормированное решение системы (2), получаемое при $\lambda=\lambda_1$, аналогично через l_2, m_2, n_2 и l_3, m_3, n_3 будут обозначаться нормированные решения, которые получаются при $\lambda=\lambda_2$ и $\lambda=\lambda_3$. Тогда

$$\boldsymbol{a}_1=\{l_1;\ m_1;\ n_1\},\quad \boldsymbol{a}_2=\{l_2;\ m_2;\ n_2\},\quad \boldsymbol{a}_3=\{l_3;\ m_3;\ n_3\}$$

— λ_1, λ_2, λ_3 единичные собственные векторы с собственными числами

Возможны три случая:

1) λ_1, λ_2, λ_3 — все разные. В этом случае векторы $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ обязательно окажутся перпендикулярными друг к другу; тройка векторов $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ определится единственным образом сточностью до знаков (см. п° 72). Искомый базис: $\tilde{\boldsymbol{e}}_1=\boldsymbol{a}_1$, $\tilde{\boldsymbol{e}}_2=\boldsymbol{a}_2$, $\tilde{\boldsymbol{e}}_3=\boldsymbol{a}_3$.

2) Среди чисел λ_1, λ_2, λ_3 —два одинаковых; например, $\lambda_1\neq\lambda_2=\lambda_3$.

В этом случае согласно п° 72 числу $\lambda=\lambda_2=\lambda_3$ соответствует бесконечное множество единичных собственных векторов, причем все они лежат в одной плоскости. Следовательно, если мы подставим в систему (2) число $\lambda=\lambda_2=\lambda_3$, то в этой системе будет только одно существенное уравнение (два других должны быть ему пропорциональны). Допустим, что существенным уравнением системы (2) является

$$(a_{11}-\lambda)\,l+a_{12}m+a_{13}n=0;\qquad(3)$$

когда мы говорим, что уравнение {3) существенно, то имеемв виду, что оно —не тождество, т. е. что среди его коэффициентов a_{12} — λ, a_{12}, a_{13} имеются отличные от нуля.

Всякое ненулевое решение *l, m, n* уравнения (3) определяет собственный вектор $\{l;\ m;\ n\}$ с собственным числом $\lambda\,(\lambda = \lambda_2 = \lambda_3)$. Уравнение (3) означает, что все такие векторы перпендикулярны к вектору $\boldsymbol{a} = \{a_{11} - \lambda;\ a_{12};\ a_{13}\}$. Следовательно, вектор $\boldsymbol{a}$ есть собственный вектор с собственным числом λ_1 (λ_1 — простой корень характеристического уравнения). Задача может быть завершена следующим способом. Разделим вектор $\boldsymbol{a}$ на его модуль; мы получим единичный собственный вектор $\boldsymbol{a}_1$ с собственным числом λ_1. Затем найдем любое нормированное решение l_2, m_2, n_2 уравнения (3). Вектор $\boldsymbol{a}_2 = \{l_2;\ m_2;\ n_2\}$ будет единичным собственным вектором с собственным числом $\lambda\,(\lambda = \lambda_2 = \lambda_3)$. Умножим векторно $\boldsymbol{a}_1$ на $\boldsymbol{a}_2$; мы получим единичный вектор $\boldsymbol{a}_3$, перпендикулярный к вектору $\boldsymbol{a}_1$. Следовательно, $\boldsymbol{a}_3$ есть собственный вектор с собственным числом $\lambda = \lambda_2 = \lambda_3$; кроме того, a_3 перпендикулярен к вектору $\boldsymbol{a}_2$. Таким образом, искомый базис получен:

$$\tilde{\boldsymbol{e}}_1 = \boldsymbol{a}_1,\ \ \tilde{\boldsymbol{e}}_2 = \boldsymbol{a}_2,\ \ \tilde{\boldsymbol{e}}_3 = \boldsymbol{a}_3.$$

3) Числа $\lambda_1, \lambda_2, \lambda_3$ всео динаковы, т. е. — $\lambda = \lambda_1 = \lambda_2 = \lambda_3$ трехкратный корень характеристического уравнения.

В этом случае данное преобразование есть подобие в пространстве с коэффициентом λ. Для такого преобразования все векторы пространства являются собственными векторами с собственным числом λ. В качестве искомого базиса можно взять любую тройку единичных попарно перпендикулярных векторов $\tilde{\boldsymbol{e}}_1,\ \tilde{\boldsymbol{e}}_2,\ \tilde{\boldsymbol{e}}_3$ (например, первоначально данную тройку $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$.

74. Мы показали, как для любого симметрического преобразования в пространстве найти базис, состоящий из единичных перпендикулярных друг к другу собственных векторов $\tilde{\boldsymbol{e}}_1,\ \tilde{\boldsymbol{e}}_2,\ \tilde{\boldsymbol{e}}_3$. Если

$$\tilde{\boldsymbol{e}}_1 = \{l_1;\ m_1;\ n_1\},\ \ \tilde{\boldsymbol{e}}_2 = \{l_2;\ m_2;\ n_2\},\ \ \tilde{\boldsymbol{e}}_3 = \{l_3;\ m_3;\ n_3\},$$

то при переходе к этому оазису координаты всех векторов пространства преобразуются по формулам п. 6.8, где

$$L = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{pmatrix};$$

матрица *L* — ортогональная.

75. Пример. В ортогональном базисе e_1, e_2, e_3 дано линейное преобразование:

$$\begin{aligned} x_1' &= 7x_1 - 2x_2, \\ x_2' &= -2x_1 + 6x_2 - 2x_3, \\ x_3' &= -2x_2 + 5x_3. \end{aligned}$$

Путем перехода к новому ортогональному базису привести матрицу данного преобразования к диагональному виду.

