СПРАВОЧНАЯ
МАТЕМАТИЧЕСКАЯ
БИБЛИОТЕКА

ВЫСШАЯ АЛГЕБРА

ФИЗМАТГИЗ · 1962

СПРАВОЧНАЯ МАТЕМАТИЧЕСКАЯ БИБЛИОТЕКА

ПОД ОБЩЕЙ РЕДАКЦИЕЙ
Л. А. ЛЮСТЕРНИКА
И
А. Р. ЯНПОЛЬСКОГО

А. П. МИШИНА и И. В. ПРОСКУРЯКОВ

ВЫСШАЯ АЛГЕБРА

ЛИНЕЙНАЯ АЛГЕБРА, МНОГОЧЛЕНЫ, ОБЩАЯ АЛГЕБРА

Под редакцией
П. К. РАШЕВСКОГО

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1962

АННОТАЦИЯ

Настоящий выпуск посвящен основным понятиям алгебры. В нем излагаются теория определителей, решение систем линейных уравнений, теория матриц, свойства квадратичных форм, алгебра многочленов, отыскание корней многочленов с числовыми коэффициентами, а также элементы теории групп, колец и полей.

Книга дает возможность читателю навести нужную справку и ознакомиться с основными фактами высшей алгебры, изложенными конспективно, без доказательств, но в систематической форме и с разбором ряда примеров.

Книга предназначена для математиков, физиков и инженеров, применяющих в своей работе математические методы, а также для студентов и аспирантов.

Справочная математическая библиотека
под общей редакцией
Л. А. Люстерника и А. Р. Янпольского.

Высшая алгебра
(линейная алгебра, многочлены, общая алгебра)
М., Физматгиз, 1962 г., 300 стр. с илл.

Редактор *В. Н. Латышев.*

Техн. редактор *Л. Ю. Плакше.* Корректор *Т. С. Плетнева.*

Сдано в набор 4/I 1962 г. Подписано к печати 20/IV 1962 г. Бумага $84\times108/_{32}$. Физ. печ. л. 9,375. Условн. печ. л. 15,38. Уч.-изд. л. 16,01. Тираж 25 000 экз. Т-04732. Цена книги 90 коп. Заказ № 1096.

Государственное издательство физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Ленинградский Совет народного хозяйства. Управление полиграфической промышленности. Типография № 1 «Печатный Двор» имени А. М. Горького. Ленинград, Гатчинская, 26.

СОДЕРЖАНИЕ

Глава II

Матрицы и квадратичные формы

Глава III

Алгебра многочленов

Глава IV

Общая алгебра

ПРЕДИСЛОВИЕ

В деятельности научных работников технических специальностей, а также математиков, физиков, химиков и др. нередки случаи, когда необходимо навести справку по тому или иному математическому вопросу без детального его изучения. Не всегда бывает легко отыскать соответствующую литературу, в особенности если вопрос не элементарный, а иногда непросто найти нужный результат в тексте какой-либо монографии и разобраться в нем в отрыве от предшествующего изложения.

Издание серии справочников по высшей математике, как нам кажется, будет весьма полезным в этом отношении. Предлагаемая книга посвящена высшей алгебре. При этом в главах I и II собраны результаты, относящиеся к линейной алгебре (матрицы, определители, системы линейных уравнений, квадратичные формы); в главе III изложена алгебра многочленов (действия над многочленами, алгебраические уравнения с одним неизвестным и свойства их корней, дробно-рациональные функции, многочлены от нескольких переменных); глава IV посвящена важнейшим понятиям алгебры как современной математической науки (группы, кольца, поля и тела, алгебры, структуры).

В кратком введении даны некоторые основные определения, используемые в дальнейшем, и его следует прочесть до пользования справочником.

Доказательства отсутствуют, но формулировки результатов сопровождаются рядом примеров, разъясняющих их смысл. Интересующий читателя вопрос следует искать по подробному оглавлению или по алфавитному указателю в конце книги.

Добавим еще, что материал, наиболее часто встречающийся в приложениях, содержится в первых трех главах; четвертая глава имеет целью помочь читателю разобраться в незнакомых ему понятиях современной алгебры.

При ссылке на параграф той же главы номер главы опускается; аналогично при ссылке на пункт того же параграфа опускается номер этого параграфа.

Главы I и II написаны И. В. Проскуряковым, главы III и IV — А. П. Мишиной.

П. К. Рашевский

ВВЕДЕНИЕ

На протяжении всей книги встречаются следующие основные понятия: множество, числовое поле, матрица. Основные понятия и обозначения теории множеств даны в [34], стр. 22—24.

Числовым полем K называется любое множество чисел, содержащее хотя бы одно число, отличное от нуля, и вместе с любыми двумя числами a и b содержащее их сумму $a+b$, разность $a-b$, произведение ab и частное $\frac{a}{b}$ (если $b \neq 0$). Можно проверить, что все рациональные числа образуют числовое поле. Точно так же все действительные числа или все комплексные числа образуют числовое поле. Эти три числовых поля играют важнейшую роль во всей математике. Поле всех комплексных чисел обладает следующим важным свойством: любой многочлен степени $n>0$ от одного переменного x с комплексными (в частности, действительными) коэффициентами имеет в том же поле комплексных чисел все n своих корней (см. гл. III). Это свойство поля комплексных чисел называется его алгебраической замкнутостью. Кроме числовых полей, рассматривается также общее понятие поля, элементами которого не обязательно являются числа (см. гл. IV). В первых трех главах (за исключением специально оговоренных мест) речь будет идти только о числовых полях.

Матрицей называется прямоугольная таблица чисел. Для обозначения матрицы применяются круглые скобки или двойные вертикальные линейки, например:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \quad \text{или} \quad \begin{Vmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{Vmatrix}.$$

В данном выпуске применяется первое обозначение. Числа, составляющие матрицу, называются ее элементами. Горизонтальные ряды элементов матрицы называются строками, а вертикальные — столбцами. Если число строк матрицы равно числу столбцов, то матрица называется *квадратной*, а число строк — *порядком* этой матрицы. В дальнейшем для краткости под матрицами данного порядка подразумеваются квадратные матрицы этого порядка. При записи в общем виде элементы матрицы обозначаются обычно одной буквой с двумя индексами, из которых первый указывает номер строки, а второй — номер столбца, например, матрица порядка n в общем виде запишется так:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

(читается: «a один, один, a один, два» и т. д., а не «a одиннадцать, a двенадцать»). Ряд элементов квадратной матрицы, лежащих на отрезке, соединяющем левый верхний угол с правым нижним, называется *главной* (или первой) диагональю, а на отрезке, соединяющем правый верхний угол с левым нижним, — *побочной* (или второй) диагональю матрицы.

ГЛАВА I
ОПРЕДЕЛИТЕЛИ И СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 1. Определители

1. Перестановки и подстановки. *Перестановкой* n чисел $1, 2, \ldots, n$ (или n любых различных между собой символов $a_1, a_2, \ldots, a_n$) называется любое расположение этих чисел (или символов) в определенном порядке. Так как данные n символов можно занумеровать числами $1, 2, \ldots, n$, то изучение перестановок любых n символов сводится к изучению перестановок этих чисел. Число всех перестановок из n чисел равно $1 \cdot 2 \cdot 3 \ldots n = n!$ (читается: «n-факториал»).

Пример 1. Все перестановки чисел 1, 2, 3 имеют вид: 123, 132, 213, 231, 312, 321. Число их $3! = 6$.

Говорят, что два числа в перестановке образуют *инверсию*, если большее число стоит впереди меньшего, если же меньшее число стоит впереди большего, то — *порядок*. Способ подсчета числа инверсий: читаем числа перестановки в порядке их записи (слева направо), для каждого из чисел считаем, сколько чисел меньших данного стоит правее него, и все полученные числа складываем.

Пример 2. В перестановке 528371964 число инверсий равно $4+1+5+1+3+2+1=17$.

Перестановка называется *четной* или *нечетной*, смотря по тому, будет ли число инверсий в ней четно или нечетно.

Транспозицией называется перемена местами двух чисел перестановки. Транспозиция чисел i и j обозначается через (i, j). От любой перестановки n чисел к любой другой перестановке тех же чисел можно перейти путем ряда транспозиций, причем можно обойтись не более чем $n-1$ транспозициями.

Пример 3. От перестановки 25134 к перестановке 42513 можно перейти путем четырех транспозиций: (2, 4), (2, 5), (1, 5), (1, 3).

Все $n!$ перестановок чисел $1, 2, \ldots, n$ можно так расположить одну вслед за другой (без пропусков и без повторений), что каждая следующая перестановка получается из предыдущей путем одной транспозиции. Каждая транспозиция меняет четность перестановки. При любом $n \geqslant 2$ число четных перестановок из n чисел равно числу нечетных, т. е. $\frac{1}{2}n!$.

Подстановкой n чисел $1, 2, \ldots, n$ или *подстановкой n-й степени* называется взаимно однозначное отображение совокупности этих чисел на себя, т. е. такое отображение, при котором каждому числу от 1 до n соответствует какое-то из этих чисел и двум различным числам всегда соответствуют два различных числа. Подстановка записывается двумя строками в общих скобках, причем каждому числу верхней строки соответствует стоящее под ним число нижней. Например, $\begin{pmatrix} 2 & 1 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ обозначает подстановку, в которой: $1 \to 1$, $2 \to 3$, $3 \to 4$, $4 \to 2$.

В зависимости от расположения чисел в верхней строке одну и ту же подстановку можно записывать многими способами. Например, записи

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}$$

обозначают одну и ту же подстановку, при которой 1 переходит в 2, 2 в 3, 3 в 1. Каждая подстановка n чисел допускает $n!$ различных записей. Число различных подстановок n элементов также равно $n!$.

Подстановка называется *четной*, если общее число инверсий в обеих ее строках четно, и *нечетной*, если нечетно. Иначе говоря, подстановка четна, если ее строки имеют одинаковую четность, и нечетна, если — противоположную четность.

Четность подстановки не зависит от ее записи в виде двух строк, т. е. все записи одной и той же подстановки имеют одинаковую четность.

Пример 4. Подстановка

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

содержит в первой записи шесть, а во второй — две инверсии и, значит, четна.

Число четных подстановок n элементов равно числу нечетных, т. е. $\frac{1}{2}n!$ (если $n\geqslant 2$).

Существует другой метод определения четности подстановки. *Циклом* называется последовательность нескольких чисел, в которой первое число при данной подстановке переходит во второе, второе в третье и т. д., а последнее — в первое. Цикл обозначается заключением его чисел в общие скобки. Если число переходит само в себя, то оно одно образует цикл. Циклы, не имеющие общих чисел, называются *независимыми.* Любую подстановку можно разложить на независимые циклы.

Пример 5. $\begin{pmatrix}1&2&3&4&5&6\\6&5&1&4&2&3\end{pmatrix}=(163)\ (25)\ (4).$

Декрементом d подстановки называется разность между числом n всех ее элементов и числом k циклов в ее разложении: $d=n-k$. Четность подстановки совпадает с четностью ее декремента. Так, в примере 5: $n=6$, $k=3$, $d=3$, т. е. подстановка нечетна.

Наименьшее число транспозиций, необходимое для перехода от одной перестановки n чисел $a_1, a_2, \ldots, a_n$ к другой перестановке тех же чисел $b_1, b_2, \ldots, b_n$, равно декременту подстановки

$$\begin{pmatrix}a_1&a_2&\ldots&a_n\\b_1&b_2&\ldots&b_n\end{pmatrix}$$

(см. [20], задача 183). Следовательно, для любой перестановки n чисел существуют перестановки тех же чисел, которые нельзя перевести в данную перестановку менее чем $n-1$ транспозициями.

Можно ввести операцию умножения подстановок, дать другое определение цикла, как некоторой подстановки, и истолковать разложение подстановки на независимые циклы как представление ее в виде произведения этих циклов. Об этом см. гл. IV, § 1, пп. 1 и 9, а также [11], стр. 33—36.

2. Понятие определителя любого порядка. Пусть дана квадратная матрица порядка n:

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}=(a_{ij})_1^n.$$

Определителем n-го порядка, или определителем матрицы A, при $n>1$ называется число, полученное из элементов этой матрицы по формулам

$$|A|=|a_{ij}|_1^n=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}=\sum(-1)^{s+t}a_{i_1j_1}a_{i_2j_2}\dots a_{i_nj_n}=$$

$$=\sum(-1)^k a_{1k_1}a_{2k_2}\dots a_{nk_n}=\sum(-1)^l a_{l_11}a_{l_22}\dots a_{l_nn}.$$

Здесь первые три выражения — обозначения определителя; первая сумма берется по всем различным между собой подстановкам

$$\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}, \qquad (*)$$

причем s — число инверсий в верхней, а t — в нижней строке этой подстановки, вторая сумма берется по всем перестановкам $k_1, k_2, \dots, k_n$ и k — число инверсий в этой перестановке. Аналогичный смысл имеет третья сумма. Все три суммы тождественно равны. Их слагаемые называются *членами определителя*; каждый член определителя равен произведению n элементов матрицы, взятых по одному из каждой строки и из каждого столбца, причем это произведение берется со своим знаком, если подстановка (*) четна, и с противоположным, если нечетна. Определитель первого порядка равен своему единственному элементу.

Число всех членов определителя n-го порядка равно $n!$. Элементы, строки, столбцы и т. д. матрицы A называются соответственно элементами, строками, столбцами и т. д. ее определителя $|A|$.

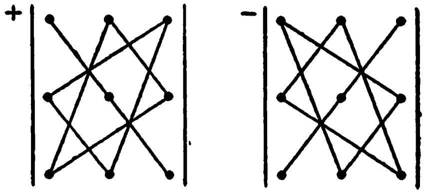
Пример 1. Определитель второго порядка:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}.$$

Пример 2. Определитель третьего порядка:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}.$$

Это выражение получается по правилу треугольников (правилу Саррюса), его можно пояснить следующими схемами, на которых элементы, входящие в одно произведение с указанным знаком, соединены отрезками:



Пример 3. Произведение элементов главной диагонали $a_{11}a_{22} \dots a_{nn}$ при любом n входит в определитель n-го порядка со знаком плюс, так как ни первые, ни вторые индексы не образуют инверсий.

Пример 4. Определить знак произведения элементов побочной диагонали $a_{1n}a_{2,n-1}a_{3,n-2} \dots a_{n1}$ определителя n-го порядка.

Так как индексы строк идут в порядке возрастания, то знак определяется четностью перестановки индексов столбцов n, $n-1$, $n-2$, ..., 1, в которой любые два числа образуют инверсию, поэтому число всех инверсий равно $C_n^2 = \frac{n(n-1)}{2}$, и произведение элементов побочной диагонали входит в определитель n-го порядка со знаком $(-1)^{\frac{n(n-1)}{2}}$. При $n = 4k$, $4k+1$ получим знак плюс, а при $n = 4k+2$, $4k+3$ — минус. Таким образом, для порядков 2, 3 получим минус, для 4, 5 — плюс, для 6, 7 — минус, для 8, 9 — плюс и т. д.

Пример 5. Вычислить определитель

$$d = \begin{vmatrix} 0 & a_{12} & a_{13} & 0 \\ a_{21} & 0 & 0 & a_{24} \\ a_{31} & 0 & 0 & a_{34} \\ 0 & a_{42} & a_{43} & 0 \end{vmatrix}.$$

Если не брать слагаемых, равных нулю, то из первой строки надо взять a_{12} или a_{13}. Если берем a_{12}, то из второй строки надо взять a_{21} или a_{24}; если берем a_{21}, то из третьей строки надо взять a_{34} и из четвертой a_{43}. Перебирая так все возможности и учитывая знаки, получим:

$$d = a_{12}a_{21}a_{34}a_{43} - a_{12}a_{24}a_{31}a_{43} - a_{13}a_{21}a_{34}a_{42} + a_{13}a_{24}a_{31}a_{42}.$$

Вычислять более сложные определители путем непосредственного применения определения было бы весьма неудобно. Для этого применяются специальные методы (см. пп. 5, 6).

3. Простейшие свойства определителей. *Транспонированием* матрицы (а также определителя) называется перемена ролями строк и столбцов с сохранением их номеров. Таким образом, строки данной матрицы будут в той же последовательности столбцами транспонированной матрицы, и наоборот.

Пример 1. Транспонируя матрицу $A=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{pmatrix}$, получим матрицу

$$A'=\begin{pmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{pmatrix}.$$

В случае квадратной матрицы (или в случае определителя) транспонирование сводится к повороту матрицы (или определителя) на 180° вокруг главной диагонали.

Суммой нескольких строк одинаковой длины называется строка, каждый элемент которой равен сумме соответствующих элементов данных строк.

Произведением строки на число называется строка, каждый элемент которой получен из соответствующего элемента данной строки умножением его на данное число.

Линейной комбинацией нескольких строк одинаковой длины называется строка, равная сумме произведений данных строк на некоторые числа, называемые коэффициентами этой линейной комбинации. Если одна строка является линейной комбинацией других, то говорят также, что она *линейно выражается* через эти строки.

Несколько строк одинаковой длины называются *линейно независимыми*, если ни одна из них не выражается линейно через остальные.

Пример 2. Равенство

$$(1,\ -1,\ -3,\ -5)=3\,(1,\ 1,\ 1,\ 1)-2\,(1,\ 2,\ 3,\ 4)$$

означает, что первая строка является линейной комбинацией двух других.

Простейшие свойства определителей.

1. При транспонировании определителя его значение не изменяется.

2. Если все элементы какой-нибудь строки (или столбца) определителя равны нулю, то определитель равен нулю.

3. При перестановке двух любых строк (или столбцов) определитель меняет знак.

4. Если две строки (или два столбца) определителя одинаковы, то он равен нулю.

5. Если все элементы одной строки (или столбца) определителя умножить на одно и то же число, то и весь определитель умножится на это число. Иными словами, общий множитель всех элементов строки (или столбца) можно выносить за знак определителя.