Решение. Составим матрицу преобразования в данном базисе:

$$A = \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}.$$

Матрица *A* симметрична, следовательно, задача может быть решена по изложенному выше плану.

Имеем характеристическое уравнение

$$\begin{vmatrix} 7-\lambda & -2 & 0 \\ -2 & 6-\lambda & -2 \\ 0 & -2 & 5-\lambda \end{vmatrix} = 0,$$

или $\lambda^3 - 18\lambda^2 + 99\lambda - 162 = 0.$ Корни этого уравнения $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$. Корни различны, значит, мы находимся в условиях первого случая п° 73. Напишем по данным задачи систему (2):

$$\begin{aligned} (7-\lambda)\, l - 2m &= 0, \\ -2l + (6-\lambda)\, m - 2n &= 0, \\ -2m + (5-\lambda)\, n &= 0. \end{aligned}$$

Подставляя сюда поочередно $\lambda = 3, 6, 9$ и каждый раз находя нормированные решения этой системы, получим три единичных вектора:

$$\begin{aligned} \tilde{e}_1 &= \{l_1;\ m_1;\ n_1\} = \left\{\frac{1}{3};\ \frac{2}{3};\ \frac{2}{3}\right\}, \\ \tilde{e}_2 &= \{l_2;\ m_2;\ n_2\} = \left\{\frac{2}{3};\ \frac{1}{3};\ -\frac{2}{3}\right\}, \\ \tilde{e}_3 &= \{l_3;\ m_3;\ n_3\} = \left\{-\frac{2}{3};\ \frac{2}{3};\ -\frac{1}{3}\right\}. \end{aligned}$$

Векторы $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_3$ являются собственными векторами данного преобразования с собственными числами $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$. В базисе $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_3$ данное преобразование имеет матрицу

$$\tilde{A}=\begin{pmatrix}3&0&0\\0&6&0\\0&0&9\end{pmatrix};$$

координатное представление

$$\begin{aligned}\tilde{x}_1' &= 3\tilde{x}_1,\\ \tilde{x}_2' &= \qquad 6\tilde{x}_2,\\ \tilde{x}_3' &= \qquad\qquad 9\tilde{x}_3.\end{aligned}$$

При переходе к базису $\tilde{e}_1$, $\tilde{e}_2$, $\tilde{e}_3$ координаты всех векторов могут быть пересчитаны с помощью матричного равенства

$$\begin{pmatrix}\tilde{x}_1\\ \tilde{x}_2\\ \tilde{x}_3\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\\ \frac{2}{3}&\frac{1}{3}&-\frac{2}{3}\\ -\frac{2}{3}&\frac{2}{3}&-\frac{1}{3}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}.$$

Пример. В ортогональном базисе e_1, e_2, e_3 дано линейное преобразование

$$\begin{aligned}x_1' &= x_1+2x_2-4x_3,\\ x_2' &= 2x_1-2x_2-2x_3,\\ x_3' &= -4x_1-2x_2+x_3.\end{aligned}$$

Путем перехода к новому ортогональному базису привести матрицу данного преобразования к диагональному виду.

Решение. В данном базисе имеем симметричную матрицу

$$A=\begin{pmatrix}1&2&-4\\2&-2&-2\\-4&-2&1\end{pmatrix}.$$

Следовательно, задачу можно решить по плану п° 73. Составим характеристическое уравнение

$$\begin{vmatrix}1-\lambda&2&-4\\2&-2-\lambda&-2\\-4&-2&1-\lambda\end{vmatrix}=0,$$

или $\lambda^3-27\lambda-54=0$. Корни этого уравнения $\lambda_1=6$, $\lambda_2=-3$, $\lambda_3=-3$. Среди корней — два одинаковых; следовательно, мы находимся в условиях второго случая п° 73. Напишем по данным задачи систему (2):

$$\begin{aligned}(1-\lambda)\,l+{}&2m&-4n=0,\\ 2\,l+{}&(-2-\lambda)\,m&-2n=0,\\ -4\,l&-2m+(1-\lambda)\,n=0.\end{aligned}$$

Подставим сюда двукратный корень $\lambda=\lambda_2=\lambda_3=-3$; мы получим

$$\begin{aligned}4l+2m-4n=0,\\ 2l+m-2n=0,\\ -4l-2m+4n=0.\end{aligned}$$

Система свелась к одному существенному уравнению:

$$2l+m-2n=0. \qquad (*)$$

Согласно п° 73 заключаем, что вектор {2; 1; —2}, координатами которого служат коэффициенты уравнения (*), является собственным вектором с собственным числом $\lambda_1 = 6$; деля этот вектор на его модуль, получим единичный вектор:

$$\tilde{e}_1=\left\{\frac{2}{3};\ \frac{1}{3};\ -\frac{2}{3}\right\}.$$

Возьмем теперь любое решение уравнения (*), например l=1, m = 2, n = 2; нормируя это решение, получим вектор

$$\tilde{e}_2=\left\{\frac{1}{3};\ \frac{2}{3};\ \frac{2}{3}\right\}.$$

Вектор $\tilde{e}_2$ является единичным собственным вектором с собственным числом —3. Умножая векторно $\tilde{e}_1$ на $\tilde{e}_2$, найдем еще один единичный собственный вектор с собственным числом —3:

$$\tilde{e}_3=\left\{\frac{2}{3};\ -\frac{2}{3};\ \frac{1}{3}\right\}.$$

Векторы $\tilde{e}_1, \tilde{e}_2, \tilde{e}_3$ составляют ортогональный базис. После перехода к этому базису матрица данного преобразования примет диагональный вид

$$\tilde{A}=\begin{pmatrix}6&0&0\\0&-3&0\\0&0&-3\end{pmatrix}.$$

Замечание. Мы имеем

$$\begin{pmatrix}6&0&0\\0&-3&0\\0&0&-3\end{pmatrix}=\begin{pmatrix}6&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-3&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-3\end{pmatrix}.$$

Следовательно, преобразование, которое дано в рассмотренной задаче, есть произведение трех сжатий: 1) к плоскости $O\tilde{e}_2\tilde{e}_3$ с коэффициентом

6; 2) к плоскости $O\tilde{e}_1\tilde{e}_3$ с коэффициентом —3; 3) к плоскости $O\tilde{e}_1\tilde{e}_2$ с коэффициентом —3. Заметим прежде всего, что коэффициенты этих сжатий по модулю больше единицы; поэтому фактически мы имеем дело с растяжениями пространства (например, в первом случае каждая точка удаляется от плоскости $O\tilde{e}_2\tilde{e}_3$ так, что ее первоначальное расстояние от этой плоскости увеличивается в шесть раз). Кроме того, во втором и в третьем случаях коэффициент сжатия является отрицательным числом. Чтобы пояснить, что это означает геометрически, рассмотрим, например, преобразование, которое определяется последней матрицей в правой части (**), его координатное представление:

$$\begin{aligned} \tilde{x}'_1 &= \tilde{x}_1, \\ \tilde{x}'_2 &= \tilde{x}_2, \\ \tilde{x}'_3 &= -3\tilde{x}_3. \end{aligned}$$

Отсюда видно, что при таком преобразовании не только меняется расстояние от точки до плоскости $O\tilde{e}_1\tilde{e}_2$ (в три раза), но образ $M'(\tilde{x}'_1;\ \tilde{x}'_2;\ \tilde{x}'_3)$ и прообраз $M(\tilde{x}_1;\ \tilde{x}_2;\ \tilde{x}_3)$ оказываются по разные стороны от этой плоскости.

6.13. Приведение к каноническому виду квадратичной формы.

133. *Квадратичной формой от двух переменных x_1, x_2 называется однородный многочлен второй степени относительно этих переменных:*

$$\Phi = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2. \qquad (1)$$

Здесь a_{11}, a_{12}, a_{22} — числа, задание которых определяет форму; их называют *коэффициентами* формы. Вместо a_{12} можно писать a_{21} считая $a_{21}=a_{12}$. Пусть на некоторой плоскости введена система координат с началом O и масштабными векторами $\mathbf{e_1}$, $\mathbf{e_2}$. Тогда x_1, x_2 можно рассматривать как координаты некоторой точки M, а число Ф — как значение формы в точке $M(x_1;\ x_2)$. Величины x_1, x_2 мы будем рассматривать также как координаты вектора $\mathbf{x} = \overline{OM}$.

77. Предположим, что базис $\mathbf{e_1}$, $\mathbf{e_2}$ заменяется новым базисом $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$. При этом координаты всех точек и векторов изменятся по формулам п.6.5. Величина Ф выразится через новые координаты

$\bar{x}_1, \bar{x}_2$ точки M в виде некоторой новой квадратичной формы от $\tilde{x}_1, \tilde{x}_2$. Коэффициенты получаемой таким путем формы определенным образом выражаются через коэффициенты первоначально данной формы и через коэффициенты формул преобразования координат. Наша ближайшая цель — найти эти выражения.

Для упрощения дела условимся употреблять только ортогональные базисы (из единичных векторов).

78. Перепишем данную квадратичную форму следующим образом:

$$\Phi = x_1(a_{11}x_1 + a_{12}x_2) + x_2(a_{21}x_1 + a_{22}x_2). \tag{2}$$

Составим матрицу из коэффициентов формы, беря строчки соответственно скобкам в правой части последнего равенства:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{3}$$

Эта матрица называется *матрицей квадратичной формы* (1) в заданном базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$. Очевидно, матрица A симметрична $(a_{21} = a_{12})$. Вместе с данной формой рассмотрим линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$, которое в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2$ имеет матрицу A; координатным представлением преобразования $\boldsymbol{x}' =A\boldsymbol{x}$ будет

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2, \\ x_2' &= a_{21}x_1 + a_{22}x_2. \end{aligned}\right\} \tag{4}$$

Так как $\boldsymbol{e}_1, \boldsymbol{e}_2$ — ортогональный базис, а матрица A симметрична, то и преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ является симметрическим. Выражение (2) теперь можем переписать так:

$$\Phi = x_1x_1' + x_2x_2'. \tag{5}$$

Вследствие ортогональности базиса $\boldsymbol{e}_1, \boldsymbol{e}_2$ правая часть (5) представляет собой не что иное, как скалярное произведение вектора $\boldsymbol{x}$ на вектор $\boldsymbol{x}'=A\boldsymbol{x}$:

$$\Phi = \boldsymbol{x}\boldsymbol{x}' = \boldsymbol{x}A\boldsymbol{x}. \tag{6}$$

Перейдем теперь к новому ортогональному базису $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2$:

$$\left.\begin{aligned} \tilde{\boldsymbol{e}}_1 &= l_1\boldsymbol{e}_1 + m_1\boldsymbol{e}_2, \\ \tilde{\boldsymbol{e}}_2 &= l_2\boldsymbol{e}_1 + m_2\boldsymbol{e}_2. \end{aligned}\right\} \tag{7}$$