6. Если все элементы одной строки определителя пропорциональны соответствующим элементам другой строки, то определитель равен нулю. То же верно и для столбцов.

7. Если все элементы i-й строки определителя представлены в виде суммы k слагаемых, то и весь определитель представляется в виде суммы k определителей, у которых все строки, кроме i-й, такие же, как в исходном определителе, а i-я строка в первом определителе состоит из первых слагаемых, во втором — из вторых и т. д. То же верно и для столбцов.

Например,

$$\begin{vmatrix} 5 & -3 & 1 \\ 3 & 2 & -7 \\ 4 & 0 & -2 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 1 \\ 1 & 1 & -6 \\ 4 & 0 & -2 \end{vmatrix} + \begin{vmatrix} 5 & -3 & 1 \\ 2 & 1 & -1 \\ 4 & 0 & -2 \end{vmatrix}.$$

8. Если к одной строке определителя прибавить одну или несколько других строк, умноженных на любые числа, то значение определителя не изменится. То же верно и для столбцов. В частности, к одной строке (столбцу) можно прибавить (или из нее вычесть) другую строку (соответственно столбец).

9. Если хотя бы одна строка определителя линейно выражается через другие его строки, то он равен нулю. Обратно, если определитель порядка $n \geqslant 2$ равен нулю, то хотя бы одна из его строк линейно выражается через другие строки. То же верно для столбцов.

Пример 3. Определитель D, в котором любые два элемента a_{jk} и a_{kj}, симметрично лежащие относительно главной диагонали, комплексно сопряжены, является сам числом действительным.

В самом деле, пусть $D = a + bi$. Транспонируя D, получим определитель D', все элементы которого сопряжены соответствующим элементам определителя D. Но определитель равен сумме произведений (с надлежащими знаками) своих элементов, и по свойствам комплексных чисел, сопряженных сумме и произведению нескольких чисел, получим: $D' = a - bi$. Так как $D = D'$, то $a + bi = a - bi$, откуда $b = 0$ и $D = a$ — действительное число.

Пример 4. Определитель D называется *кососимметрическим*, если любые два элемента, симметрично лежащие относительно главной диагонали, различаются только знаком, т. е. $a_{jk} = -a_{kj}$ (в частности, элементы главной диагонали равны нулю). Кососимметрический определитель D нечетного порядка n равен нулю. Действительно, вынося из каждой строки определителя D множитель -1, получим транспонированный определитель, равный D, т. е. $D = (-1)^n D$ и $D = 0$ в силу нечетности n.

4. Миноры и алгебраические дополнения элементов. Разложение определителя по строке или столбцу. *Минором* M_{ij} элемента a_{ij} определителя D порядка $n \geqslant 2$ называется определитель порядка $n-1$, полученный из D вычеркиванием i-й строки и j-го столбца, в пересечении которых стоит данный элемент a_{ij}.

Алгебраическим дополнением элемента a_{ij} называется число $A_{ij} = (-1)^{i+j} M_{ij}$.

Определитель равен сумме произведений элементов любой его строки (или столбца) на их алгебраические дополнения. В частности, если все элементы строки (или столбца), кроме одного, равны нулю, то определитель равен этому элементу, умноженному на его алгебраическое дополнение.

Пример 1. Минор элемента a_{23} определителя

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

есть

$$M_{23} = \begin{vmatrix} a_{11} & a_{12} & a_{14} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix}.$$

Алгебраическое дополнение элемента a_{23} есть $A_{23} = -M_{23}$.

Пример 2. Вычислить определитель

$$D = \begin{vmatrix} 2 & -3 & 4 & 1 \\ 4 & -2 & 3 & 2 \\ a & b & c & d \\ 3 & -1 & 4 & 3 \end{vmatrix}.$$

Разлагая по третьей строке, найдем

$$D=a(-1)^4\begin{vmatrix}-3 & 4 & 1\\ -2 & 3 & 2\\ -1 & 4 & 3\end{vmatrix}+b(-1)^5\begin{vmatrix}2 & 4 & 1\\ 4 & 3 & 2\\ 3 & 4 & 3\end{vmatrix}+c(-1)^6\begin{vmatrix}2 & -3 & 1\\ 4 & -2 & 2\\ 3 & -1 & 3\end{vmatrix}+$$

$$+d(-1)^7\begin{vmatrix}2 & -3 & 4\\ 4 & -2 & 3\\ 3 & -1 & 4\end{vmatrix}=8a+15b+12c-19d.$$

5. Вычисление определителей с числовыми элементами. Вычисление определителя порядка n можно свести к вычислению одного определителя порядка $n-1$. Пусть надо вычислить определитель

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

Если все элементы первого столбца равны нулю, то $D=0$ (см. стр. 18, свойство 2); если $a_{11}=0$, но $a_{k1}\neq 0$, то, переставив первую и k-ю строки, получим в левом верхнем углу элемент, отличный от нуля. Итак, без ограничения общности можно считать, что $a_{11}\neq 0$. Ко второй строке прибавим первую, умноженную на $-\frac{a_{21}}{a_{11}}$, к третьей строке первую, умноженную на $-\frac{a_{31}}{a_{11}}, \dots,$ к n-й строке первую, умноженную на $-\frac{a_{n1}}{a_{11}}$.

Определитель не изменит своего значения (см. стр. 19, свойство 8) и примет вид

$$D=\begin{vmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ 0 & b_{22} & \dots & b_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ 0 & b_{n2} & \dots & b_{nn}\end{vmatrix}=a_{11}|b_{ij}|_1^{n-1} \quad \text{(см. п. 4).}$$

Пример. Вычислить определитель

$$D=\begin{vmatrix}6 & -5 & 8 & 4\\ 9 & 7 & 5 & 2\\ 7 & 5 & 3 & 7\\ -4 & 8 & -8 & -3\end{vmatrix}.$$

Имеем

$$D=\begin{vmatrix} -12 & -19 & -2 & 0 \\ 9 & 7 & 5 & 2 \\ -20 & -16 & -12 & 1 \\ 5 & 15 & -3 & -1 \end{vmatrix}=\begin{vmatrix} -12 & -19 & -2 & 0 \\ 19 & 37 & -1 & 0 \\ -15 & -1 & -15 & 0 \\ 5 & 15 & -3 & -1 \end{vmatrix}=$$

$$=-\begin{vmatrix} -12 & -19 & -2 \\ 19 & 37 & -1 \\ -15 & -1 & -15 \end{vmatrix}=-\begin{vmatrix} -10 & -19 & 283 \\ 20 & 37 & -556 \\ 0 & -1 & 0 \end{vmatrix}=$$

$$=-\begin{vmatrix} -10 & 283 \\ 20 & -556 \end{vmatrix}=-\begin{vmatrix} -10 & 283 \\ 0 & 10 \end{vmatrix}=100.$$

6. Методы вычисления определителей n-го порядка. Изложенный в предыдущем пункте метод вычисления определителей данного порядка с числовыми элементами неудобен в случае определителей с буквенными элементами или определителей с числовыми элементами, но высокого порядка n. Общего метода вычисления таких определителей не существует, если не считать вычисления определителя заданного порядка непосредственно по его определению (стр. 16). К определителям того или иного специального вида применяются различные методы вычисления, приводящие к выражениям более простым, чем выражения, получаемые прямо из определения. Ниже указаны некоторые из этих методов.

1) М е т о д п р и в е д е н и я к т р е у г о л ь н о м у в и д у. Пользуясь свойствами определителя (см. стр. 18—19), приводят его к так называемому «треугольному» виду, когда все элементы, стоящие по одну сторону от главной диагонали, равны нулю. Полученный определитель треугольного вида равен произведению элементов главной диагонали, т. е.

$$D=\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}=a_{11}a_{22}a_{33}\dots a_{nn}.$$

Случай, когда нули стоят справа от главной диагонали, сводится транспонированием к предыдущему или разбирается аналогично. Если удобнее получить нули по одну сторону от побочной диагонали, то, меняя в таком определителе D порядок строк или столбцов на обратный, найдем $D=(-1)^{\frac{n(n-1)}{2}}D_1$, где n — порядок D и D_1 уже треугольного вида.

Пример 1.

$$D=\begin{vmatrix} n & n-1 & \dots & 4 & 3 & 2 & 1 \\ n & n-1 & \dots & 4 & 3 & 3 & 1 \\ n & n-1 & \dots & 4 & 5 & 2 & 1 \\ n & n-1 & \dots & 7 & 3 & 2 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 2n-1 & n-1 & \dots & 4 & 3 & 2 & 1 \end{vmatrix}$$

(по побочной диагонали стоят первые n нечетных чисел). Вычитая первую строку из всех остальных, найдем

$$D=\begin{vmatrix} n & n-1 & \dots & 4 & 3 & 2 & 1 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 2 & 0 & 0 \\ 0 & 0 & \dots & 3 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ n-1 & 0 & \dots & 0 & 0 & 0 & 0 \end{vmatrix}=$$

$$=(-1)^{\frac{n(n-1)}{2}}\begin{vmatrix} 1 & 2 & 3 & 4 & \dots & n-1 & n \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & n-1 \end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}(n-1)!$$

Пример 2. Вычислить определитель порядка n:

$$D=\begin{vmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \dots & \dots & \dots & \dots & \dots \\ b & b & b & \dots & a \end{vmatrix}.$$

К первому столбцу прибавляем все остальные:

$$D=\begin{vmatrix} a+(n-1)b & b & b & \dots & b \\ a+(n-1)b & a & b & \dots & b \\ a+(n-1)b & b & a & \dots & b \\ \dots & \dots & \dots & \dots & \dots \\ a+(n-1)b & b & b & \dots & a \end{vmatrix}=$$

(первую строку вычитаем из всех остальных)

$$=\begin{vmatrix} a+(n-1)b & b & b & \dots & b \\ 0 & a-b & 0 & \dots & 0 \\ 0 & 0 & a-b & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a-b \end{vmatrix} = [a+(n-1)b](a-b)^{n-1}.$$

2) Метод рекуррентных соотношений. Данный определитель D_n порядка n, после преобразования и разложения по строке или столбцу, выражается через определители того же вида, но более низких порядков. Полученное равенство называется *рекуррентным соотношением*. Вычисляют столько определителей данного вида начальных порядков, сколько их входит в правую часть рекуррентного соотношения. Далее вычисляют определители следующих порядков, используя рекуррентное соотношение, до тех пор, пока не удастся заметить общую закономерность для получаемых выражений. Затем, пользуясь рекуррентным соотношением, можно доказать эту закономерность для D_n индукцией по n.

Пример 3. Вычислить определитель

$$J_n=\begin{vmatrix} 3 & 2 & 0 & 0 & \dots & 0 & 0 \\ 1 & 3 & 2 & 0 & \dots & 0 & 0 \\ 0 & 1 & 3 & 2 & \dots & 0 & 0 \\ 0 & 0 & 1 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & 3 \end{vmatrix}.$$

Рекуррентное соотношение:

$$J_n=3J_{n-1}-2J_{n-2}.$$

Находим $J_1=3$, $J_2=7$, $J_3=3J_2-2J_1=15$, $J_4=3J_3-2J_2=31$.

Найденные определители вычисляются по формуле $J_m=2^{m+1}-1$, докажем ее справедливость индукцией по числу m. Для $m=1$ формула верна; предполагая ее верной для всех $m<n$, в частности, полагая $J_{n-1}=2^n-1$, $J_{n-2}=2^{n-1}-1$, из рекуррентного соотношения будем иметь

$$J_n=3J_{n-1}-2J_{n-2}=3(2^n-1)-2(2^{n-1}-1)=2^{n+1}-1,$$

что и требовалось.

3) Специальный случай рекуррентных соотношений. Пусть рекуррентное соотношение имеет вид

$$D_n = pD_{n-1} + qD_{n-2},$$

где p и q — постоянные (т. е. не зависящие от n) числа. В этом случае можно вывести формулу для вычисления D_n ([20], стр. 33, 34).

Если $q = 0$, то $D_n = p^{n-1}D_1$, где D_1 — определитель первого порядка данного вида.

Если $q \neq 0$, то решаем квадратное уравнение $x^2 - px - q = 0$. Пусть α и β — его корни. Если $\alpha \neq \beta$, то $D_n = c_1\alpha^n + c_2\beta^n$, где

$$c_1 = \frac{D_2 - \beta D_1}{\alpha(\alpha - \beta)}, \qquad c_2 = -\frac{D_2 - \alpha D_1}{\beta(\alpha - \beta)}.$$

D_1, D_2 — определители первого и второго порядков данного вида. Выражения для c_1 и c_2 можно найти непосредственно из равенств

$$D_1 = c_1\alpha + c_2\beta, \quad D_2 = c_1\alpha^2 + c_2\beta^2.$$

Если $q \neq 0$, но $\alpha = \beta$, то

$$D_n = \alpha^n [(n-1)c_1 + c_2],$$

где

$$c_1 = \frac{D_2 - \alpha D_1}{\alpha^2} \quad \text{и} \quad c_2 = \frac{D_1}{\alpha}.$$

Пример 4. Вычислить определитель Якоби (стр. 28) n-го порядка:

$$J_n = \begin{vmatrix} 5 & 3 & 0 & 0 & \dots & 0 & 0 \\ 2 & 5 & 3 & 0 & \dots & 0 & 0 \\ 0 & 2 & 5 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 2 & 5 \end{vmatrix}.$$

Рекуррентное соотношение:

$$J_n = 5J_{n-1} - 6J_{n-2}.$$

Решаем уравнение $x^2 - 5x + 6 = 0$. Его корни: $\alpha = 2$, $\beta = 3$. Далее находим $J_1 = 5$, $J_2 = 19$. Поэтому для c_1 и c_2 имеем систему уравнений

$$2c_1 + 3c_2 = 5, \quad 4c_1 + 9c_2 = 19,$$

откуда $c_1 = -2$, $c_2 = 3$ и, значит,

$$J_n = c_1\alpha^n + c_2\beta^n = 3^{n+1} - 2^{n+1}.$$

7. Определитель Вандермонда и связанные с ним определители. 1) *Определителем Вандермонда* называется определитель вида

$$D_n=\begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & x_2^3 & \dots & x_2^{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & x_n^2 & x_n^3 & \dots & x_n^{n-1} \end{vmatrix}.$$

Он вычисляется по формуле

$$D_n=(x_2-x_1)(x_3-x_1)\dots(x_n-x_1)(x_3-x_2)(x_4-x_2)\dots$$
$$\dots(x_n-x_2)\dots(x_n-x_{n-1})=\prod_{n\geq i>k\geq 1}(x_i-x_k)$$

(см. [11], стр. 50 или [20], стр. 32).

Пример 1.

$$D=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 3 & \dots & n+1 \\ 1 & 2^2 & 3^2 & \dots & (n+1)^2 \\ 1 & 2^3 & 3^3 & \dots & (n+1)^3 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 2^n & 3^n & \dots & (n+1)^n \end{vmatrix}.$$

Этот определитель получается транспонированием из определителя Вандермонда с определяющими элементами

$$x_1=1,\quad x_2=2,\dots,x_{n+1}=n+1.$$

Поэтому

$$D=(2-1)(3-1)\dots[(n+1)-1]\times$$
$$\times(3-2)(4-2)\dots[(n+1)-2]\dots[(n+1)-n]=1!\,2!\,3!\dots n!.$$

Пример 2.

$$D=\begin{vmatrix} a_1^n & a_1^{n-1}b_1 & a_1^{n-2}b_1^2 & \dots & b_1^n \\ a_2^n & a_2^{n-1}b_2 & a_2^{n-2}b_2^2 & \dots & b_2^n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n+1}^n & a_{n+1}^{n-1}b_{n+1} & a_{n+1}^{n-2}b_{n+1}^2 & \dots & b_{n+1}^n \end{vmatrix}.$$

Вынося за определитель из первой строки a_1^n, из второй — a_2^n и т. д., придем к определителю Вандермонда с определяющими элементами $\frac{b_1}{a_1}, \frac{b_2}{a_2}, \ldots, \frac{b_{n+1}}{a_{n+1}}$. Поэтому

$$D = a_1^n a_2^n \ldots a_{n+1}^n \prod_{n+1 \geq i > k \geq 1} \left(\frac{b_i}{a_i} - \frac{b_k}{a_k} \right) = \prod_{n+1 \geq i > k \geq 1} (a_k b_i - a_i b_k).$$

2) К определителю Вандермонда приводится определитель вида

$$D = \begin{vmatrix} 1 & x_1 & x_1^2 \ldots & x_1^{p-1} & x_1^{p+1} & \ldots & x_1^n \\ 1 & x_2 & x_2^2 \ldots & x_2^{p-1} & x_2^{p+1} & \ldots & x_2^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 \ldots & x_n^{p-1} & x_n^{p+1} & \ldots & x_n^n \end{vmatrix} \quad (1 \leqslant p \leqslant n-1).$$

Если через s_{n-p} обозначим сумму всевозможных произведений элементов $x_1, x_2, \ldots, x_n$, взятых в количестве $n-p$, и через D_n определитель Вандермонда из тех же элементов, то $D = s_{n-p} D_n$ ([20], задачи 344, 346).

Пример 3.

$$D = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-2} & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-2} & x_2^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-2} & x_n^n \end{vmatrix} =$$

$$= (x_1 + x_2 + \ldots + x_n) \prod_{n \geq i > k \geq 1} (x_i - x_k).$$

Здесь $p = n - 1$.

3) *Обобщенным определителем Вандермонда* называется определитель

$$D_n = \begin{vmatrix} a_1^{\alpha_1} & a_1^{\alpha_2} & \ldots & a_1^{\alpha_n} \\ a_2^{\alpha_1} & a_2^{\alpha_2} & \ldots & a_2^{\alpha_n} \\ \cdot & \cdot & \cdot & \cdot \\ a_n^{\alpha_1} & a_n^{\alpha_2} & \ldots & a_n^{\alpha_n} \end{vmatrix},$$

где $a_1, a_2, \ldots, a_n$ — положительные и $\alpha_1, \alpha_2, \ldots, \alpha_n$ — действительные числа. Этот определитель тогда и только тогда равен нулю, когда среди чисел a_i или среди чисел α_i имеются равные. Достаточно рассмотреть случай, когда $0 < a_1 < a_2 < \ldots < a_n$ и $\alpha_1 < \alpha_2 < \ldots < \alpha_n$.

Для этого случая можно доказать, что $D_n > 0$ (см [7], стр. 93).

8. Якобиевы определители. Континуант и связь с числами Фибоначчи. О якобиевых матрицах см. гл. II, § 1, п. 13. *Якобиевым определителем* называется определитель вида

$$J_n=\begin{vmatrix} a_1 & b_1 & 0 & 0 & \dots & 0 & 0 \\ -c_2 & a_2 & b_2 & 0 & \dots & 0 & 0 \\ 0 & -c_3 & a_3 & b_3 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & 0 & \dots & -c_n & a_n \end{vmatrix}.$$

Разлагая его по последней строке, найдем рекуррентное соотношение $J_n=a_nJ_{n-1}+b_{n-1}c_nJ_{n-2}$.

Можно доказать следующий закон составления развернутого выражения для J_n: берется сумма всевозможных произведений, одно из которых есть $a_1a_2a_3\dots a_n$, а другие получаются из него путем замены одной или нескольких пар соседних множителей $a_i a_{i+1}$ на $b_i c_{i+1}$.

Пример 1.

$$J_5=a_1a_2a_3a_4a_5+b_1c_2a_3a_4a_5+a_1b_2c_3a_4a_5+ \\ +a_1a_2b_3c_4a_5+a_1a_2a_3b_4c_5+b_1c_2b_3c_4a_5+ \\ +b_1c_2a_3b_4c_5+a_1b_2c_3b_4c_5.$$

Частным случаем якобиевого определителя является так называемый *континуант*

$$(a_1a_2a_3\dots a_n)=\begin{vmatrix} a_1 & 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & a_2 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & a_3 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & 1 \\ 0 & 0 & 0 & 0 & \dots & -1 & a_n \end{vmatrix}.$$

Его название объясняется связью с цепными дробями, выражаемой равенством

$$\frac{(a_1a_2a_3\dots a_n)}{(a_2a_3\dots a_n)}=a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\ddots+\cfrac{1}{a_n}}}.$$

(см. [30], стр. 29 или [20], задача 539). Он равен сумме произведения $a_1 a_2 \ldots a_n$ и всевозможных произведений, полученных из него вычеркиванием пар соседних множителей (при четном n результат вычеркивания всех n сомножителей считается равным единице).

Пример 2.

$$(a_1a_2a_3a_4) = a_1a_2a_3a_4 + a_3a_4 + a_1a_4 + a_1a_2 + 1;$$

$$(a_1a_2a_3a_4a_5) = \\ = a_1a_2a_3a_4a_5 + a_3a_4a_5 + a_1a_4a_5 + a_1a_2a_5 + a_1a_2a_3 + a_5 + a_3 + a_1.$$

Частный вид континуанта

$$u_n = \begin{vmatrix} 1 & 1 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \ldots & -1 & 1 \end{vmatrix}$$

равен n-му числу так называемого *ряда Фибоначчи*

$$1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ \ldots,$$

в котором каждое число, начиная с третьего, равно сумме двух предшествующих чисел ([20], задача 365).

О других определителях специального вида см. § 1, пп. 7, 12, 13, 14, а также [20] и [24].

9. Миноры любого порядка, их алгебраические дополнения и теорема Лапласа. *Минором k-го порядка* в определителе D порядка n $(1 \leqslant k \leqslant n)$ называется определитель M, составленный из элементов, стоящих на пересечении любых k строк и любых k столбцов определителя D. В частности, минор n-го порядка в D равен самому D. Минор нулевого порядка по определению считается равным единице. *Дополнительным минором* для минора M в определителе D называется минор M', полученный из D вычеркиванием тех k строк и k столбцов, в которых стоит данный минор M. *Алгебраическим дополнением* минора M в определителе D называется его дополнительный минор M', взятый со знаком плюс, если сумма номеров строк и столбцов определителя D, в которых стоит исходный минор M, четна, и со знаком минус, если нечетна.

Пример 1. В определителе $D_4=|a_{ij}|_1^4$ для минора второго порядка

$$M=\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

дополнительным минором будет

$$M'=\begin{vmatrix} a_{12} & a_{14} \\ a_{42} & a_{44} \end{vmatrix}$$

и алгебраическим дополнением $A=(-1)^9M'=-M'$.

Теорема Лапласа. Пусть в определителе D порядка n выделены любые k строк (или столбцов) $(1\leqslant k\leqslant n-1)$. Тогда определитель D равен сумме произведений всех миноров k-го порядка, стоящих в выделенных k строках (или столбцах), на их алгебраические дополнения (см. [11], стр. 51).

Обобщение теоремы Лапласа на случай разложения по нескольким группам строк (или столбцов) см. в [20], задача 466.

Пример 2. Вычислить определитель

$$D=\begin{vmatrix} 2 & 1 & 4 & 3 & 5 \\ 3 & 4 & 0 & 5 & 0 \\ 3 & 4 & 5 & 2 & 1 \\ 1 & 5 & 2 & 4 & 3 \\ 4 & 6 & 0 & 7 & 0 \end{vmatrix}.$$

Здесь десять миноров второго порядка, стоящих во второй и пятой строках, но только три из них отличны от нуля. Разлагая по этим двум строкам, найдем

$$D=(-1)^{10}\begin{vmatrix} 3 & 4 \\ 4 & 6 \end{vmatrix}\cdot\begin{vmatrix} 4 & 3 & 5 \\ 5 & 2 & 1 \\ 2 & 4 & 3 \end{vmatrix}+(-1)^{12}\begin{vmatrix} 3 & 5 \\ 4 & 7 \end{vmatrix}\cdot\begin{vmatrix} 1 & 4 & 5 \\ 4 & 5 & 1 \\ 5 & 2 & 3 \end{vmatrix}+$$

$$+(-1)^{13}\begin{vmatrix} 4 & 5 \\ 6 & 7 \end{vmatrix}\cdot\begin{vmatrix} 2 & 4 & 5 \\ 3 & 5 & 1 \\ 1 & 2 & 3 \end{vmatrix}=2\cdot 49+1(-100)+2(-1)=-4.$$

В частности, если в определителе D главную диагональ покрывают две квадратные матрицы без общих элементов с определителями D_1 и D_2 и по одну сторону от них все элементы равны нулю, то $D=D_1D_2$.

Пример 3.

$$\begin{vmatrix} 3 & 2 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 6 & 7 & 2 & 3 \\ 8 & 9 & 3 & 7 \end{vmatrix} = \begin{vmatrix} 3 & 2 \\ 4 & 5 \end{vmatrix} \cdot \begin{vmatrix} 2 & 3 \\ 3 & 7 \end{vmatrix} = 7 \cdot 5 = 35.$$

Более общим является случай так называемых *ступенчатых* определителей: если в определителе D на главной диагонали стоит цепочка квадратных матриц с определителями $D_1, D_2, \ldots, D_k$, а по одну сторону от этой цепочки все элементы равны нулю, то $D = D_1 D_2 \ldots D_k$.

Пример 4.

$$\begin{vmatrix} 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 & 0 \\ 5 & 2 & 3 & 0 & 0 & 0 & 0 \\ 7 & 6 & 2 & 5 & 4 & 0 & 0 \\ 3 & 2 & 7 & 4 & 5 & 0 & 0 \\ 5 & 1 & 3 & 2 & 6 & 7 & 3 \\ 2 & 7 & 1 & 5 & 3 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \cdot 3 \cdot \begin{vmatrix} 5 & 4 \\ 4 & 5 \end{vmatrix} \cdot \begin{vmatrix} 7 & 3 \\ 4 & 1 \end{vmatrix} =$$

$$= (-2) \cdot 3 \cdot 9 \cdot (-5) = 270.$$

10. Разложение определителя по строке и столбцу. Окаймленные определители. *Окаймленным определителем для определителя*

$$D = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

называется определитель

$$D_1 = \begin{vmatrix} a_{11} & \ldots & a_{1n} & x_1 \\ \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} & x_n \\ y_1 & \ldots & y_n & z \end{vmatrix}.$$

Любой определитель порядка $n > 1$ можно рассматривать как окаймленный относительно минора M_{nn} элемента, стоящего в правом нижнем углу. Разлагая D_1 по последнему столбцу, а затем все полученные миноры, кроме минора для z, по последней строке, получим разложение этого определителя по последнему столбцу и последней строке:

$$D_1 = Dz - \sum_{i, j=1}^{n} A_{ij} x_i y_j,$$

где A_{ij} — алгебраическое дополнение элемента a_{ij} в определителе D (см. [30], стр. 22 или [23], стр. 61).

Если надо разложить определитель по любой строке и любому столбцу, то, переставляя данные строку и столбец со всеми последующими, приходят к рассмотренному случаю.

Разложение по строке и столбцу удобно применять в случае, когда легко вычисляются алгебраические дополнения всех элементов определителя D.

Пример 1. Вычислить определитель

$$D_1 = \begin{vmatrix} a_1 & 0 & \ldots & 0 & x_1 \\ 0 & a_2 & \ldots & 0 & x_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_n & x_n \\ y_1 & y_2 & \ldots & y_n & z \end{vmatrix}.$$

Здесь

$$D = \begin{vmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_n \end{vmatrix},$$

$A_{ij} = 0$ при $i \neq j$, $A_{ii} = \dfrac{D}{a_i}$. Поэтому

$$D_1 = D\left(z - \sum_{i=1}^{n} \frac{x_i y_i}{a_i}\right),$$

где $D = a_1 a_2 \ldots a_n$.

Рассмотрим окаймленный определитель

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & x_{11} & \ldots & x_{1p} \\ a_{21} & a_{22} & \ldots & a_{2n} & x_{21} & \ldots & x_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & x_{n1} & \ldots & x_{np} \\ y_{11} & y_{12} & \ldots & y_{1n} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ y_{p1} & y_{p2} & \ldots & y_{pn} & 0 & \ldots & 0 \end{vmatrix}.$$

Разлагая этот определитель по теореме Лапласа по последним p столбцам, а затем (при $n \geqslant p$) каждый из дополнительных миноров порядка n по последним p строкам, получим:

1) Если $p \leqslant n$, то

$$\Delta = (-1)^p \sum A_{i_1 i_2 \ldots i_p j_1 j_2 \ldots j_p} X_{i_1 i_2 \ldots i_p} Y_{j_1 j_2 \ldots j_p},$$

где $A_{i_1 i_2 \ldots i_p j_1 j_2 \ldots j_p}$ — алгебраическое дополнение минора $M_{i_1 i_2 \ldots i_p j_1 j_2 \ldots j_p}$ в определителе $D = |a_{ij}|_1^n$,

$$X_{i_1 i_2 \ldots i_p} = \begin{vmatrix} x_{i_1 1} & x_{i_1 2} & \ldots & x_{i_1 p} \\ x_{i_2 1} & x_{i_2 2} & \ldots & x_{i_2 p} \\ \cdot & \cdot & \cdot & \cdot \\ x_{i_p 1} & x_{i_p 2} & \ldots & x_{i_p p} \end{vmatrix},$$

$$Y_{j_1 j_2 \ldots j_p} = \begin{vmatrix} y_{1 j_1} & y_{1 j_2} & \ldots & y_{1 j_p} \\ y_{2 j_1} & y_{2 j_2} & \ldots & y_{2 j_p} \\ \cdot & \cdot & \cdot & \cdot \\ y_{p j_1} & y_{p j_2} & \ldots & y_{p j_p} \end{vmatrix}$$

и сумма берется по всем индексам $i_1, i_2, \ldots, i_p$; $j_1, j_2, \ldots, j_p$, удовлетворяющим условиям

$$1 \leqslant \begin{matrix} i_1 < i_2 < \ldots < i_p \\ j_1 < j_2 < \ldots < j_p \end{matrix} \leqslant n.$$

Такое же разложение имеет место при окаймлении определителя D с помощью первых p строк и первых p столбцов.

2) Если $p > n$, то $\Delta = 0$.

Пример 2.

$$\Delta = \begin{vmatrix} a_1 & 0 & 0 & x_{11} & x_{12} \\ 0 & a_2 & 0 & x_{21} & x_{22} \\ 0 & 0 & a_3 & x_{31} & x_{32} \\ y_{11} & y_{12} & y_{13} & 0 & 0 \\ y_{21} & y_{22} & y_{23} & 0 & 0 \end{vmatrix} =$$

$$= a_1 \begin{vmatrix} x_{21} & x_{22} \\ x_{31} & x_{32} \end{vmatrix} \begin{vmatrix} y_{12} & y_{13} \\ y_{22} & y_{23} \end{vmatrix} + a_2 \begin{vmatrix} x_{11} & x_{12} \\ x_{31} & x_{32} \end{vmatrix} \begin{vmatrix} y_{11} & y_{13} \\ y_{21} & y_{23} \end{vmatrix} +$$

$$+ a_3 \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} \begin{vmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{vmatrix}.$$

11. Умножение определителей. Теорема Бине — Коши. *Произведением определителей* $|A|=|a_{ij}|_1^n$ и $|B|=|b_{ij}|_1^n$ одного и того же порядка n называется определитель $|C|=$ $=|c_{ij}|_1^n$ того же порядка, все элементы которого вычисляются по одной из следующих четырех формул:

$$1)\ c_{ij}=a_{i1}b_{j1}+a_{i2}b_{j2}+\ldots+a_{in}b_{jn};$$
$$2)\ c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{in}b_{nj};$$
$$3)\ c_{ij}=a_{1i}b_{j1}+a_{2i}b_{j2}+\ldots+a_{ni}b_{jn};$$
$$4)\ c_{ij}=a_{1i}b_{1j}+a_{2i}b_{2j}+\ldots+a_{ni}b_{nj}.$$

В случае 1) элемент c_{ij} получается как сумма произведений элементов i-й строки определителя $|A|$ на соответствующие элементы j-й строки определителя $|B|$. Поэтому говорят, что в данном случае произведение получено умножением строк первого определителя на строки второго. Во втором случае — путем умножения строк на столбцы, в третьем — столбцов на строки и в четвертом — столбцов на столбцы. Значение определителя $|C|$ одно и то же во всех четырех случаях, хотя элементы c_{ij} различны: $|C|=|A|\cdot|B|$. Умножение определителей тесно связано с умножением матриц (см. стр. 67—68).

Пример 1. Перемножая четырьмя способами определители

$$\begin{vmatrix} 2 & 1 \\ 4 & 3 \end{vmatrix}=2, \quad \begin{vmatrix} 4 & -3 \\ -1 & 2 \end{vmatrix}=5,$$

получим:

$$1)\ \begin{vmatrix} 5 & 0 \\ 7 & 2 \end{vmatrix}=10, \quad 2)\ \begin{vmatrix} 7 & -4 \\ 13 & -6 \end{vmatrix}=10,$$
$$3)\ \begin{vmatrix} -4 & 6 \\ -5 & 5 \end{vmatrix}=10, \quad 4)\ \begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix}=10.$$

Умножение определителей можно использовать для вычисления некоторых определителей специального вида.

Пример 2. Вычислить определитель

$$D_x=\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ -x_2 & x_1 & x_4 & -x_3 \\ -x_3 & -x_4 & x_1 & x_2 \\ -x_4 & x_3 & -x_2 & x_1 \end{vmatrix}.$$

Возводя его в квадрат путем умножения строк на строки, получим на главной диагонали одно и то же выражение $x_1^2+x_2^2+$ $+x_3^2+x_4^2$, а вне главной диагонали нули. Поэтому $D_x^2=(x_1^2+x_2^2+$ $+x_3^2+x_4^2)^4$. Так как D_x содержит произведение по главной диагонали, равное x_1^4, то из обеих частей последнего равенства можно извлечь корень со знаком плюс, откуда $D_x=(x_1^2+x_2^2+x_3^2+x_4^2)^2$.

Пример 3. Вычислить определитель

$$D=\begin{vmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & s_4 & \dots & s_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{n-1} & s_n & s_{n+1} & \dots & s_{2n-2} \end{vmatrix},$$

где $s_k=x_1^k+x_2^k+\dots+x_n^k$ $(k=0, 1, 2, \dots)$ — так называемые *степенные суммы* переменных $x_1, x_2, \dots x_n$ (в частности, $s_0=n$).

Умножая сам на себя определитель Вандермонда

$$V_n=\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}$$

путем умножения столбцов на столбцы и применяя выражение для определителя Вандермонда, данное на стр. 26, получим

$$D=V_n^2=\prod_{n\geqslant i>j\geqslant 1}(x_i-x_j)^2.$$

Теорема Бине—Коши. Пусть даны две прямоугольные матрицы

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{p1} & b_{p2} & \dots & b_{pn} \end{pmatrix}$$

одинаковых размеров $p\times n$. Умножая строки A на строки B, составим p^2 чисел

$$c_{ij}=a_{i1}b_{j1}+a_{i2}b_{j2}+\dots+a_{in}b_{jn} \qquad (i, j=1, 2, \dots, p)$$

и рассмотрим определитель $D=|c_{ij}|_1^p$ порядка p из этих чисел.

Тогда: 1) Если $p\leqslant n$, то

$$D=\sum_{1\leqslant i_1<i_2<\dots<i_p\leqslant n} A(i_1i_2\dots i_p)\,B(i_1i_2\dots i_p),$$

где $A(i_1 i_2 \ldots i_p)$ и $B(i_1 i_2 \ldots i_p)$ — миноры порядка p соответственно матриц A и B, составленные из столбцов с номерами $i_1, i_2, \ldots, i_p$.