Как обычно, обозначим через L матрицу, которая получается транспонированием матрицы коэффициентов (7):

$$L=\begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix}.$$

Матрица L ортогональна, т.е. $L^{-1}=L^*$. В новом базисе линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$ будет иметь новое координатное представление:

$$\left.\begin{aligned} \tilde{x}_1' &= \tilde{a}_{11}\tilde{x}_1+\tilde{a}_{12}\tilde{x}_2, \\ \tilde{x}_2' &= \tilde{a}_{21}\tilde{x}_1+\tilde{a}_{22}\tilde{x}_2, \end{aligned}\right\} \tag{8}$$

с матрицей

$$\tilde{A}=\begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{pmatrix}. \tag{9}$$

Вследствие симметрии преобразования $\boldsymbol{x}'=A\boldsymbol{x}$ матрица A также будет симметричной ($a_{12}=a_{21}$). В равенствах (8) $\tilde{x}_1, \tilde{x}_2$ являются новыми координатами вектора $\boldsymbol{x}$, а $\tilde{x}_1', \tilde{x}_2'$ — новыми координатами вектора $\boldsymbol{x}'=A\boldsymbol{x}$. Вследствие ортогональности нового базиса можем написать

$$\Phi=\boldsymbol{x}\boldsymbol{x}'=\tilde{x}_1\tilde{x}_1'+\tilde{x}_2\tilde{x}_2'.$$

Заменяя здесь $\tilde{x}_1', \tilde{x}_2'$ по формулам (8), получим

$$\Phi=\tilde{x}_1(\tilde{a}_{11}\tilde{x}_1+\tilde{a}_{12}\tilde{x}_2)+\tilde{x}_2(\tilde{a}_{21}\tilde{x}_1+\tilde{a}_{22}\tilde{x}_2)= \\ =\tilde{a}_{11}\tilde{x}_1^2+2\tilde{a}_{12}\tilde{x}_1\tilde{x}_2+\tilde{a}_{22}\tilde{x}_2^2.$$

Мы видим, что элементы матрицы $\tilde{A}$ являются в то же время коэффициентами квадратичной формы Ф в новых координатах. Следовательно, определение этих коэффициентов равносильно определению матрицы $\tilde{A}$. Тем самым дело сводится к прямому применению результатов п.6.6; имеем

$$\tilde{A}=L^{-1}AL.$$

Так как $L^{-1} = L^*$, то

$$\tilde{A}=L^*AL. \tag{10}$$

В подробной записи:

$$\begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{pmatrix}=\begin{pmatrix} l_1 & m_1 \\ l_2 & m_2 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix}.$$

Равенство (10) *выражает матрицу А данной квадратичной формы в новом базисе* $\tilde{e}_1, \tilde{e}_2$ *через матрицу А этой формы в старом базисе* e_1, e_2 *и через матрицу L преобразования координат.*

79. Важной задачей математики и ее приложений является задача приведения данной квадратичной формы к каноническому виду.

Говорят, что квадратичная форма имеет *канонический вид,* если она содержит только члены с квадратами текущих координат, т.е. если $a_{12} = a_{21} = 0$. В эгом случае матрица формы является диагональной:

$$A = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}.$$

Чтобы произвольную форму привести к каноническому виду, нужно найти такой базис, в котором матрица формы получает диагональный вид. Выше, в п° 78, мы видели, что при переходе от ортогонального базиса $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ к новому ортогональному базису $\tilde{e}_1, \tilde{e}_2$ матрица квадратичной формы изменяется точно так же, как матрица симметрического линейного преобразования. Поэтому задача о приведении к каноническому виду формы (1) сводится к задаче о приведении к диагональному виду матрицы линейного преобразования (4). Решение этой последней задачи изложено в п.6.11.

Согласно сказанному, приведение квадратичной формы от двух переменных x_1, x_2 к каноническому виду можно осуществить по следующему плану.

Пусть в базисе $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ дана форма

$$\Phi = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_{22}^2$$

с матрицей

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

1) Составим характеристическое уравнение матрицы *А:*

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0.$$

Найдем корни этого уравнения λ_1, λ_2.

2) Следуя указаниям п.6.11, найдем для симметрического преобразования (4) два единичных собственных вектора с собственными числами λ_1, λ_2 так, чтобы эти векторы были пер-

пендикулярны друг к другу; обозначим их соответственно через $\tilde{e}_1$, $\tilde{e}_2$:

$$\tilde{e}_1 = \{l_1;\ m_1\},\ \tilde{e}_2 = \{l_2;\ m_2\}.$$

3) Перейдем к базису $\tilde{e}_1$, $\tilde{e}_2$. В этом базисе

$$\tilde{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

и данная форма получает канонический вид:

$$\Phi = \lambda_1 \tilde{x}_1^2 + \lambda_2 \tilde{x}_2^2. \qquad (11)$$

При переходе к базису e_1, e_2 координаты всех векторов преобразуются согласно матричному равенству

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} l_1 & l_2 \\ m_1 & m_2 \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix}.$$

Пример. Привести к каноническому виду квадратичную форму

$$\Phi = 5x_1^2 + 8x_1x_2 + 5x_2^2.$$

Решение. Составим характеристическое уравнение

$$\begin{vmatrix} 5-\lambda & 4 \\ 4 & 5-\lambda \end{vmatrix} = 0,$$

или $(\lambda - 5)^2 - 16 = 0$. Отсюда $\lambda_1 = 1$, $\lambda_2 = 9$. Канонический вид данной квадратичной формы можем написать сразу

$$\Phi = \tilde{x}_1^2 + 9\tilde{x}_2^2.$$

Чтобы найти базис, в котором форма имеет такой вид, придется искать собственные векторы преобразования (4). Напишем систему уравнений, определяющую искомые собственные векторы.

$$\begin{aligned} (5-\lambda)\, l + 4m &= 0, \\ 4l + (5-\lambda)\, m &= 0 \end{aligned} \qquad (*)$$

Подставляя сюда $\lambda = \lambda_1 = 1$, $\lambda = \lambda_2 = 9$ и беря каждый раз нормированное решение системы (*), найдем векторы

$$\tilde{e}_{-1} = \left\{\frac{1}{\sqrt{2}};\ -\frac{1}{\sqrt{2}}\right\},\quad \tilde{e}_2 = \left\{\frac{1}{\sqrt{2}};\ \frac{1}{\sqrt{2}}\right\}.$$

Они составляют нужный базис.