2) Если $p > n$, то $D = 0$ (см. [6], стр. 17).

Правило умножения определителей является частным случаем теоремы Бине — Коши при $p = n$.

Пример 4. Применяя теорему Бине — Коши к матрицам

$$\begin{pmatrix} a_1 & a_2 \ldots a_n \\ b_1 & b_2 \ldots b_n \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} c_1 & c_2 \ldots c_n \\ d_1 & d_2 \ldots d_n \end{pmatrix},$$

получим равенство

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

12. Циркулянты и связанные с ними определители. *Циркулянтом* (или циклическим определителем) называется каждый из двух следующих определителей:

$$D_n = \begin{vmatrix} a_1 & a_2 & a_3 & \ldots & a_{n-1} & a_n \\ a_n & a_1 & a_2 & \ldots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_n & a_1 & \ldots & a_{n-3} & a_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_2 & a_3 & a_4 & \ldots & a_n & a_1 \end{vmatrix},$$

$$D'_n = \begin{vmatrix} a_1 & a_2 & a_3 & \ldots & a_{n-1} & a_n \\ a_2 & a_3 & a_4 & \ldots & a_n & a_1 \\ a_3 & a_4 & a_5 & \ldots & a_1 & a_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & a_1 & a_2 & \ldots & a_{n-2} & a_{n-1} \end{vmatrix}.$$

Для первого из них имеет место выражение

$$D_n = f(\varepsilon_0) f(\varepsilon_1) f(\varepsilon_2) \ldots f(\varepsilon_{n-1}),$$

где

$$f(x) = a_1 + a_2 x + a_3 x^2 + \ldots + a_n x^{n-1}$$

и

$$\varepsilon_k = \cos\frac{2\pi k}{n} + i \sin\frac{2\pi k}{n}, \qquad k = 0, 1, 2, \ldots, n-1,$$

— все значения $\sqrt[n]{1}$ (см. [30], стр. 36 или [20], задача 479).

Изменением порядка строк с номерами 2, 3, ..., n определителя D_n на обратный получается выражение для второго определителя:

$$D'_n = (-1)^{\frac{(n-1)(n-2)}{2}} D_n.$$

Пример. Вычислить определитель

$$D = \begin{vmatrix} 1 & a & a^2 & \dots & a^{n-1} \\ a^{n-1} & 1 & a & \dots & a^{n-2} \\ a^{n-2} & a^{n-1} & 1 & \dots & a^{n-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a & a^2 & a^3 & \dots & 1 \end{vmatrix}.$$

Здесь

$$f(x) = 1 + ax + a^2x^2 + \dots + a^{n-1}x^{n-1} = \frac{1 - a^n x^n}{1 - ax}.$$

Пусть сначала $a \neq 0$ и $a^n \neq 1$. Так как $\varepsilon_i^n = 1$, $i = 0, 1, 2, \dots, n-1$, то

$$D = f(\varepsilon_0) f(\varepsilon_1) \dots f(\varepsilon_{n-1}) = \frac{(1-a^n)^n}{(1-a\varepsilon_0)(1-a\varepsilon_1)\dots(1-a\varepsilon_{n-1})}.$$

Для упрощения знаменателя в равенстве

$$x^n - 1 = (x - \varepsilon_0)(x - \varepsilon_1) \dots (x - \varepsilon_{n-1})$$

положим $x = \frac{1}{a}$. Получим

$$\frac{1}{a^n} - 1 = \left(\frac{1}{a} - \varepsilon_0\right)\left(\frac{1}{a} - \varepsilon_1\right) \dots \left(\frac{1}{a} - \varepsilon_{n-1}\right)$$

или, умиожая на a^n, найдем

$$1 - a^n = (1 - a\varepsilon_0)(1 - a\varepsilon_1) \dots (1 - a\varepsilon_{n-1}).$$

Иначе это можно доказать, раскрывая скобки в правой части и применяя формулы Виета для $x^n - 1$. Следовательно,

$$D_n = (1 - a^n)^{n-1}.$$

Эта формула верна также при $a = 0$, так как обе ее части обращаются в единицу, и при $a^n = 1$, так как при этом условии, умножая вторую строку D на a, получим строку, совпадающую с первой, и потому $D = 0$.

Косым циркулянтом называется определитель вида

$$Z_n = \begin{vmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ -a_n & a_1 & a_2 & \dots & a_{n-1} \\ -a_{n-1} & -a_n & a_1 & \dots & a_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_2 & -a_3 & -a_4 & \dots & a_1 \end{vmatrix}.$$

Для него имеет место выражение

$$Z_n = f(\eta_0) f(\eta_1) f(\eta_2) \ldots f(\eta_{n-1}),$$

где

$$f(x) = a_1 + a_2 x + a_3 x^2 + \ldots + a_n x^{n-1}$$

и

$$\eta_k = \cos\frac{(2k+1)\pi}{n} + i\sin\frac{(2k+1)\pi}{n}, \qquad k = 0,\ 1,\ 2,\ \ldots,\ n-1,$$

— все значения корня $\sqrt[n]{-1}$ (см. [20], задача 493).

Как циркулянт, так и косой циркулянт являются частными видами определителя

$$\begin{vmatrix} a_1 & a_2 & a_3 & \ldots & a_n \\ a_n z & a_1 & a_2 & \ldots & a_{n-1} \\ a_{n-1} z & a_n z & a_1 & \ldots & a_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_2 z & a_3 z & a_4 z & \ldots & a_1 \end{vmatrix} = f(\alpha_0) f(\alpha_1) f(\alpha_2) \ldots f(\alpha_{n-1}),$$

где $z = r(\cos\varphi + i\sin\varphi)$ — любое комплексное число и

$$\alpha_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right), \qquad k = 0, 1, 2, \ldots, n-1,$$

— все значения корня $\sqrt[n]{z}$ (см. [20], задача 494). Выражения для различных циркулянтов частного вида даны в [20], задачи 375, 376, 484—492, 495.

13. Взаимные (или присоединенные) и ассоциированные определители, их миноры и формула Сильвестра. Твердо установившихся обозначений и терминологии здесь не существует.

Взаимным (или *присоединенным*) определителем для определителя $D = |a_{ij}|_1^n$ порядка $n \geqslant 2$ называется определитель $\hat{D} = |\hat{a}_{ij}|_1^n$, где $\hat{a}_{ij} = A_{ij}$ $(i, j = 1, 2, \ldots, n)$, т. е. определитель, полученный из данного определителя D путем замены его элементов на их алгебраические дополнения (с сохранением их расположения). Значения данного и взаимного определителей связаны равенством $\hat{D} = D^{n-1}$, где n — порядок D и $\hat{D}$ (см. [20], задача 506). В частности, если $D = 0$, то и $\hat{D} = 0$, и обратно.

Миноры данного и взаимного определителей связаны так: пусть M — минор порядка p определителя D порядка n, A — алгебраическое дополнение минора M, $\hat{M}$ — минор взаимного определителя $\hat{D}$, соответствующий минору M, т. е. стоящий в тех же строках и столбцах, что и M. Тогда $\hat{M} = D^{p-1}A$ (см. [3], стр. 37).

Обозначая через $\hat{A}$ алгебраическое дополнение минора $\hat{M}$ и применяя последнее равенство с заменой p на $n-p$, получим $\hat{A} = D^{n-p-1}M$.

Если M — минор второго порядка, стоящий в i-й и j-й строках и k-м и l-м столбцах определителя D, и M' — дополнительный минор, полученный из определителя D вычеркиванием указанных строк и столбцов, то

$$\begin{vmatrix} A_{ik} & A_{il} \\ A_{jk} & A_{jl} \end{vmatrix} = (-1)^{i+j+k+l} DM',$$

где A_{st} — алгебраическое дополнение элемента a_{st} (см. [20], задача 508 или [3], стр. 39).

Из последнего равенства вытекает, что если данный определитель D равен нулю, то все миноры второго порядка взаимного определителя $\hat{D}$ равны нулю, т. е. любые две строки (или два столбца) определителя $\hat{D}$ пропорциональны.

Обозначая через $\hat{M}_{ij}$ минор и через $\hat{A}_{ij}$ алгебраическое дополнение элемента $\hat{a}_{ij}$ взаимного определителя $\hat{D}$, получим

$$\hat{M}_{ij} = (-1)^{i+j} D^{n-2} a_{ij}$$

или

$$\hat{A}_{ij} = D^{n-2} a_{ij}$$

(см. [3], стр. 39).

Это равенство позволяет по элементам взаимного определителя $\hat{D}$ (или по минорам определителя D) найти элементы самого определителя D (см. [20], задача 512).

Ассоциированным определителем для определителя $D = |a_{ij}|_1^n$ $(n \geqslant 2)$ называется определитель $\tilde{D} = |\tilde{a}_{ij}|_1^n$, где $\tilde{a}_{ij} = M_{ij}$ $(i, j = 1, 2, \ldots, n)$, т. е. определитель, полученный из данного определителя D путем замены его элементов на их миноры. Его свойства похожи на свойства взаимного определителя $\hat{D}$ и могут быть выведены из последних. Именно:

$\tilde{D}=\hat{D}=D^{n-1}$. Если M — минор порядка p определителя D, M' — дополнительный минор для M и $\tilde{M}$ — минор ассоциированного с D определителя $\tilde{D}$, стоящий в тех же строках и столбцах, что и минор M, то $\tilde{M}=D^{p-1}M'$ (см. [7], стр. 19).

Понятия взаимного и ассоциированного определителей можно обобщить, беря определители, составленные из миноров любого порядка k определителя D или из их алгебраических дополнений.

Пусть D — определитель порядка $n\geqslant 2$ и k равно одному из чисел 1, 2, ..., n. Пусть

$$N=C_n^k=\frac{n!}{k!\,(n-k)!}.$$

Все сочетания по k чисел из 1, 2, ..., n занумеруем в произвольном (но в дальнейшем фиксированном) порядке: s_1, s_2, ..., s_N. Пусть μ_{ij} — минор порядка k, стоящий в строках определителя D, номера которых составляют сочетание s_i, и в столбцах, номера которых составляют сочетание s_j, α_{ij} — алгебраическое дополнение минора μ_{ij} в D. Тогда *k-м ассоциированным* определителем для определителя D называется определитель $\tilde{D}_k=|\mu_{ij}|_1^N$.

k-м взаимным (или *k-м присоединенным*) для определителя D называется определитель $\hat{D}_k=|\alpha_{ij}|_1^N$. $\tilde{D}_k$ и $\hat{D}_k$ имеют порядок $N=C_n^k$. Эти определители обладают следующими свойствами (см. [20], задача 551):

1) значения определителей $\tilde{D}_k$ и $\hat{D}_k$ не зависят от выбора нумерации сочетаний из чисел 1, 2, ..., n по k;

$$2)\ \tilde{D}_k=\hat{D}_{n-k};$$

$$3)\ \tilde{D}_k\hat{D}_k=D^N=D^{C_n^k};$$

$$4)\ \tilde{D}_k=D^{C_{n-1}^{k-1}};$$

$$5)\ \hat{D}_k=D^{C_{n-1}^k}.$$

Теорема Сильвестра. Пусть $D=|a_{ij}|_1^n$ — определитель порядка $n\geqslant 2$, $M=|a_{ij}|_1^p$ — его минор порядка $p\,(1\leqslant p<n)$, стоящий в левом верхнем углу, b_{ij} — минор, полученный путем приписывания к минору M снизу i-й строки и справа j-го столбца определителя D $(i,\ j=p+1,$

$p+2, \ldots, n$), $\Delta = |b_{ij}|_{p+1}^{n}$ — определитель порядка $n-p$ из всех элементов b_{ij}. Тогда имеет место *формула Сильвестра*

$$\Delta = M^{n-p-1} D$$

(см. [7], стр. 21).

14. Симметрические, кососимметрические и псевдосимметрические определители. Определитель $D = |a_{ij}|_1^n$ называется *симметрическим*, *кососимметрическим* или *псевдосимметрическим*, если его элементы связаны соответственно условиями:

$$\begin{aligned}
&1)\ a_{ij} = a_{ji}, && i, j = 1, 2, \ldots, n;\\
&2)\ a_{ij} = -a_{ji}, && i, j = 1, 2, \ldots, n;\\
&3)\ a_{ij} = -a_{ji}, && i, j = 1, 2, \ldots, n;\ i \neq j.
\end{aligned}$$

Главные миноры (т. е. миноры, лежащие симметрично относительно главной диагонали, иными словами, лежащие в строках и столбцах с одинаковыми номерами) определителя одного из трех указанных типов сами имеют тот же тип.

В симметрическом определителе элементы a_{ij} и a_{ji}, лежащие симметрично относительно главной диагонали, имеют равные миноры и равные алгебраические дополнения:

$$M_{ij} = M_{ji}, \quad A_{ij} = A_{ji}.$$

Иными словами, для симметрического определителя D взаимный определитель $\hat{D}$ (стр. 38) и ассоциированный определитель $\tilde{D}$ (стр. 39) также будут симметрическими. То же верно и для k-го взаимного и k-го ассоциированного определителей $\hat{D}_k$ и $\tilde{D}_k$ (стр. 40).

Если D — кососимметрический определитель порядка n, то взаимный определитель $\hat{D}$ и ассоциированный определитель $\tilde{D}$ будут симметрическими, если n нечетно, и кососимметрическими, если n четно, и, вообще, k-й взаимный определитель $\hat{D}_k$ будет симметрическим, если разность $n-k$ четна, и кососимметрическим, если нечетна, а k-й ассоциированный определитель $\tilde{D}_k$ будет симметрическим, если k четно, и кососимметрическим, если k нечетно.

Кососимметрический определитель нечетного порядка равен нулю (см. [11], стр. 43).

Кососимметрический определитель четного порядка n равен квадрату некоторого однородного многочлена степени $n/2$ от его элементов (см. [30], стр. 45 или [20], задача 543). Ввиду условий $a_{ij}=-a_{ji}$ этот многочлен можно выразить лишь через элементы определителя, лежащие выше главной диагонали. Тогда он будет определен однозначно с точностью до знака и называется *пфаффовым агрегатом* или *пфаффианом*. О точном определении пфаффова агрегата см. [10], стр. 147—156 или [20], задача 545.

Пусть $D_n=|a_{ij}|_1^n$ — кососимметрический определитель четного порядка n и p_n — соответствующий ему пфаффов агрегат, так что $D_n=p_n^2$. Обозначим через p_{in} пфаффов агрегат кососимметрического определителя D_{in}, полученного из определителя D_n вычеркиванием n-й и i-й строк, а также n-го и i-го столбцов. Тогда имеет место рекуррентная формула:

$$p_2=a_{12},\quad p_n=\sum_{i=1}^{n-1}(-1)^{i-1}p_{in}a_{in}\quad (n\geqslant 4)$$

(см. [20], задача 546).

Пример 1. $p_2=a_{12}$. Поэтому $p_{14}=a_{23}$, $p_{24}=a_{13}$; $p_{34}=a_{12}$ и по рекуррентной формуле

$$p_4=p_{14}a_{14}-p_{24}a_{24}+p_{34}a_{34}=a_{23}a_{14}-a_{13}a_{24}+a_{12}a_{34}.$$

Далее,

$$\begin{aligned}p_6&=p_{16}a_{16}-p_{26}a_{26}+p_{36}a_{36}-p_{46}a_{46}+p_{56}a_{56}=\\&=a_{34}a_{25}a_{16}-a_{24}a_{35}a_{16}+a_{23}a_{45}a_{16}-a_{34}a_{15}a_{26}+\\&+a_{14}a_{35}a_{26}-a_{13}a_{45}a_{26}+a_{24}a_{15}a_{36}-a_{14}a_{25}a_{36}+\\&+a_{12}a_{45}a_{36}-a_{23}a_{15}a_{46}+a_{13}a_{25}a_{46}-a_{12}a_{35}a_{46}+a_{23}a_{14}a_{56}-\\&-a_{13}a_{24}a_{56}+a_{12}a_{34}a_{56}.\end{aligned}$$

Из рекуррентной формулы ясно, что число слагаемых пфаффова агрегата p_n равно произведению $1\cdot 3\cdot 5\ldots(n-1)$ (n — число четное).

Псевдосимметрический определитель, у которого все элементы главной диагонали равны между собой, равен сумме квадратов некоторых многочленов от его элементов, если он четного порядка, и сумме указанного вида, умноженной на элемент главной диагонали, если он нечетного порядка (см. [30], стр. 46).

Пример 2.

$$\begin{vmatrix} x & r & q \\ -r & x & p \\ -q & -p & x \end{vmatrix} = x(x^2+p^2+q^2+r^2);$$

$$\begin{vmatrix} x & a & b & c \\ -a & x & r & q \\ -b & -r & x & p \\ -c & -q & -p & x \end{vmatrix} = x^4+(a^2+b^2+c^2+p^2+q^2+r^2)x^2+ \\ +(ap-bq+cr)^2.$$

15. Кронекеровское произведение определителей. О кронекеровском произведении матриц см. стр. 120—122.

Кронекеровским или *прямым произведением* двух определителей $|A|=|a_{ij}|_1^n$ порядка n и $|B|=|b_{ij}|_1^p$ порядка p называется определитель порядка np, матрица которого строится по матрицам A и B данных определителей одним из следующих двух способов: 1) эта матрица состоит из p^2 клеток, расположенных в p строк и p столбцов, причем клетка в i-й клеточной строке и j-м клеточном столбце получается из матрицы A умножением всех ее элементов на число b_{ij}, или 2) эта матрица состоит из n^2 клеток, расположенных в n строк и n столбцов, причем клетка в i-й клеточной строке и j-м клеточном столбце получается из матрицы B умножением всех ее элементов на число a_{ij}. В первом случае произведение называется *левым*, а во втором — *правым*. В обоих случаях кронекеровское произведение имеет одно и то же значение $|A|^p|B|^n$.