При переходе к базису $\tilde{e}_1$, $\tilde{e}_2$ координаты всех векторов преобразуются по формулам

$$x_1 = \frac{1}{\sqrt{2}}\tilde{x}_1 + \frac{1}{\sqrt{2}}\tilde{x}_2,$$
$$x_2 = -\frac{1}{\sqrt{2}}\tilde{x}_1 + \frac{1}{\sqrt{2}}\tilde{x}_2.$$

Непосредственной проверкой легко убедиться, что эти формулы действительно дают преобразование нашей квадратичной формы к каноническому виду.

80. Вернемся к п°78. Вместе с данной квадратичной формой мы ввели там некоторое симметрическое линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$; оно определяется формулами (4) и имеет ту же матрицу, что и данная квадратичная форма.

Направления собственных векторов преобразования $\boldsymbol{x}'=A\boldsymbol{x}$ называются *главными направлениями* данной формы.

Когда мы приводим квадратичную форму к каноническому виду, то при этом мы переходим к ортогональному базису, векторы которого имеют главные направления данной формы.

81. Коэффициенты λ_1, λ_2 в каноническом виде (11) квадратичной формы называются *характеристическими числами* квадратичной формы. Они же — корни характеристического уравнения; они же и собственные числа собственных векторов линейного преобразования $\boldsymbol{x}'=A\boldsymbol{x}$, о котором шла речь в предыдущем пункте.

82. Если $\lambda_1 \neq \lambda_2$, то форма имеет единственную пару главных направлений, причем они перпендикулярны друг к другу. Если $\lambda_1=\lambda_2$, то любое направление является главным (см. п.6.11). В последнем случае форма в любом ортогональном базисе имеет канонический вид $\lambda x_1^2 + \lambda x_2^2$, где

$$\lambda = \lambda_1 = \lambda_2.$$

83. Зная характеристические числа формы, можно установить оценки величины этой формы.

В самом деле, пусть в ортогональном базисе $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ дана форма

$$\Phi = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

Переходя к новому ортогональному базису $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$, приведем ее к каноническому виду

$$\Phi = \lambda_1\tilde{x}_1^2 + \lambda_2\tilde{x}_2^2.$$

Допустим, что $\lambda_1 \leqslant \lambda_2$. Тогда

$$\lambda_1(\tilde{x}_1^2+\tilde{x}_2^2)\leqslant\lambda_1\tilde{x}_1^2+\lambda_2\tilde{x}_2^2\leqslant\lambda_2(\tilde{x}_1^2+\tilde{x}_2^2).$$

Так как базисы $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ и $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$ оба ортогональны, то

$$\tilde{x}_1^2+\tilde{x}_2^2=x_1^2+x_2^2.$$

Следовательно,

$$\lambda_1(x_1^2+x_2^2)\leqslant\Phi\leqslant\lambda_2(x_1^2+x_2^2). \qquad (12)$$

Это и есть оценки, которые мы имели в виду получить. Соотношения (12) являются точными оценками. Именно, в точке, где $\tilde{x}_2=0$, Φ совпадает с левой частью этих соотношений; в точке, где $\tilde{x}_1=0$, Φ равно правой части.

В точках единичной окружности $x^2_1+x^2_2=1$ оценки (12) принимают особенно простой вид:

$$\lambda_1\leqslant\Phi\leqslant\lambda_2.$$

Таким образом, характеристические числа являются границами изменения численных значений квадратичной формы на единичной окружности.

84. Если $\lambda_1>0$, $\lambda_2>0$, то согласно (12) форма Ф положительна во всех точках (кроме начала координат, где она равна нулю). Такая форма называется *положительно определенной.* Если $\lambda_1<0$, $\lambda_2<0$, форма называется *отрицательно определенной;* в этом случае форма имеет во всех точках отрицательные значения (не считая начала координат). Если λ_1, λ_2 — числа разных знаков, то форма в некоторых точках положительна, в некоторых — отрицательна (в некоторых обращается в нуль). Такая форма называется *знакопеременной.*

Пример. Формая

$$x_1^2-x_1x_2+x_2^2$$

является положительно определенной, так как $\lambda_1=\dfrac{1}{2},\lambda_2=\dfrac{3}{2}$; имеем $\lambda_1>0$, $\lambda_2>0$.

На единичной окружности $x^2_1+x^2_2=1$ эта форма изменяется в границах

$$\frac{1}{2}\leqslant x_1^2-x_1x_2+x_2^2\leqslant\frac{3}{2}.$$

85 Теперь мы перейдем к рассмотрению квадратичной формы от трех переменных:

$$\Phi = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 +$$
$$+ 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3. \quad (13)$$

Ввиду полной аналогии с предыдущим, мы лишь кратко отметим основные положения, касающиеся таких форм.

Мы полагаем, что в пространстве введена система координат с началом *O* и ортогональным базисом ***e₁, e₂, e₃*** (из единичных векторов). Тогда x_1, x_2, x_3 можно рассматривать как координаты некоторой точки $M(x_1;\ x_2;\ x_3)$ или вектора $x = \overline{OM}$. Величина Ф есть значение формы в точке $M(x_1;\ x_2;\ x_3)$. В дальнейшем будем употреблять только ортогональные базисы (из единичных векторов).