Пример. Левое произведение определителей

$$|A|=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \text{ и } |B|=\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

имеет вид

$$\begin{vmatrix} a_{11}b_{11} & a_{12}b_{11} & a_{13}b_{11} & a_{11}b_{12} & a_{12}b_{12} & a_{13}b_{12} \\ a_{21}b_{11} & a_{22}b_{11} & a_{23}b_{11} & a_{21}b_{12} & a_{22}b_{12} & a_{23}b_{12} \\ a_{31}b_{11} & a_{32}b_{11} & a_{33}b_{11} & a_{31}b_{12} & a_{32}b_{12} & a_{33}b_{12} \\ a_{11}b_{21} & a_{12}b_{21} & a_{13}b_{21} & a_{11}b_{22} & a_{12}b_{22} & a_{13}b_{22} \\ a_{21}b_{21} & a_{22}b_{21} & a_{23}b_{21} & a_{21}b_{22} & a_{22}b_{22} & a_{23}b_{22} \\ a_{31}b_{21} & a_{32}b_{21} & a_{33}b_{21} & a_{31}b_{22} & a_{32}b_{22} & a_{33}b_{22} \end{vmatrix} = |A|^2|B|^3.$$

16. Неравенство Адамара и другие неравенства, связанные с определителями. Пусть $A=(a_{ij})_1^n$ — квадратная действительная матрица. Для ее определителя имеет место следующее *неравенство Адамара*:

$$|A|^2 \leqslant \prod_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2.$$

Если $|A| \neq 0$, то неравенство Адамара обращается в равенство тогда и только тогда, когда строки матрицы A попарно ортогональны, т. е. когда выполняются равенства

$$\sum_{k=1}^{n} a_{ik}a_{jk} = 0; \quad i, j = 1, 2, \ldots, n; \; i \neq j$$

(см. [7], стр. 51).

Неравенство Адамара имеет следующий геометрический смысл. Если элементы строк матрицы A считать координатами векторов n-мерного евклидова пространства ([11], стр. 211—212) в ортонормированной базе, то определитель $|A|$ с точностью до знака равен объему n-мерного параллелепипеда, построенного на этих векторах, как на ребрах. Неравенство Адамара означает, что объем параллелепипеда не превосходит произведения длин его ребер, выходящих из одной вершины, и равен этому произведению тогда и только тогда, когда ребра попарно ортогональны.

Назовем квадратную действительную матрицу A *положительно определенной*, если все ее миноры, стоящие в левом верхнем углу, положительны, т. е. если *квадратичная форма* с матрицей A положительно определенна. Для такой матрицы неравенство Адамара можно записать в другом виде.

Вторая форма неравенства Адамара. Если $A = (a_{ij})_1^n$ — положительно определенная матрица, то

$$|A| \leqslant a_{11}a_{22} \ldots a_{nn},$$

причем знак равенства имеет место тогда и только тогда, когда матрица A диагональна, т. е. когда $a_{ij} = 0$; $i \neq j$; $i, j = 1, 2, \ldots, n$ ([7], стр. 48).

Определитель, составленный из элементов матрицы A, стоящих на пересечении k строк с номерами $i_1, i_2, \ldots, i_k$ и k столбцов с номерами $j_1, j_2, \ldots, j_k$, называется *минором k-го порядка* матрицы A и обозначается через

$$A\begin{pmatrix} i_1 i_2 \ldots i_k \\ j_1 j_2 \ldots j_k \end{pmatrix}.$$

Обобщенное неравенство Адамара. Если $A = (a_{ij})_1^n$ — положительно определенная матрица и p — натуральное число $(1 \leqslant p \leqslant n-1)$, то

$$|A| \leqslant A\begin{pmatrix}1 & 2 \ldots & p\\ 1 & 2 \ldots & p\end{pmatrix} A\begin{pmatrix}p+1 \ldots n\\ p+1 \ldots n\end{pmatrix}, \qquad (1.1)$$

причем знак равенства имеет место тогда и только тогда, когда $a_{ij} = a_{ji} = 0$; $i = 1, 2, \ldots, p$; $j = p+1, \ldots, n$ (см. [7], стр. 50).

Если использовать умножение прямоугольных матриц (стр. 67—68) и через A' обозначить транспонированную матрицу A, то обобщенное неравенство Адамара можно записать в следующей матричной форме.

Пусть $A = (B, C)$ — квадратная вещественная матрица, где B — матрица из первых p, C — матрица из последних $n-p$ столбцов матрицы A $(1 \leqslant p \leqslant n-1)$. Тогда

$$|A'A| \leqslant |B'B| \cdot |C'C|.$$

Если $B'C = 0$, то имеет место знак равенства. Если $|A| \neq 0$, то, обратно, в случае знака равенства получаем $B'C = 0$ (роль матрицы A в (1.1) теперь выполняет матрица $A'A$). В этой форме неравенство допускает следующие обобщения.

Пусть $A = (B, C)$ — вещественная прямоугольная матрица, причем B и C — матрицы соответственно из нескольких первых столбцов и остальных столбцов матрицы A. Тогда

$$|A'A| \leqslant |B'B| |C'C|.$$

Если $B'C = 0$, то имеет место знак равенства. Если столбцы матрицы A (как векторы) линейно независимы ([34], стр. 56—71), то, обратно, в случае знака равенства получаем $B'C = 0$ ([20], задача 922 или [24], задача 518).

Аналогичное неравенство верно и для матрицы с комплексными элементами. Обозначим через A^* матрицу, полученную из матрицы A путем транспонирования и замены всех элементов числами, комплексно сопряженными. Пусть $A = (B, C)$ — прямоугольная комплексная матрица, где B и C — матрицы соответственно из первых и остальных столбцов матрицы A. Тогда определители

$$|A^*A|, \; |B^*B|, \; |C^*C|$$

будут неотрицательными действительными числами, причем

$$|A^*A| \leqslant |B^*B|\,|C^*C|.$$

Если $B^*C=0$, то имеет место знак равенства. Если столбцы матрицы A линейно независимы, то, обратно, из справедливости знака равенства следует $B^*C=0$ ([20], задача 925 или [24], задача 521).

Приведем еще две оценки величины определителя.

Если все элементы определителя $D=|a_{ij}|_1^n$ порядка n по модулю не превосходят положительное число M, то для модуля этого определителя выполняется оценка $|D| \leqslant M^n\sqrt{n^n}$, причем эта оценка является точной ([24], задачи 522, 524 или [20], задача 926).

Если все элементы определителя $D=|a_{ij}|_1^n$ действительны и удовлетворяют условиям $0 \leqslant a_{ij} \leqslant M$, то

$$|D| \leqslant \frac{M^n\sqrt{(n+1)^{n+1}}}{2^n}$$

(см. [24], задача 523).

§ 2. Системы линейных уравнений

1. Общие сведения. *Система линейных уравнений* имеет вид

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \cdots\cdots\cdots&\cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n&=b_s. \end{aligned}\right\} \tag{1.2}$$

Здесь $x_1, x_2, \ldots, x_n$ — неизвестные, а коэффициенты при неизвестных a_{ij} и свободные члены b_i — числа из данного числового поля K (см. введение).

Решением системы уравнений (1.2) называется упорядоченная система чисел

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \tag{1.3}$$

удовлетворяющая всем уравнениям данной системы, т. е. обращающая при замене неизвестных на соответствующие числа все уравнения в верные равенства. Числа (1.3) называются *значениями неизвестных* $x_1, x_2, \ldots, x_n$ в данном решении.

Система линейных уравнений называется *совместной*, если она имеет хотя бы одно решение, и *несовместной* (или *противоречивой*), если она не имеет решений. Совместная система называется *определенной*, если она имеет единственное решение, и *неопределенной*, если более одного решения. Неопределенная система линейных уравнений всегда имеет бесконечное множество решений.

Две системы линейных уравнений называются *эквивалентными* (или *равносильными*), если каждое решение первой системы является решением второй, и обратно. Любые несовместные системы линейных уравнений с одинаковым числом неизвестных считаются эквивалентными. Для того чтобы две совместные системы линейных уравнений с одинаковым числом неизвестных были эквивалентны, необходимо и достаточно, чтобы каждое уравнение первой системы было линейной комбинацией уравнений второй системы, и обратно.

Если каждое решение первой системы линейных уравнений является решением второй (но не обязательно наоборот), то вторая система называется *следствием* первой. Любая система линейных уравнений считается следствием несовместной системы. Если каждое уравнение второй системы является линейной комбинацией уравнений первой системы, то вторая система будет следствием первой. Если первая система совместна и вторая система является следствием первой, то каждое уравнение второй системы будет линейной комбинацией уравнений первой системы.

2. Метод Гаусса исключения неизвестных. *Элементарными преобразованиями* системы линейных уравнений называются преобразования следующих трех типов:

1) перестановка двух уравнений;

2) умножение обеих частей одного из уравнений на любое число, отличное от нуля;

3) прибавление к обеим частям одного уравнения соответствующих частей другого, умноженных на любое число.

Элементарные преобразования переводят данную систему уравнений в эквивалентную систему.

Системой ступенчатого вида (или *ступенчатой системой*) называется система вида

$$
\begin{aligned}
a_{1,k_1}x_{k_1} + a_{1,k_1+1}x_{k_1+1} + \ldots + a_{1n}x_n &= b_1,\\
a_{2,k_2}x_{k_2} + a_{2,k_2+1}x_{k_2+1} + \ldots + a_{2n}x_n &= b_2,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots &\\
a_{r,k_r}x_{k_r} + a_{r,k_r+1}x_{k_r+1} + \ldots + a_{rn}x_n &= b_r,\\
0 &= 1,\\
0 &= 0,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots &\\
0 &= 0,
\end{aligned}
$$

где $1 \leqslant k_1 < k_2 < \ldots < k_r \leqslant n$ и $a_{i,k_i} \neq 0$, $i = 1, 2, \ldots, r$. Здесь через $0 = 1$ и $0 = 0$ обозначены уравнения, в которых коэффициенты при всех неизвестных равны нулю. При этом уравнения вида $0 = 1$ и $0 = 0$ могут и отсутствовать, но уравнение вида $0 = 1$ входит не более одного раза.

Пример 1. Система

$$
\begin{aligned}
x_1 + 2x_2 \qquad + x_4 - x_5 + x_6 + 3x_7 &= 5,\\
x_2 - x_3 \qquad + x_5 \qquad - x_7 &= 2,\\
3x_4 + x_5 - x_6 + x_7 &= 1,\\
x_5 + 2x_6 + 3x_7 &= 0,\\
0 &= 1,\\
0 &= 0
\end{aligned}
$$

имеет ступенчатый вид.

Любую систему линейных уравнений при помощи конечного числа элементарных преобразований можно привести к ступенчатому виду. Этим решение любой системы сводится к решению системы ступенчатого вида.

Уравнения ступенчатой системы вида $0 = 0$ можно отбросить, так как это приводит к системе уравнений, эквивалентной прежней.

Ступенчатая система тогда и только тогда совместна, когда она не содержит уравнения вида $0 = 1$.

Неизвестные $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$, начинающие уравнения ступенчатой системы, называются главными, а остальные неизвестные — свободными.

Решение совместной системы ступенчатого вида находят так: отбрасывают уравнения вида $0 = 0$, затем из последнего уравнения находят выражение последнего главного неизвестного через свободные и подставляют это выражение во все предыдущие уравнения; в последнем из полученных уравнений снова находят выражение главного неизвестного через свободные и т. д. Полученные выражения главных неизвестных через свободные называются *общим решением* системы уравнений. Подставляя в общее решение вместо свободных неизвестных произвольные числа, найдем вполне определенные значения главных неизвестных.

Ступенчатая система, содержащая свободные неизвестные и не содержащая уравнений вида $0=1$, имеет бесконечное множество решений. Если же ступенчатая система не содержит свободных неизвестных и не содержит уравнений вида $0=1$, то она имеет единственное решение.

Ступенчатая система называется *треугольной*, если она не содержит уравнений, в которых коэффициенты при всех неизвестных равны нулю, и если число уравнений равно числу неизвестных. Система линейных уравнений тогда и только тогда имеет единственное решение, когда путем элементарных преобразований и отбрасывания уравнений вида $0=0$ она приводится к треугольному виду.

Практически удобнее приводить к ступенчатому виду не саму систему уравнений, а матрицу из коэффициентов при неизвестных и свободных членов, соединяя последовательно получающиеся матрицы знаком эквивалентности $\sim$.

Пример 2. Решить систему уравнений

$$\begin{aligned}
x_1+x_2\phantom{{}-3x_3}+2x_4+3x_5&=5,\\
2x_1+4x_2-x_3+5x_4+4x_5&=-1,\\
x_1+3x_2\phantom{{}-3x_3}+5x_4+2x_5&=-3,\\
3x_1+7x_2-3x_3+9x_4+2x_5&=-14,\\
2x_1+8x_2-4x_3+2x_4+7x_5&=-10.
\end{aligned}$$

Преобразуем матрицу

$$\left(\begin{array}{ccccc|c}
1 & 1 & 0 & 2 & 3 & 5\\
2 & 4 & -1 & 5 & 4 & -1\\
1 & 3 & 0 & 5 & 2 & -3\\
3 & 7 & -3 & 9 & 2 & -14\\
2 & 8 & -4 & 2 & 7 & -10
\end{array}\right)\sim$$

(первую строку, умноженную соответственно на 2, 1, 3, 2, вычитаем из остальных)

$$\sim\left(\begin{array}{ccccc|c}
1 & 1 & 0 & 2 & 3 & 5\\
0 & 2 & -1 & 1 & -2 & -11\\
0 & 2 & 0 & 3 & -1 & -8\\
0 & 4 & -3 & 3 & -7 & -29\\
0 & 6 & -4 & -2 & 1 & -20
\end{array}\right)\sim$$

(вторую строку, умноженную соответственно на 1, 2, 3, вычитаем из следующих)

$$\sim\left(\begin{array}{rrrrr|r} 1 & 1 & 0 & 2 & 3 & 5\\ 0 & 2 & -1 & 1 & -2 & -11\\ 0 & 0 & 1 & 2 & 1 & 3\\ 0 & 0 & -1 & 1 & -3 & -7\\ 0 & 0 & -1 & -5 & 7 & 13\end{array}\right)\sim$$

(третью строку прибавляем к следующим)

$$\sim\left(\begin{array}{rrrrr|r} 1 & 1 & 0 & 2 & 3 & 5\\ 0 & 2 & -1 & 1 & -2 & -11\\ 0 & 0 & 1 & 2 & 1 & 3\\ 0 & 0 & 0 & 3 & -2 & -4\\ 0 & 0 & 0 & -3 & 8 & 16\end{array}\right)\sim$$

(четвертую строку прибавляем к пятой)

$$\sim\left(\begin{array}{rrrrr|r} 1 & 1 & 0 & 2 & 3 & 5\\ 0 & 2 & -1 & 1 & -2 & -11\\ 0 & 0 & 1 & 2 & 1 & 3\\ 0 & 0 & 0 & 3 & -2 & -4\\ 0 & 0 & 0 & 0 & 6 & 12\end{array}\right).$$

Система уравнений привелась к треугольному виду

$$\begin{aligned} x_1 + x_2 \qquad + 2x_4 + 3x_5 &= 5,\\ 2x_2 - x_3 + x_4 - 2x_5 &= -11,\\ x_3 + 2x_4 + x_5 &= 3,\\ 3x_4 - 2x_5 &= -4,\\ 6x_5 &= 12.\end{aligned}$$

Она имеет единственное решение. Из последнего уравнения находим $x_5 = 2$; подставляем это значение x_5 в предыдущее уравнение и находим из него $x_4 = 0$; подставляем эти значения x_4 и x_5 в третье уравнение и находим $x_3 = 1$; так же находим $x_2 = -3$, $x_1 = 2$.

Пример 8. Решить систему уравнений

$$\begin{aligned} 2x_1 - x_2 - x_3 + 3x_4 &= 1,\\ 4x_1 - 2x_2 - x_3 + x_4 &= 5,\\ 6x_1 - 3x_2 - x_3 - x_4 &= 9,\\ 2x_1 - x_2 + 2x_3 - 12x_4 &= 10.\end{aligned}$$

Преобразуем матрицу

$$\left(\begin{array}{rrrr|r} 2 & -1 & -1 & 3 & 1\\ 4 & -2 & -1 & 1 & 5\\ 6 & -3 & -1 & -1 & 9\\ 2 & -1 & 2 & -12 & 10\end{array}\right)\sim$$

(первую строку, умноженную соответственно на 2, 3, 1, вычитаем из следующих)

$$\sim\left(\begin{array}{cccc|c} 2 & -1 & -1 & 3 & 1 \\ 0 & 0 & 1 & -5 & 3 \\ 0 & 0 & 2 & -10 & 6 \\ 0 & 0 & 3 & -15 & 9 \end{array}\right)\sim$$

(вторую строку, умноженную соответственно на 2, 3, вычитаем из следующих)

$$\sim\left(\begin{array}{cccc|c} 2 & -1 & -1 & 3 & 1 \\ 0 & 0 & 1 & -5 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Система привелась к ступенчатой, которая после отбрасывания двух уравнений вида $0=0$ превращается в следующую:

$$\begin{aligned} 2x_1 - x_2 - x_3 + 3x_4 &= 1, \\ x_3 - 5x_4 &= 3. \end{aligned}$$

Неизвестные x_1 и x_3 — главные, а x_2 и x_4 — свободные. Система неопределенна. Из второго уравнения найдем выражение x_3 через x_4; подставив его в первое уравнение, найдем выражение x_1 через x_2 и x_4. Общее решение имеет вид

$$\begin{aligned} x_1 &= \frac{1}{2}x_2 + x_4 + 2, \\ x_3 &= 5x_4 + 3. \end{aligned}$$

Полагая в общем решении, например, $x_2 = 2$, $x_4 = -1$, найдем одно из решений:

$$x_1 = 2, \quad x_2 = 2, \quad x_3 = -2, \quad x_4 = -1,$$

которое в отличие от общего решения называется *частным решением* данной системы уравнений.