Данную форму напишем в следующем виде:

$$\Phi = x_1(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) +$$
$$+ x_2(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) +$$
$$+ x_3(a_{31}x_1 + a_{32}x_2 + a_{33}x_3),$$

считая $a_{21} = a_{12}$, $a_{31} = a_{13}$, $a_{32} = a_{23}$.

Вместе с данной формой рассмотрим линейное преобразование $\boldsymbol{x}'=A\boldsymbol{x}$, которое в базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ имеет координатное представление

$$x_1' = a_{11}x_1 + a_{12}x_2 + a_{13}x_3,$$
$$x_2' = a_{21}x_1 + a_{22}x_2 + a_{23}x_3,$$
$$x_3' = a_{31}x_1 + a_{32}x_2 + a_{33}x_3.$$

Матрица этого преобразования

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

называется матрицей данной квадратичной формы в данном базисе $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. Форма Ф теперь может быть записана так:

$$\Phi = x_1x_1' + x_2x_2' + x_3x_3' = \boldsymbol{x}\boldsymbol{x}' = \boldsymbol{x}A\boldsymbol{x};$$

здесь $\boldsymbol{x}\boldsymbol{x}'$ — скалярное произведение вектора $\boldsymbol{x}$ на вектор $\boldsymbol{x}' = A\boldsymbol{x}$.

Пусть $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$, $\tilde{\boldsymbol{e}}_3$ — единичные попарно перпендикулярные собственные векторы преобразования $\boldsymbol{x}' = A\boldsymbol{x}$, имеющие собственные числа λ_1, λ_2, λ_3. Если перейти к базису $\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3$, то матрица квадратичной формы (как матрица преобразования $\boldsymbol{x}' = A\boldsymbol{x}$) станет диагональной:

$$\tilde{A}=\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}.$$

Одновременно сама форма получит канонический вид:

$$\Phi=\lambda_1\tilde{x}_1^2+\lambda_2\tilde{x}_2^2+\lambda_3\tilde{x}_3^2.$$

Числа λ_1, λ_2, λ_3 суть корни характеристического уравнения матрицы *А;* их называют *характеристическими числами* данной квадратичной формы.

Направления собственных векторов преобразования ***x***$'$=*А****x*** называются *главными направлениями* квадратичной формы. Если вектор имеет собственное число

$$\lambda\ (=\lambda_1,\ \lambda_2,\ \lambda_3),$$

то определяемое им главное направлениеназывается соответствующим этому числу λ.

Когда мы приводим квадратичную форму к каноническому виду, то при этом мы переходим к ортогональному базису, векторы которого имеют главные направления данной формы. Способ разыскания такого базиса изложен в п.6.12.

Если числа λ_1, λ_2, λ_3 — все разные, то квадратичная форма имеет точно три главных направления; они перпендикулярны друг к другу. Если $\lambda_1\neq\lambda_2=\lambda_3$, то числу λ_1 соответствует одно определенное главное направление; условимся называть его первым. Числу $\lambda=\lambda_2=\lambda_3$ в этом случае соответствует бесконечное множество главных направлений; именно, любое направление, перпендикулярное к первому главному направлению, будет главным направлением и будет соответствовать числу $\lambda=\lambda_2=\lambda_3$. Если $\lambda_1=\lambda_2=\lambda_3$, то любое направление в пространстве является главным направлением формы; такая квадратичная форма в любом ортогональном базисе имеет канонический вид $\lambda x_1^2+\lambda x_2^2+\lambda x_3^2$, где $\lambda=\lambda_1=\lambda_2=\lambda_3$. Все утверждения, которые мы сейчас высказали, непосредственно вытекают из результатов п.6.12.

Пример. Дана квадратичная форма

$$\Phi=7x_1^2+6x_2^2+5x_3^2-4x_1x_2-4x_2x_3;$$

привести ее к каноническому виду.

Решение. Матрица данной формы

$$A=\begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}$$

совпадает с матрицей линейного преобразования, которое рассмотрено в первом примере п° 75. Имеем $\lambda_1=3,\ \lambda_2=6,\ \lambda_3=9$. В качестве нового базиса следует взять векторы e_1, e_2, e_3, которые найдены в указанном примере. В базисе $\tilde{e}_1,\ \tilde{e}_2,\ \tilde{e}_3$ квадратичная форма имеет канонический вид:

$$\Phi=3\tilde{x}_1^2+6\tilde{x}_2^2+9\tilde{x}_3^2.$$

86. Методы приведения квадратичной формы к каноническому виду прежде всего находят приложение к аналитической геометрии.

Займемся задачей упрощения уравнения кривой второго порядка. Будем, как принято в аналитической геометрии, векторы ортогонального базиса на плоскости обозначать через $\boldsymbol{i}$, $\boldsymbol{j}$, а координаты текущей точки — буквами x, y. Чтобы не усложнять дело лишними деталями, рассмотрим уравнение второй степени без линейных членов

$$a_{11}x^2+2a_{12}xy+a_{22}y^2=H. \tag{14}$$

Такое уравнение определяет кривую второго порядка, центр которой находится в начале координат. Задача о приведении уравнения (14) к каноническому виду полностью сводится к задаче о приведении к каноническому виду квадратичной формы старших членов, которая стоит в левой части этого уравнения. Главные направления этой формы называются главными направлениями данной кривой.

Пусть A — матрица квадратичной формы старших членов, λ_1, λ_2 — корни ее характеристического уравнения, $\tilde{\boldsymbol{i}},\ \tilde{\boldsymbol{j}}$ — перпендикулярные друг к другу единичные векторы главных направлений, отвечающие числам λ_1, λ_2. Тогда, если мы, не меняя начала координат, перейдем к базису $\tilde{\boldsymbol{i}},\ \tilde{\boldsymbol{j}}$, то в новых координатах уравнение кривой примет канонический вид:

$$\lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2=H.$$

Пример. Привести к каноническому виду уравнение кривой второго порядка

$$5x^2+8xy+5y^2=9;$$

определить вид кривой.

Решение. Квадратичная форма старших членов данного уравнения совпадает с квадратичной формой, которая была рассмотрена в примере п° 79. Имеем $\lambda_1=1$, $\lambda_2=9$. Искомый базис

$$\tilde{i}=\left\{\frac{1}{\sqrt{2}};\ \frac{-1}{\sqrt{2}}\right\},\quad \tilde{j}=\left\{\frac{1}{\sqrt{2}};\ \frac{1}{\sqrt{2}}\right\}.$$

Переход к новой системе координат достигается поворотом осей на угол α=—45°; соответствующие формулы преобразования координат имеют вид

$$x=\frac{\tilde{x}+\tilde{y}}{\sqrt{2}},\quad y=\frac{-\tilde{x}+\tilde{y}}{\sqrt{2}}.$$

Каноническое уравнение кривой

$$\tilde{x}^2+9\tilde{y}^2=9,$$

или

$$\frac{\tilde{x}^2}{9}+\frac{\tilde{y}^2}{1}=1.$$

Кривая является эллипсом с полуосями 3 и 1.

87. Рассмотрим в пространственных координатах с ортогональным базисом ***i, j, k*** уравнение

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz=H. \quad (15)$$

Уравнение (15) определяет поверхность второго порядка с центром в начале координат.

Данное уравнение приводится к каноническому виду одновременно с квадратичной формой старших членов, которая стоит в его левой части. Чтобы привести уравнение (15) к каноническому виду, нужно (не меняя начала координат) перейти к ортогональному базису $\tilde{i}$, $\tilde{j}$, $\tilde{k}$, векторы которого имеют главные направления поверхности. При этом главными направлениями поверхности называются главные направления квадратичной формы старших членов ее уравнения.

Пример. Привести к каноническому виду уравнение поверхности второго порядка

$$7x^2+6y^2+5z^2-4xy-4yz=18;$$

определить вид поверхности.

Решение. Составим матрицу квадратичной формы старших членов:

$$A=\begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}.$$

Эта матрица совпадает с матрицей линейного преобразования, которое рассматривалось в первом примере п. 75. Имеем $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$. Искомый базис

$$\tilde{i} = \left\{\frac{1}{3};\ \frac{2}{3};\ \frac{2}{3}\right\},$$

$$\tilde{j} = \left\{\frac{2}{3};\ \frac{1}{3};\ -\frac{2}{3}\right\},$$

$$\tilde{k} = \left\{-\frac{2}{3};\ \frac{2}{3};\ -\frac{1}{3}\right\}.$$

Формулы преобразования координат при переходе к этому базису

$$x = \frac{1}{3}\tilde{x} + \frac{2}{3}\tilde{y} - \frac{2}{3}\tilde{z},$$

$$y = \frac{2}{3}\tilde{x} + \frac{1}{3}\tilde{y} + \frac{2}{3}\tilde{z},$$

$$z = \frac{2}{3}\tilde{x} - \frac{2}{3}\tilde{y} - \frac{1}{3}\tilde{z}.$$

Каноническое уравнение данной поверхности

$$3\tilde{x}^2 + 6\tilde{y}^2 + 9\tilde{z}^2 = 18,$$

или

$$\frac{\tilde{x}^2}{6} + \frac{\tilde{y}^2}{3} + \frac{\tilde{z}^2}{2} = 1.$$

Данная поверхность является эллипсоидом с полуосями $\sqrt{6}$, $\sqrt{3}$, $\sqrt{2}$.

88. Если расположение базиса $\tilde{i}$, $\tilde{j}$, $\tilde{k}$ нас не интересует, то каноническое уравнение поверхности можно написать сразу, зная λ_1, λ_2, λ_3. Числа λ_1, λ_2, λ_3. суть корни характеристического уравнения, которое является уравнением третьей степени. Нахождение корней такого уравнения может оказаться довольно хлопотливым. Между тем очень часто требуется определить только вид поверхности, а для этого достаточно знать лишь знаки чисел λ_1, λ_2, λ_3. Знаки корней характеристического уравнения можно установить весьма просто при помощи известного правила Декарта.

Литература

Основная

1. Горбатов В.А. Основы дискретной математики. - М.: Высшая школа, 1986.

2. Коршунов Ю.М. Математические основы кибернетики. - М.:Энергия, 1980.

3. Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженера. - М.:Энергия, 1980.

4. Кук Д., Бейз Г. Компьютерная математика.- М.: Наука, 1990.

5. Сигорский В.П. Математический аппарат инженера.-К.: Техника, 1977.

6. Кузичев А.С. Диаграммы Венна.- М.: Наука, 1968.

7. Кононюк А.Ю. Высшая математика. В 2 ч. Ч.1, 2 - К:Кольори, 2007.

8. Аверкин А.Н., Батыршин И.З. и др. Нечеткие множества в моделях управления и икусственного интелекта.- М.: Наука, 1986.

9. Кофман А. Введение в теорию нечетких множеств.- М.: Радио и связь, 1982.

10. Кофман А. Введение в прикладную комбинаторику.- М.: Наука, 1975.