Пример 4. Решить систему уравнений

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 + x_4 &= 3, \\ 3x_1 + 4x_2 + 2x_3 + 3x_4 &= -2, \\ x_1 + 2x_2 + 8x_3 - x_4 &= 8, \\ 7x_1 + 9x_2 + x_3 + 8x_4 &= 0. \end{aligned}$$

Преобразуем матрицу

$$\left(\begin{array}{cccc|c} 2 & 3 & 5 & 1 & 3 \\ 3 & 4 & 2 & 3 & -2 \\ 1 & 2 & 8 & -1 & 8 \\ 7 & 9 & 1 & 8 & 0 \end{array}\right)\sim$$

(меняем местами первую и третью строки и новую первую строку, умноженную соответственно на 3, 2, 7, вычитаем из остальных)

$$\sim\left(\begin{array}{rrrr|r} 1 & 2 & 8 & -1 & 8 \\ 0 & -2 & -22 & 6 & -26 \\ 0 & -1 & -11 & 3 & -13 \\ 0 & -5 & -55 & 15 & -56 \end{array}\right)\sim$$

(делим вторую строку на -2 и, умножив ее соответственно на 1 и 5, прибавляем к следующим)

$$\sim\left(\begin{array}{rrrr|r} 1 & 2 & 8 & -1 & 8 \\ 0 & 1 & 11 & -3 & 13 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 \end{array}\right)\sim$$

(четвертую строку делим на 9 и переставляем ее с третьей)

$$\sim\left(\begin{array}{rrrr|r} 1 & 2 & 8 & -1 & 8 \\ 0 & 1 & 11 & -3 & 13 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Система привелась к ступенчатому виду, содержащему уравнение вида $0=1$, и потому несовместна.

3. Правило Крамера. Если дана система уравнений $a_{i1}x_1+\ldots+a_{in}x_n=b_i$, $i=1, \ldots, n$ (т. е. число уравнений равно числу неизвестных) и определитель этой системы $d=|a_{ij}|_1^n\neq 0$, то система имеет единственное решение, которое находится по *формулам Крамера* $x_j=\frac{d_j}{d}$, $j=1, \ldots, n$, где определитель d_j получен из определителя системы d заменой j-го столбца на столбец из свободных членов.

Пример 1. Решить систему уравнений

$$\begin{aligned} x_1-2x_2&=1,\\ 3x_1-4x_2&=7. \end{aligned}$$

Определитель системы

$$d=\begin{vmatrix} 1 & -2 \\ 3 & -4 \end{vmatrix}=2\neq 0.$$

Поэтому можно применить правило Крамера. Находим

$$d_1 = \begin{vmatrix} 1 & -2 \\ 7 & -4 \end{vmatrix} = 10, \quad d_2 = \begin{vmatrix} 1 & 1 \\ 3 & 7 \end{vmatrix} = 4.$$

Отсюда

$$x_1 = \frac{d_1}{d} = 5, \quad x_2 = \frac{d_2}{d} = 2.$$

В случае, когда определитель d из коэффициентов при неизвестных равен нулю, правило Крамера неприменимо. Из общей теории линейных уравнений (п. 6) следует, что такая система будет либо неопределенной, либо противоречивой.

Пример 2. Решить систему уравнений

$$\begin{aligned} x_1 + x_2 + x_3 &= 1, \\ x_1 + 2x_2 + 2x_3 &= 2, \\ 2x_1 + 3x_2 + 3x_3 &= 3. \end{aligned}$$

Здесь $d = d_1 = d_2 = d_3 = 0$, так как третья строка равна сумме первых двух. Система приводится к ступенчатому виду

$$\begin{aligned} x_1 + x_2 + x_3 &= 1, \\ x_2 + x_3 &= 1, \\ 0 &= 0 \end{aligned}$$

и неопределенна.

Пример 3. Решить систему уравнений

$$\begin{aligned} x_1 + x_2 + x_3 &= 1, \\ 2x_1 + 2x_2 + 2x_3 &= 1, \\ 3x_1 + 3x_2 + 3x_3 &= 2. \end{aligned}$$

Здесь снова $d = d_1 = d_2 = d_3 = 0$, так как третья строка равна сумме первых двух. Система приводится к ступенчатому виду

$$\begin{aligned} x_1 + x_2 + x_3 &= 1, \\ 0 &= 1, \\ 0 &= 0 \end{aligned}$$

и противоречива.

Если определитель d из коэффициентов при неизвестных равен нулю, а хотя бы один из определителей, стоящих в числителях формул Крамера, отличен от нуля, то система несовместна.

4. Ранг матрицы. Рассмотрим прямоугольную матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & \dots & a_{sn} \end{pmatrix}$$

из s строк и n столбцов (см. введение).

Рангом системы строк (соответственно *столбцов*) матрицы A называется наибольшее число линейно независимых (§ 1, п, 3) среди них.

Оказывается, что ранг системы строк матрицы равен рангу системы ее столбцов. Это число называется *рангом матрицы.*

При транспонировании матрицы (стр. 18) ранг ее не изменяется.

Элементарными преобразованиями матрицы A называются следующие преобразования: 1) перестановка двух любых строк, 2) умножение строки на число c, отличное от нуля, 3) прибавление к одной строке другой строки, умноженной на любое число, и аналогичные преобразования столбцов.

Две матрицы называются *эквивалентными,* если от одной из них к другой можно перейти путем конечного числа элементарных преобразований.

При элементарных преобразованиях матрицы ее ранг не изменяется, иными словами, ранги эквивалентных матриц равны.

Ступенчатой матрицей называется матрица, обладающая тем свойством, что если в какой-либо из ее строк первый отличный от нуля элемент стоит на k-м месте, то во всех следующих строках на первых k местах стоят нули, например:

$$\begin{pmatrix} 2 & 1 & 3 & 2 & 5 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Любая матрица при помощи элементарных преобразований только строк приводится к ступенчатой. Ранг ступенчатой матрицы равен числу ее ненулевых строк.

Канонической матрицей называется матрица, у которой в начале главной диагонали стоят подряд несколько единиц (число которых может равняться нулю), а все остальные элементы равны нулю, например:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Любую матрицу при помощи элементарных преобразований (как строк, так и столбцов) можно привести к канонической.

Ранг канонической матрицы равен числу единиц на ее главной диагонали.

Простейший метод вычисления ранга матрицы с числовыми элементами состоит в приведении ее элементарными преобразованиями к ступенчатой или канонической матрице.

Пример 1. Найти ранг матрицы

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

При помощи элементарных преобразований приводим эту матрицу к ступенчатому виду, соединяя соседние матрицы знаком эквивалентности $\sim$. Из второй строки вычтем первую и переставим эти строки:

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

(из второй и третьей строк вычтем первую, умноженную соответственно на 2 и 5)

$$\sim \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 1 & 9 & -7 & 0 \end{pmatrix} \sim$$

(из третьей строки вычтем вторую)

$$\sim \begin{pmatrix} 1 & 1 & -2 & 2 & -1 \\ 0 & 1 & 9 & -7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ранг матрицы равен двум. Последнюю матрицу легко привести к канонической. Вычитая первый столбец, умноженный на подходящие числа, из следующих, обратим в нуль все элементы первой строки, кроме первого, причем элементы остальных строк не изменятся. Затем, вычитая второй столбец, умноженный на подходящие числа, из следующих, обратим в нуль все элементы второй строки, кроме второго, и получим каноническую матрицу

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Другой метод вычисления ранга матрицы связан с вычислением определителей.

Теорема о ранге матрицы. Если матрица имеет минор порядка r (см. стр. 44), отличный от нуля, для которого все содержащие его миноры порядка $r+1$ (окаймляющие миноры) равны нулю, то ранг этой матрицы равен r.

Из этой теоремы следует, что ранг матрицы (не все элементы которой равны нулю) равен наивысшему порядку отличных от нуля миноров этой матрицы.

Вычисление ранга матрицы с помощью метода окаймления надо вести от низших порядков к высшим. Сначала ищем минор первого порядка (т. е. элемент матрицы) или сразу второго порядка, отличный от нуля. Затем вычисляем окаймляющие его миноры следующего порядка, пока не найдем среди них отличного от нуля. Переходим к вычислению миноров, окаймляющих этот последний, и т. д., пока не найдем минор порядка r, отличный от нуля, для которого либо все окаймляющие его миноры порядка $r+1$ равны нулю, либо таких миноров вообще не будет (если матрица содержит r строк или r столбцов). Тогда ранг матрицы равен r.

Пример 2. Найти ранг матрицы

$$\begin{pmatrix} 1 & -2 & 3 & -2 & 4 \\ 2 & -4 & 5 & 1 & 7 \\ 1 & -2 & 1 & 8 & 2 \end{pmatrix}.$$

Минор первого порядка в левом верхнем углу равен $1 \neq 0$. Окаймляющий его минор второго порядка равен нулю:

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

Добавляя вторую строку и третий столбец, получим окаймляющий минор

$$\begin{vmatrix} 1 & 3 \\ 2 & 5 \end{vmatrix} = -1 \neq 0.$$

Вычисляем окаймляющие его миноры третьего порядка. Второй столбец можно не добавлять, так как он пропорционален первому, и получится минор, равный нулю. Добавляя третью строку и либо четвертый, либо пятый столбец, получим миноры

$$\begin{vmatrix} 1 & 3 & -2 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} 1 & 3 & 4 \\ 2 & 5 & 7 \\ 1 & 1 & 2 \end{vmatrix},$$

равные нулю; значит, ранг матрицы равен двум.

Ранг суммы и произведения двух матриц. Сложение и умножение матриц определены в гл. II, § 1, п. 1.

Ранг суммы двух (или нескольких) матриц не больше суммы их рангов ([20], задача 626). Любую матрицу ранга r можно представить в виде суммы r матриц ранга 1, но нельзя представить в виде суммы менее чем r таких матриц ([20], задача 627).

Любую матрицу C ранга r можно представить в виде произведения $C=AB$, где A состоит из r линейно независимых столбцов и B — из r линейно независимых строк ([20], задача 939).

Если обозначим ранг матрицы A через r_A, то для ранга произведения AB двух квадратных матриц A и B порядка n имеет место

Неравенство Сильвестра:

$$r_A + r_B - n \leqslant r_{AB} \leqslant \min(r_A, r_B)$$

[(6), стр. 61 или [20], задача 931).

Такое же неравенство остается верным и для произведения прямоугольных матриц при условии, что n обозначает число столбцов матрицы A и число строк матрицы B.

Ранг симметрической и кососимметрической матриц. Определения этих матриц даны на стр. 129.

Главным минором матрицы называется любой минор, стоящий на пересечении строк и столбцов с одинаковыми номерами, т. е. расположенный симметрично относительно главной диагонали матрицы.

Ранг симметрической или кососимметрической матрицы равен наивысшему порядку отличных от нуля главных миноров. Точнее, если условимся минор «нулевого порядка» считать равным 1, то: если в симметрической (или кососимметрической) матрице существует главный минор D порядка r, отличный от нуля, а все содержащие его главные миноры порядков $r+1$ и $r+2$ (или только порядка $r+2$ в случае кососимметрической матрицы) равны нулю, то ранг матрицы равен r ([3], стр. 59—60 или [20], задачи 631, 633).

Если обозначим через D_k минор k-го порядка, стоящий в левом верхнем углу матрицы (по определению $D_0=1$), то путем некоторой перестановки строк и такой же перестановки столбцов симметрической матрицы можно получить такую симметрическую матрицу, в которой в ряду миноров D_0, D_1, D_2, ..., D_r не будет двух нулей подряд и $D_r \neq 0$, $D_k=0$ для всех $k>r$.

В случае кососимметрической матрицы ранга r аналогичным путем можно получить условия: миноры D_0, D_2, ..., D_r отличны от нуля, а все остальные миноры D_k равны нулю. Ранг r кососимметрической матрицы — всегда число четное.

Пример 3. Найти ранг матрицы

$$\begin{pmatrix} 0 & 4 & 10 & 1 \\ 4 & 8 & 18 & 7 \\ 10 & 18 & 40 & 17 \\ 1 & 7 & 17 & 3 \end{pmatrix}.$$

Так как матрица симметрическая, то достаточно знать лишь главные миноры

$$D_0=1, \quad D_1=8,$$

$$D_2=\begin{vmatrix} 0 & 4 \\ 4 & 8 \end{vmatrix}=-16, \qquad D_3=\begin{vmatrix} 0 & 4 & 10 \\ 4 & 8 & 18 \\ 10 & 18 & 40 \end{vmatrix}=0.$$

Другой главный минор, содержащий D_2, будет

$$D_3'=\begin{vmatrix} 0 & 4 & 1 \\ 4 & 8 & 7 \\ 1 & 7 & 3 \end{vmatrix}=0.$$

Наконец,

$$D_4=\begin{vmatrix} 0 & 4 & 10 & 1 \\ 4 & 8 & 18 & 7 \\ 10 & 18 & 40 & 17 \\ 1 & 7 & 17 & 3 \end{vmatrix}=$$

(из второй и третьей строк вычитаем четвертую, умноженную соответственно на 4 и 10)

$$=\begin{vmatrix} 0 & 4 & 10 & 1 \\ 0 & -20 & -50 & -5 \\ 0 & -52 & -130 & -13 \\ 1 & 7 & 17 & 3 \end{vmatrix}=0,$$

так как имеются пропорциональные строки. Значит, ранг матрицы равен двум.

5. Общая теория линейных уравнений. Укажем метод, позволяющий формулировать свойства системы линейных уравнений в общем виде, не производя предварительных преобразований данной системы.

Пусть дана система линейных уравнений

$$a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n=b_i, \quad i=1, 2, \ldots, s. \quad (1.4)$$

Матрицей системы уравнений (1.4) называется матрица из коэффициентов при неизвестных

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & \ldots & a_{sn} \end{pmatrix}.$$

Расширенной матрицей называется матрица, полученная из матрицы системы добавлением столбца свободных членов, т. е. матрица

$$B=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & \ldots & a_{sn} & b_s \end{pmatrix}.$$

Вопрос о совместности системы (1.4) решает следующий

Критерий Кронекера—Капелли. Для того чтобы система линейных уравнений (1.4) была совместна, необходимо и достаточно, чтобы ранг матрицы A системы был равен рангу расширенной матрицы B.

Пусть ранг матрицы системы равен $r \geqslant 1$. В этом случае любой отличный от нуля минор порядка r матрицы системы будем называть *главным минором системы*, а любой минор порядка $r+1$, полученный из главного минора добавлением столбца соответствующих свободных членов и одной из оставшихся строк (если она существует), — *характеристическим минором* для данного главного минора. Критерий совместности можно формулировать иначе.

Для того чтобы система линейных уравнений была совместна, необходимо и достаточно, чтобы существовал главный минор, для которого все характеристические миноры равны нулю или же вовсе не существуют (последнее имеет место при $r=s$).

Совместная система линейных уравнений имеет единственное решение, если ранг r матрицы системы равен числу неизвестных n, и бесконечно много решений, если $r<n$.

Если при этом $r>0$, то за главные неизвестные можно принять любые r неизвестных, из коэффициентов которых в некоторых r уравнениях системы можно составить минор порядка r, отличный от нуля. Остальные $n-r$ неизвестных являются свободными. Свободным неизвестным можно дать любые числовые значения. При заданных значениях свободных неизвестных значения главных неизвестных определяются однозначно. Иными словами, существует единственное решение совместной системы линейных уравнений, в котором свободные неизвестные имеют заданные числовые значения.

Правило для решения системы линейных уравнений. 1) Путем вычисления ранга матрицы системы и расширенной матрицы выясняют вопрос о совместности системы; если система совместна, то находят один из главных миноров $D\neq 0$ порядка r, равного рангу обеих матриц.

2) Берут r уравнений, в которых лежит минор D; остальные уравнения отбрасывают; r неизвестных, коэффициенты которых входят в главный минор D, объявляют главными и оставляют слева, а остальные $n-r$ неизвестных объявляют свободными и переносят в правую часть.

3) Методом исключения неизвестных (п. 2) или по правилу Крамера (п. 3) находят выражения главных неизвестных через свободные. Полученные равенства называются *общим решением*.

4) Придавая свободным неизвестным любые числовые значения, из общего решения находят соответствующие значения главных неизвестных и тем самым находят решение исходной системы уравнений, которое в отличие от общего решения можно назвать *частным решением*. Указанным путем можно получить любое решение. Для этого свободным неизвестным

в общем решении надо придать те значения, которые они имеют в данном решении.