11. Згуровский М.З. Интегрированные системы оптимального управления и проектирования.- К.: Высшая школа, 1990.

13. Минский М. Фреймы для представления знаний. - М.:Энергия, 1979.

14. Вильсон А. Дж. Энтропийные методы моделирования сложных систем.- М.: Наука, 1978.

15. В.Ю. Юрков, О.В. Лукина /Прикладная геометрия, вып. 8, N 18 (2006), стр. 9-36.

16. И.И. Ежев, А.В. Скороход, М.И. Ядренко. Элементы комбинаторики.-М.: Наука, 1977.

17. Алгебраическая теория автоматов, языков и полугрупп. Под ред. М.А. Арбиба. - М.: Статистика, 1975.

18. Минк Х. Перманенты. М.: Мир, 1982.

19. Холл. М. М.: Мир, 1970.

20. Ф.Р. Гантмахер. Теория матриц. - М.: Наука, 1967.

21. М.А. Акивис, В.В. Гольдберг. Тензорное исчисление. - М.: Наука, 1972.

Дополнительная

1. Айне Э. Л., Обыкновенные дифференциальные уравнения, ДНТВУ, 1939, гл. XIX.
2. Ахиезер Н. И., Лекции по теории аппроксимации, ГостехизДат, 1948, гл. I.
3. Ахиезер Н. И. и Крейн М. Г., О некоторых вопросах теории моментов, ДНТВУ, 1938.
4. Беркштейн С. Н., Теория вероятностей, 4-е изд., ГостехизДат, 1946.
5. Б о хер М., Введение в высшую алгебру, ОНТИ, 1933.
6. Булгаков Б. В., Колебания, т. I, ГостехизДат, 1949, гл. I.
7. Г а н т м а х е р Ф. Р. и Крейн М. Г., Осцилляционные матрицы и ядра и малые колебания механических систем, 2-е изд., Гостехиздат, 1950.
8. Гельфанд И. М., Лекции по линейной алгебре, 3-е изд., Изд-во «Наука», 1966.
9. Геронимус Я. Л., Теория ортогональных многочленов, Гостехиздат, 1950, гл. II.
10. Граве Д. А., Элементы высшей алгебры, Киев, 1914.
11. Еругин Н. П., Приводимые системы, Труды Матем. ин-та им. В. А. Стеклова, т. XIII (1946).
12. Еругин Н. П., Метод Лаппо-Данилевского в теории линейных дифференциальных уравнений, Изд-во Лешшгр. ун-та, 1956.
13. Каган В. Ф., Основания теории определителей, Гос. изд-во Украины, 1922.
14. Клейн Ф., Высшая геометрия, ГОНТИ, 1939, §§ 96—99.
15. Крейн М. Г., Основные положения теории Х-зон устойчивости канонической системы линейных дифференциальных уравнений с периодическими коэффициентами, Сб. «Памяти А. А. Андронова», Изд-во Академии наук СССР, 1955.
16. Крейн М. Г. и Наймарк М. А., Метод симметрических и эрмитовых форм в теории отделения корней алгебраических уравнений, Харьков, 1936.
17. Крейн М. Г. и Рутман М. А., Линейные операторы, оставляющие инвариантным конус в пространстве Банаха, УМН 3, № 1 (23) (1948).
18. Кудрявцев Л. Д., О некоторых математических вопросах теории электрических цепей, УМН 3, № 4 (26) (1948).
19. Курант Р. и Гильберт Д., Методы математической физики, т. I, 3-е изд., Гостехиздат, 1951, гл. I, П.
20. К у р о ш А. Г., Курс высшей алгебры, 3-е изд., Гостехиздат, 1952.

21. Л а п п о-Данилевский И. А., а) Теория функций от матриц и систем линейных дифференциальных уравнений, Гостехиздат, 1934. б) Mernoires sur la theorie des systemes des equations differentielles lineaires, т. I, 11, III, Труды Физико-матем. ин-та им. В. А. Стеклова, VI—VIII (1934—1936).
22. Ляпунов А. М., Общая задача об устойчивости движения, Гостехиздат, 1950.
23. Малкин И. Г., Теория устойчивости движения, Гостехиздат, 1952.
24. Мальцев А. И., Основы линейной алгебры, Гостехиздат, 1948.
25. Марков А. А., Избранные труды, Гостехиздат, 1948.
26. М е й м а н Н. Н., Некоторые вопросы расположения корней полиномов, УМН 4, № 6 (34) (1949).
27. Наймарк Ю. И., Устойчивость линеаризованных систем, Изд-во Ленингр. Военно-воздушной инж. академии, 1949.
28. П о т а п о в В. П., Мультипликативная структура J-нерастягивающих матриц функций/ Труды Моск. матем. о-ва 4 (1955).
29. Романовский В. И., Дискретные цепи Маркова, Гостехиздат, 1948.
30. Смирнов В. И., Курс высшей математики, т. III, Физматгиз, 1958.
31. Стильтьес Т. И., Исследование о непрерывных дробях, ОНТИ, Харьков, 1936.

Научно-практическое издание

Кононюк Анатолий Ефимович

Дискретно непрерывная математика

Книга 5

Матрицы

Часть 1

Авторская редакция

Подписано в печать 30.09.2011 г.
Формат 60x84/16.
Усл. печ. л. 16,5. Тираж 300 экз.

Издатель и изготовитель:
Издательство «Освита Украины»
04214, г. Киев, ул. Героев Днепра, 63, к. 40
Свидетельство о внесении в Государственный реестр
издателей ДК №1957 от 23.04.2009 г.
Тел./факс (044) 411-4397; 237-5992
E-mail: osvita2005@ukr.net, www.rambook.ru