Пример. Найти общее решение и одно частное решение для системы уравнений

$$\begin{aligned} 2x_1 - x_2 + x_3 + 2x_4 + 3x_5 &= 2, \\ 6x_1 - 3x_2 + 2x_3 + 4x_4 + 5x_5 &= 3, \\ 6x_1 - 3x_2 + 4x_3 + 8x_4 + 13x_5 &= 9, \\ 4x_1 - 2x_2 + x_3 + x_4 + 2x_5 &= 1. \end{aligned}$$

Выписываем расширенную матрицу, отделяя чертой столбец свободных членов:

$$\left(\begin{array}{rrrrr|r} 2 & -1 & 1 & 2 & 3 & 2 \\ 6 & -3 & 2 & 4 & 5 & 3 \\ 6 & -3 & 4 & 8 & 13 & 9 \\ 4 & -2 & 1 & 1 & 2 & 1 \end{array}\right).$$

Вычисляем ранг матрицы системы. Минор второго порядка в ее правом верхнем углу

$$\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -2 \neq 0.$$

Содержащий его минор третьего порядка

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 1 & 2 \end{vmatrix} = -1 \neq 0.$$

Добавляя к нему третью строку и второй столбец, получим минор четвертого порядка

$$\begin{vmatrix} -1 & 1 & 2 & 3 \\ -3 & 2 & 4 & 5 \\ -3 & 4 & 8 & 13 \\ -2 & 1 & 1 & 2 \end{vmatrix} = 0.$$

Добавлять первый столбец не надо, так как он пропорционален второму столбцу. Значит, ранг матрицы системы равен трем и D можно принять за главный минор. Для выяснения совместности достаточно вычислить единственный характеристический минор

$$\begin{vmatrix} 1 & 2 & 3 & 2 \\ 2 & 4 & 5 & 3 \\ 4 & 8 & 13 & 9 \\ 1 & 1 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 2 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{vmatrix} = 0.$$

Значит, ранг расширенной матрицы также равен трем и система совместна. Принимаем x_3, x_4, x_5 за главные, а x_1, x_2 — за свободные неизвестные и решаем методом исключения неизвестных систему уравнений

$$\begin{aligned} x_3+2x_4+3x_5&=2-2x_1+x_2,\\ 2x_3+4x_4+5x_5&=3-6x_1+3x_2,\\ x_3+x_4+2x_5&=1-4x_1+2x_2. \end{aligned}$$

Находим общее решение

$$\begin{aligned} x_3&=-4x_1+2x_2-1,\\ x_4&=0,\\ x_5&=2x_1-x_2+1. \end{aligned}$$

Полагая, например, $x_1=1$, $x_2=2$, найдем

$$x_3=-1;\ x_4=0;\ x_5=1.$$

Для того чтобы система линейных уравнений (1.4) была совместна при фиксированных значениях коэффициентов a_{ij} при неизвестных и любых значениях свободных членов b_i, необходимо и достаточно, чтобы ранг r матрицы A системы был равен числу s уравнений системы. При этом система имеет единственное решение, если число уравнений равно числу неизвестных, $s=n$, и бесконечно много решений, если $s<n$. В частности, для того чтобы система линейных уравнений (1.4) с фиксированными значениями коэффициентов a_{ij} при неизвестных имела единственное решение при любых значениях свободных членов b_i, необходимо и достаточно, чтобы данная система была *системой Крамера*, т. е. чтобы число уравнений было равно числу неизвестных и определитель системы был отличен от нуля.

6. Системы однородных линейных уравнений. Фундаментальные системы решений. Линейное уравнение называется *однородным*, если его свободный член равен нулю.

Система однородных линейных уравнений имеет вид

$$a_{i1}x_1+\ldots+a_{in}x_n=0,\quad i=1,\ 2,\ \ldots,\ s. \qquad (1.5)$$

Всякая система однородных линейных уравнений имеет *нулевое решение* $x_1=x_2=\ldots=x_n=0$ и, значит, совместна.

Для того чтобы система однородных линейных уравнений имела ненулевое решение, необходимо и достаточно, чтобы ранг r этой системы (т. е. ранг матрицы из коэффициентов при неизвестных) был меньше числа неизвестных n. Отсюда следует, в частности, что любая система однородных линейных уравнений, в которой число уравнений меньше числа неизвестных, $s<n$, имеет ненулевое решение.

Для того чтобы система однородных линейных уравнений, в которой число уравнений равно числу неизвестных, $s=n$, имела ненулевое решение, необходимо и достаточно, чтобы ее определитель был равен нулю.

Любое решение

$$x_1=c_1,\quad x_2=c_2,\ \ldots,\ x_n=c_n$$

системы линейных уравнений с n неизвестными можно рассматривать как строку (или столбец) $(c_1, c_2, \ldots, c_n)$. Поэтому имеют смысл такие понятия, как сумма двух решений, произведение решения на число, линейная комбинация решений, линейная зависимость или независимость системы решений (§ 1, п. 3).

Свойства решений системы однородных линейных уравнений: 1) сумма двух решений есть решение, 2) произведение решения на любое число есть решение.

Из этих свойств следует: 3) линейная комбинация решений есть решение. В частности, если существует хоть одно ненулевое решение, то из него умножением на произвольные числа можно получить бесконечно много решений.

Фундаментальной (или *основной*) системой решений для системы однородных линейных уравнений (1.5) называется линейно независимая система решений, через которую линейно выражается любое решение системы (1.5).

Если ранг r системы уравнений (1.5) равен числу n неизвестных, то эта система не имеет фундаментальной системы решений, так как единственным решением будет нулевое решение, составляющее линейно зависимую систему. Если $r<n$, то система (1.5) имеет бесконечно много фундаментальных систем решений, причем каждая из них состоит из $n-r$ решений и любые $n-r$ линейно независимых решений составляют фундаментальную систему.

Правило для построения фундаментальной системы решений. Берут любой определитель M порядка $n-r$, отличный от нуля (например, определитель, у которого элементы главной диагонали равны единице, а остальные — нулю). Свободным неизвестным придают поочередно значения, равные элементам первой, второй и т. д. строк определителя M, и каждый раз из общего решения находят соответствующие значения главных неизвестных. Полученные $n-r$ решений составляют фундаментальную систему. Меняя произвольно исходный определитель M, можно получить всевозможные фундаментальные системы решений

Пример 1. Найти общее решение и фундаментальную систему решений для системы уравнений

$$\begin{aligned} 3x_1+4x_2+x_3+2x_4+3x_5&=0,\\ 5x_1+7x_2+x_3+3x_4+4x_5&=0,\\ 4x_1+5x_2+2x_3+x_4+5x_5&=0,\\ 7x_1+10x_2+x_3+6x_4+5x_5&=0. \end{aligned}$$

Данная система приводится к ступенчатому виду (п. 2):

$$\begin{aligned} x_1+x_2+x_3-x_4+2x_5&=0,\\ x_2-2x_3+5x_4-3x_5&=0,\\ x_4&=0. \end{aligned}$$

Главными неизвестными считаем x_1, x_2 и x_4, а свободными — x_3 и x_5. Из второго уравнения находим

$$x_2=2x_3+3x_5.$$

Подставляя это выражение x_2 в первое уравнение, найдем:

$$x_1=-3x_3-5x_5.$$

Итак, общее решение имеет вид $x_1=-3x_3-5x_5$ $x_2=2x_3+3x_5$, $x_4=0$. Давая свободным неизвестным поочередно значения, равные элементам строк определителя $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$, найдем фундаментальную систему решений:

x_1	x_2	x_3	x_4	x_5
-3	2	1	0	0
-5	3	0	0	1

Если ранг системы однородных линейных уравнений единицу меньше числа неизвестных: $r=n-1$, то $n-r=$

т. е. фундаментальная система состоит из одного решения и любое ненулевое решение образует фундаментальную систему. В этом случае любые два решения различаются между собой лишь числовым множителем.

В качестве решения системы $n-1$ однородных линейных уравнений с n неизвестными, т. е. системы вида

$$a_{i1}x_1+\ldots+a_{in}x_n=0, \quad i=1, 2, \ldots, n-1,$$

можно взять совокупность миноров, полученных из матрицы системы поочередным вычеркиванием каждого столбца, причем эти миноры берутся с чередующимися знаками. Таким образом, если

$$M_i=\begin{vmatrix} a_{11} & \ldots & a_{1,i-1} & a_{1,i+1} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2,i-1} & a_{2,i+1} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-1,1} & \ldots & a_{n-1,i-1} & a_{n-1,i+1} & \ldots & a_{n-1,n} \end{vmatrix}, \quad i=1, 2, \ldots, n,$$

то

$$x_1=M_1, \quad x_2=-M_2, \quad x_3=M_3, \quad \ldots, x_n=(-1)^{n+1}M_n.$$

При этом если не все эти миноры равны нулю, то полученное решение составляет фундаментальную систему решений и все другие решения ему пропорциональны ([11], стр. 83 или [20], задача 748). В противном случае $r<n-1$, и этот метод дает лишь нулевое значение.

Пример 2. Найти общее решение и фундаментальную систему решений системы уравнений

$$\begin{aligned} 3x-2y+4z&=0, \\ 2x-3y+5z&=0. \end{aligned}$$

Составляем матрицу системы:

$$\begin{pmatrix} 3 & -2 & 4 \\ 2 & -3 & 5 \end{pmatrix}.$$

Поэтому

$$x=\begin{vmatrix} -2 & 4 \\ -3 & 5 \end{vmatrix}=2, \quad y=-\begin{vmatrix} 3 & 4 \\ 2 & 5 \end{vmatrix}=-7, \quad z=\begin{vmatrix} 3 & -2 \\ 2 & -3 \end{vmatrix}=-5.$$

Так как это решение ненулевое, то оно образует фундаментальную систему решений. Все решения ему пропорциональны. Значит, общее решение имеет вид: $x=2c$, $y=-7c$, $z=-5c$, где c — параметр, принимающий любые числовые значения.

7. Связь решений однородных и неоднородных систем. Если в неоднородной системе линейных уравнений заменить все свободные члены нулями, то получится однородная система, называемая *приведенной системой* для исходной неоднородной системы. Решения данной неоднородной системы и соответствующей ей приведенной системы связаны следующим образом.

1) Сумма любого решения неоднородной системы и любого решения ее приведенной системы является решением неоднородной системы.

2) Разность двух любых решений неоднородной системы является решением ее приведенной системы.

Из этих двух свойств следует:

3) Все решения неоднородной системы можно получить, прибавляя к одному (любому) ее решению поочередно все решения ее приведенной системы. Иными словами, общее решение неоднородной системы можно получить, если к любому частному решению этой системы прибавить общее решение ее приведенной системы.

Пример. Найти общее решение системы уравнений

$$3x-2y+4z=5,$$
$$2x-3y+5z=4.$$

Эта система имеет частное решение $x=y=z=1$. Общее решение ее приведенной системы найдено в примере 2 предыдущего пункта и имеет вид: $x=2c$, $y=-7c$, $z=-5c$, где c — параметр, принимающий любые значения. Поэтому общее решение данной неоднородной системы уравнений имеет вид

$$x=1+2c,\ y=1-7c,\ z=1-5c,$$

где c — произвольный параметр.

ГЛАВА II

МАТРИЦЫ И КВАДРАТИЧНЫЕ ФОРМЫ

§ 1. Матрицы

1. Действия с матрицами. О ранге матрицы см. гл. I, § 2, п. 4.

Определения матрицы, ее строк и столбцов, квадратной матрицы, ее порядка, главной и побочной диагоналей даны во введении.

Прямоугольную (в частности, квадратную) матрицу из m строк и n столбцов мы будем называть (m, n)-*матрицей* или матрицей *размеров* $m \times n$ и обозначать сокращенно через $A = (a_{ij})_{m,n}$, а квадратную матрицу порядка n — через $A = (a_{ij})_1^n$. Две матрицы с одинаковым числом строк и одинаковым числом столбцов будем называть матрицами одинаковых размеров.

Для матриц определяются три основные операции: сложение, умножение и умножение матрицы на число.

Суммой двух матриц $A = (a_{ij})_{m,n}$ и $B = (b_{ij})_{m,n}$ одинаковых размеров $m \times n$ называется матрица тех же размеров

$$A + B = (a_{ij})_{m,n} + (b_{ij})_{m,n} = (a_{ij} + b_{ij})_{m,n}.$$

Пример 1.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & -2 & 1 \\ 2 & 3 & -6 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 4 \\ 6 & 8 & 0 \end{pmatrix}.$$

Произведением двух матриц $A = (a_{ij})_{m,n}$ и $B = (b_{jk})_{n,p}$, которые заданы в определенном порядке (первая и вторая) и размеры которых связаны условием: число столбцов первой матрицы равно числу строк второй, называется матрица $C = (c_{i,k})_{m,p}$, элементы которой определяются

3*

равенствами

$$c_{ik}=\sum_{j=1}^{n} a_{ij}b_{jk}; \quad i=1, 2, \ldots, m; \quad k=1, 2, \ldots, p. \tag{2.1}$$

Таким образом, элемент произведения матриц A и B, стоящий в i-й строке и k-м столбце, равен сумме произведений элементов i-й строки первой матрицы A на соответствующие элементы k-го столбца второй матрицы B. Произведение матриц A и B в данном порядке обозначается через AB или $A \cdot B$. Для квадратных матриц одинакового порядка произведение AB (как и BA) всегда определено.

Пример 2. Если

$$A=\begin{pmatrix} 4 & -1 \\ 5 & -2 \end{pmatrix}, \qquad B=\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

то

$$AB=\begin{pmatrix} 0 & 3 & 6 \\ -3 & 0 & 3 \end{pmatrix}.$$

Произведение BA не определено.

Умножение матриц тесно связано с линейными преобразованиями неизвестных. Пусть даны два линейных преобразования неизвестных:

$$x_i=\sum_{j=1}^{n} a_{ij}y_j, \qquad i=1, 2, \ldots, m; \tag{2.2}$$

$$y_j=\sum_{k=1}^{p} b_{jk}z_k, \qquad j=1, 2, \ldots, n. \tag{2.3}$$

Подставляя выражения для y_j из равенств (2.3) в (2.2), получим снова линейное преобразование вида

$$x_i=\sum_{k=1}^{p} c_{ik}z_k, \qquad i=1, 2, \ldots, m. \tag{2.4}$$

Оно называется результатом последовательного выполнения или *произведением* данных преобразований (2.2) и (2.3). Если

$$A=(a_{ij})_{m,n}, \qquad B=(b_{jk})_{n,p}, \qquad C=(c_{ik})_{m,p}$$

— соответственно матрицы коэффициентов преобразований (2.2), (2.3) и их произведения (2.4), то $AB=C$.

Прозведением матрицы $A=(a_{ij})_{m,n}$ *на число* c называется матрица $cA=(ca_{ij})_{m,n}$, получаемая умножением всех элементов матрицы A на число c.

Действия над матрицами обладают следующими свойствами (во всех следующих равенствах предполагается, что все действия, встречающиеся в одной из двух частей равенства, имеют смысл, и отсюда следует, что действия в другой части равенства также имеют смысл):

1) Сложение матриц коммутативно: $A+B=B+A$.

2) Сложение матриц ассоциативно: $A+(B+C)=(A+B)+C$.

Из этих двух свойств можно вывести, что в любой сумме конечного числа матриц слагаемые можно писать в любом порядке и скобки, указывающие порядок действий, можно расставлять произвольно. Например, $[C+(D+A)]+B=[(A+B)+C]+D$, и можно писать просто $A+B+C+D$.

3) Сложение обладает однозначной обратной операцией — *вычитанием*. Это значит, что для любых двух матриц A и B одинаковых размеров существует единственная матрица C со свойством $B+C=A$. Она называется *разностью* матриц A и B и обозначается через $A-B$.

Свойство 3) эквивалентно следующим двум:

3а) Существует единственная матрица O данных размеров, называемая *нулевой*, такая, что $A+O=A$ для любой матрицы A данных размеров. Именно, нулевой матрицей служит матрица, у которой все элементы равны нулю.

3б) Для любой матрицы A данных размеров существует единственная матрица $-A$ тех же размеров со свойством $A+(-A)=O$. Именно, за матрицу $-A$ надо взять матрицу, все элементы которой отличаются знаком от соответствующих элементов матрицы A.

Умножение матриц не коммутативно, т. е. не обладает переместительным свойством. В общем случае $AB\neq BA$ даже для квадратных матриц одинакового порядка $n\geqslant 2$.

Пример 3.

$$A=\begin{pmatrix}1&2\\3&4\end{pmatrix},\quad B=\begin{pmatrix}2&1\\4&3\end{pmatrix},\quad AB=\begin{pmatrix}10&7\\22&15\end{pmatrix},$$

$$BA=\begin{pmatrix}5&8\\13&20\end{pmatrix},\quad AB\neq BA.$$

4) Умножение матриц ассоциативно: $A(BC)=(AB)C$.

5) Умножение матриц дистрибутивно относительно сложения:

$$(A+B)C=AC+BC, \qquad C(A+B)=CA+CB.$$

Отсюда следует дистрибутивность умножения матриц относительно вычитания: $(A-B)C=AC-BC$, $C(A-B)=CA-CB$.

6) $a(bA)=(ab)A$, где a, b — числа и A — матрица.

7) $a(AB)=(aA)B$, где a — число и A, B — матрицы.

Матрицу, транспонированную для матрицы A, обозначим через A' (см. гл. I, § 1, п. 3).

8) Для транспонирования произведения матриц справедлива формула $(AB)'=B'A'$, и аналогично для любого конечного числа сомножителей.

Обозначим определитель квадратной матрицы A через $|A|$.

9) Определитель произведения двух квадратных матриц одинакового порядка равен произведению их определителей:

$$|AB|=|A|\cdot|B|,$$

и аналогично для любого конечного числа сомножителей.

10) Выражение минора произведения двух матриц через миноры сомножителей (обозначения даны на стр. 44): если A — матрица размеров $m\times n$, B — матрица размеров $n\times p$ и $C=AB$, то

$$C\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ k_1 & k_2 & \dots & k_r\end{pmatrix}=$$
$$=\sum_{1\leqslant j_1<j_2<\dots<j_r\leqslant n} A\begin{pmatrix} i_1 & i_2 & \dots & i_r \\ j_1 & j_2 & \dots & j_r\end{pmatrix} B\begin{pmatrix} j_1 & j_2 & \dots & j_r \\ k_1 & k_2 & \dots & k_r\end{pmatrix}$$

при $r\leqslant n$. Все миноры матрицы C порядка $r>n$ равны нулю ([6], стр. 19).

Это равенство при $m=p=r$ превращается в теорему Бине—Коши (стр. 35).

Пример 4. Умножая ненулевой столбец

$$A=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m\end{pmatrix}$$

на ненулевую строку $B=(b_1\ b_2\ \dots\ b_p)$, получим матрицу

$$C=AB=\begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_p \\ a_2b_1 & a_2b_2 & \dots & a_2b_p \\ \dots & \dots & \dots & \dots \\ a_mb_1 & a_mb_2 & \dots & a_mb_p \end{pmatrix},$$

в которой все миноры порядка выше первого равны нулю, но имеется элемент, отличный от нуля. Таким образом, ранг матрицы C равен единице.

Об операциях над клеточными матрицами см. [6], гл. II, § 5.

2. Единичная и обратная матрицы. *Единичной матрицей* порядка n называется квадратная матрица E порядка n, обладающая свойством $AE=EA=A$ для любой матрицы A порядка n. Таким образом, матрица E играет такую же роль при умножении матриц, какую число 1 при умножении чисел. Для любого порядка n существует единственная единичная матрица данного порядка. Она имеет вид

$$E=\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Квадратная матрица A называется *вырожденной* (или *особенной*), если ее определитель равен нулю, $|A|=0$, и *невырожденной* (или *неособенной*), если $|A|\neq 0$.

Обратной матрицей для квадратной матрицы A порядка n называется квадратная матрица A^{-1} того же порядка, обладающая свойством $AA^{-1}=A^{-1}A=E$, где E — единичная матрица порядка n. Вырожденная матрица не имеет обратной, так как из свойства 9) предыдущего пункта следует, что произведение вырожденной матрицы на любую квадратную матрицу того же порядка будет снова вырожденной матрицей. Любая невырожденная матрица $A=(a_{ij})_1^n$ имеет единственную обратную матрицу, а именно:

$$A^{-1}=\begin{pmatrix} \frac{A_{11}}{|A|} & \frac{A_{21}}{|A|} & \dots & \frac{A_{n1}}{|A|} \\ \frac{A_{12}}{|A|} & \frac{A_{22}}{|A|} & \dots & \frac{A_{n2}}{|A|} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{A_{1n}}{|A|} & \frac{A_{2n}}{|A|} & \dots & \frac{A_{nn}}{|A|} \end{pmatrix},$$

где A_{ij} — алгебраическое дополнение элемента a_{ij} в матрице A (стр. 20).

Матрицей, обратной для матрицы A^{-1}, будет матрица A, т. е. $(A^{-1})^{-1}=A$. Определители двух взаимно обратных матриц являются числами взаимно обратными: $|A|\cdot|A^{-1}|=1$.

Матрица, обратная для произведения двух невырожденных матриц, равна произведению матриц, обратных сомножителям, взятым в обратном порядке:

$$(AB)^{-1}=B^{-1}A^{-1},$$

и аналогично для любого конечного числа сомножителей.

Методы вычисления обратной матрицы. 1. Используют указанный выше вид элементов обратной матрицы.

Пример 1. Найти обратную матрицу для матрицы

$$A=\begin{pmatrix}1 & 3\\ 4 & 5\end{pmatrix};\quad |A|=-7.$$

Находим:

$$A^{-1}=-\frac{1}{7}\begin{pmatrix}5 & -3\\ -4 & 1\end{pmatrix}=\begin{pmatrix}-\frac{5}{7} & \frac{3}{7}\\ \frac{4}{7} & -\frac{1}{7}\end{pmatrix}.$$

Этот метод неудобен для матриц высокого порядка.

2. Обозначим элементы i-го столбца обратной матрицы A^{-1} для матрицы $A=(a_{ij})_1^n$ через $x_1, x_2, \ldots, x_n$. Приравнивая элементы i-го столбца в матричном равенстве $AA^{-1}=E$, получаем системы уравнений

$$a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n=\begin{cases}0 \text{ при } k\neq i,\\ 1 \text{ при } k=i,\end{cases}\quad k=1,2,\ldots,n,$$

где в правой части i-го уравнения стоит единица, а во всех остальных уравнениях — нули. Так как $|A|\neq 0$, то по правилу Крамера (гл. I, § 2, п. 3) эта система имеет единственное решение, которое можно найти методом исключения неизвестных (гл. I, § 2, п. 2). Применение здесь формул Крамера привело бы к предыдущему методу вычисления обратной матрицы.

Пример 2. Найти обратную матрицу для матрицы

$$A=\begin{pmatrix}0 & 1 & \dots & 1\\ 1 & 0 & \dots & 1\\ \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \dots & 0\end{pmatrix}.$$

Для элементов i-го столбца обратной матрицы имеем систему уравнений

$$x_1+\dots+x_{k-1}+x_{k+1}+\dots+x_n=\begin{cases}0 \text{ при } k\neq i,\\ 1 \text{ при } k=i.\end{cases}$$

Сложив все уравнения и разделив на $n-1$, получим

$$\sum_{k=1}^{n} x_k=\frac{1}{n-1}.$$

Вычитая отсюда каждое из уравнений системы, найдем

$$x_1=x_2=\dots=x_{i-1}=x_{i+1}=\dots=x_n=\frac{1}{n-1};$$

$$x_i=\frac{1}{n-1}-1=\frac{2-n}{n-1}.$$

Отсюда

$$A^{-1}=\frac{1}{n-1}\begin{pmatrix}2-n & 1 & \dots & 1\\ 1 & 2-n & \dots & 1\\ \cdot & \cdot & \cdot & \cdot\\ 1 & 1 & \dots & 2-n\end{pmatrix}.$$

3. Любую невырожденную матрицу A путем элементарных преобразований (стр. 54) только строк (или только столбцов) можно привести к единичной матрице E. Применяя ту же последовательность преобразований к единичной матрице E, получим обратную матрицу A^{-1}. Удобно совершать элементарные преобразования над A и E одновременно, записывая обе матрицы рядом через черту ([20], задача 933).

Пример 3. Найдем этим методом обратную матрицу для матрицы примера 1 этого пункта. Соединяем матрицы знаком эквивалентности $\sim$:

$$\left(\begin{array}{cc|cc}1 & 3 & 1 & 0\\ 4 & 5 & 0 & 1\end{array}\right)\sim$$

(из второй строки вычитаем первую, умноженную на 4)

$$\sim\left(\begin{array}{cc|cc}1 & 3 & 1 & 0\\ 0 & -7 & -4 & 1\end{array}\right)\sim$$

$\left(\text{к первой строке прибавляем вторую, умноженную на } \frac{3}{7}, \text{ и затем вторую строку умножаем на } -\frac{1}{7}\right)$

$$\sim\left(\begin{array}{cc|cc} 1 & 0 & -\frac{5}{7} & \frac{3}{7} \\ 0 & 1 & \frac{4}{7} & -\frac{1}{7} \end{array}\right); \quad A^{-1}=\begin{pmatrix} -\frac{5}{7} & \frac{3}{7} \\ \frac{4}{7} & -\frac{1}{7} \end{pmatrix}.$$

Миноры обратной матрицы $B=A^{-1}$ для невырожденной матрицы A порядка n следующим образом выражаются через миноры матрицы A. Если

$$1 \leqslant \begin{matrix} i_1 < i_2 < \ldots < i_p \\ j_1 < j_2 < \ldots < j_p \end{matrix} \leqslant n,$$

то

$$B\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix} = \frac{(-1)^{\sum\limits_{s=1}^{p}(i_s+j_s)} A\begin{pmatrix} j'_1 & j'_2 & \ldots & j'_{n-p} \\ i'_1 & i'_2 & \ldots & i'_{n-p} \end{pmatrix}}{|A|},$$

где система индексов $i'_1 < i'_2 < \ldots < i'_{n-p}$ дополняет систему $i_1 < i_2 < \ldots < i_p$, а система $j'_1 < j'_2 < \ldots < j'_{n-p}$ дополняет систему $j_1 < j_2 < \ldots < j_p$ до полной системы индексов $1, 2, \ldots, n$ ([6], стр. 26 или [20], задача 972).

С вычислением обратной матрицы тесно связано решение *матричных уравнений* вида

$$AX=B, \qquad YA=B, \tag{2.5}$$

где A, B — данные и X, Y — искомые матрицы. Если A — прямоугольная или вырожденная квадратная матрица, то решение этих уравнений сводится к решению систем линейных уравнений для элементов каждого столбца матрицы X или каждой строки матрицы Y; для получения этих уравнений надо приравнять элементы матриц в обеих частях равенства. Но если A — невырожденная матрица, то решения даются формулами

$$X=A^{-1}B, \qquad Y=BA^{-1}. \tag{2.6}$$

В частности, если матрица B состоит из одного столбца (как и искомая матрица X или Y), то первая из формул (2.6) приводит к формулам Крамера (гл. I, § 2, п. 3). Однако эта формула удобнее формул Крамера в том случае, когда надо решать ряд систем уравнений, отличающихся лишь правыми частями. Тогда, найдя матрицу A^{-1}, легко вычислять $X = A^{-1}B$ для различных частных значений B.

Пользуясь действиями с матрицами, можно при известных условиях свести вычисление определителя четного порядка к вычислению одного определителя вдвое меньшего порядка. Пусть определитель $\Delta = \begin{vmatrix} A & B \\ C & D \end{vmatrix}$ порядка $2n$ разбит на четыре клетки A, B, C, D порядка n. Тогда имеют место

Формулы Шура:

$$\Delta = |AD - ACA^{-1}B|, \qquad \text{если } |A| \neq 0;$$
$$\Delta = |AD - BD^{-1}CD|, \qquad \text{если } |D| \neq 0;$$
$$\Delta = |AD - CB|, \qquad \text{если } AC = CA;$$
$$\Delta = |AD - BC|, \qquad \text{если } CD = DC$$ ([6], стр. 45).

3. Степени матрицы. Многочлены от матрицы. Перестановочные матрицы. Матрицу A можно умножить саму на себя тогда и только тогда, когда она квадратная. Все матрицы в этом пункте (если не оговорено противное) предполагаются квадратными одного и того же порядка n.

Степени матрицы определяются так: если k — натуральное число, большее единицы, то A^k есть произведение k матриц, равных A. Из закона ассоциативности умножения матриц (стр. 70) вытекает, что это произведение имеет однозначный смысл. Далее, $A^1 = A$, $A^0 = E$, где E — единичная матрица. Степени с отрицательными показателями определяются лишь для невырожденных матриц. Именно, A^{-1} — обратная матрица (стр. 71), и при целом положительном k, $k > 1$, матрица A^{-k} определяется как любая из двух равных матриц $(A^k)^{-1} = (A^{-1})^k$.

Определение корня из матрицы и связанного с ним понятия степени матрицы с дробным показателем см. в [6], гл. 8, §§ 6, 7. Степени матрицы с любыми действительными или комплексными показателями можно определить как специальный случай общего понятия *функции от матрицы* ([6], гл. 5).

Для действий со степенями матриц справедливы следующие правила (где k и l — любые целые числа; в случае отрицательных показателей матрицы предполагаются невырожденными):

$$A^kA^l = A^{k+l}, \tag{2.7}$$

$$(A^k)^l = A^{kl}, \tag{2.8}$$

$$(AB)^k = A^kB^k, \tag{2.9}$$

если только

$$AB = BA.$$

Если $AB = BA$, то верна формула бинома Ньютона

$$(A+B)^m = A^m + C_m^1 A^{m-1}B + C_m^2 A^{m-2}B^2 + \ldots + B^m,$$

где m — целое положительное число.

Матрицы A и B называются *перестановочными*, если оба произведения AB и BA имеют смысл и $AB = BA$. Две перестановочные матрицы обязательно являются квадратными одинакового порядка. Если A перестановочна с B, то A перестановочна с B^{-1}, A^{-1} с B и A^{-1} с B^{-1}, если только обратные матрицы существуют.

Из (2.7) следует, что все степени одной и той же матрицы A перестановочны между собой.

Скалярной матрицей называется квадратная матрица вида $A = cE$, где E — единичная матрица.

Для того чтобы квадратная матрица A порядка n была перестановочна со всеми матрицами того же порядка, необходимо и достаточно, чтобы она была скалярной ([11], стр. 105).

Если матрица A перестановочна с какой-нибудь диагональной матрицей (см. стр. 89), у которой элементы главной диагонали попарно различны между собой, то эта матрица A сама диагональна ([20], задача 820).

Значением многочлена

$$f(x) = c_0x^m + c_1x^{m-1} + \ldots + c_m$$

с числовыми коэффициентами $c_0, c_1, \ldots, c_m$ от матрицы A или значением многочлена $f(x)$ при $x = A$ называется матрица

$$f(A) = c_0A^m + c_1A^{m-1} + \ldots + c_mE,$$

где E — единичная матрица.

Из перестановочности степеней матрицы A следует, что все многочлены от матрицы A перестановочны между собой и, в частности, перестановочны с матрицей A.

Частное матриц A и B имеет смысл лишь для невырожденного делителя B, причем это частное имеет два значения: *левое частное* $B^{-1}A$ и *правое частное* AB^{-1}. Оба эти частных совпадают тогда и только тогда, когда A и B перестановочны. Только в этом случае применяются обозначения частного $A:B$ или $\frac{A}{B}$.

Рациональной функцией

$$h(x)=\frac{f(x)}{g(x)}$$

(где $f(x)$ и $g(x)$ — многочлены с числовыми коэффициентами) *от матрицы* A называется матрица

$$h(A)=\frac{f(A)}{g(A)}.$$

Это значение $h(A)$ имеет смысл тогда и только тогда, когда матрица $g(A)$ не вырождена. Все рациональные функции от матрицы A перестановочны между собой и с матрицей A.

Обратная матрица A^{-1} для невырожденной матрицы A, а потому и рациональная функция

$$h(A)=\frac{f(A)}{g(A)}=f(A)\,[g(A)]^{-1}$$

представляются в виде многочлена от матрицы A ([6], стр. 98).

При известных условиях можно определить значение любой функции $f(x)$ от матрицы A как значение некоторого многочлена от этой матрицы ([6], гл. 5, § 1).

Поэтому значение любой функции от матрицы A есть матрица, перестановочная с матрицей A.

Кроме многочленов от матрицы A, могут существовать и другие матрицы, перестановочные с A. Все матрицы, перестановочные с данной матрицей A, тогда и только тогда являются многочленами от A, когда все инвариантные множители (стр. 83) ее характеристической матрицы, кроме последнего, равны единице или (что то же самое) когда миноры $(n-1)$-го порядка характеристической матрицы $\lambda E - A$ в совокупности взаимно просты ([6], стр. 184). Без этого условия матрицы, перестановочные с A, не могут исчерпываться многочленами не только от матрицы A, но вообще ни от какой-либо одной матрицы ([6], стр. 184).

В самом общем случае все матрицы X, перестановочные с данной квадратной матрицей A порядка n, находятся следующим образом. Пусть U — какая-нибудь невырожденная матрица, трансформирующая матрицу A в *жорданову матрицу* A_0, т. е. удовлетворяющая равенству $A = UA_0U^{-1}$ (п. 11). Тогда $X = UX_0U^{-1}$, где X_0 — любая матрица, перестановочная с жордановой матрицей A_0. Все матрицы X_0 находятся по следующему правилу: назовем прямоугольную (в частности, квадратную) матрицу *правильной верхней треугольной*, если все ее элементы, стоящие ниже каждой из диагоналей, выходящих из левого верхнего и правого нижнего углов, равны нулю, а среди остальных элементов все элементы, стоящие на каждой параллели к указанным диагоналям, равны между собой, например:

$$\begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix}, \quad \begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{pmatrix}.$$

Пусть клетки Жордана, стоящие на главной диагонали матрицы A_0 в порядке сверху вниз, имеют порядки $p_1, p_2, \ldots$ $\ldots, p_s$ и на их диагонали стоят числа $\lambda_1, \lambda_2, \ldots, \lambda_s$ (среди которых могут быть и равные между собой). Проведя горизонтальные и вертикальные линии, отделяющие клетки Жордана, мы разобьем матрицу A_0 на квадратные и прямоугольные клетки. Пусть $X_0 = (X_{\alpha\beta})_1^s$ — соответствующее разбиение на клетки матрицы X_0. Здесь $X_{\alpha\beta}$ — клетка размеров $p_\alpha \times p_\beta$. Матрица X_0 определяется условиями: $X_{\alpha\beta} = 0$, если $\lambda_\alpha \neq \lambda_\beta$, и $X_{\alpha\beta}$ — любая правильная верхняя треугольная матрица, если $\lambda_\alpha = \lambda_\beta$. Число независимых параметров в матрице X_0 или число линейно независимых матриц X, перестановочных с A, определяется формулой

$$N = n_1 + 3n_2 + 5n_3 + \ldots + (2t-1)\,n_t,$$

где $n_1 \geqslant n_2 \geqslant n_3 \geqslant \ldots \geqslant n_t > 0$ — степени инвариантных множителей матрицы $\lambda E - A$, отличных от единицы ([6], стр. 183—184).

Пример. Если матрица $\lambda E - A$ имеет инвариантные множители $E_1(\lambda) = \ldots = E_5(\lambda) = 1$, $E_6(\lambda) = (\lambda - \lambda_1)^2$, $E_7(\lambda) = (\lambda - \lambda_1)^3 (\lambda - \lambda_2)^2$, то

$$X_0 = \left(\begin{array}{ccc:cc:cc} a & b & c & d & e & 0 & 0 \\ 0 & a & b & 0 & d & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 \\ \hdashline 0 & f & g & h & k & 0 & 0 \\ 0 & 0 & f & 0 & h & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & l & m \\ 0 & 0 & 0 & 0 & 0 & 0 & l \end{array}\right).$$

Число параметров $a, b, \ldots, m$ равно $N = n_1 + 3n_2 = 5 + 3 \cdot 2 = 11$.

4. Связь умножения матриц с элементарными преобразованиями. Разложение матрицы в произведение треугольных матриц. Каждое элементарное преобразование (стр. 54) строк (или столбцов) прямоугольной матрицы A размеров $m \times n$ можно заменить умножением матрицы A слева (соответственно справа) на квадратную невырожденную матрицу порядка m (соответственно n), полученную из единичной матрицы того же порядка применением того же самого элементарного преобразования.

Пример.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Прибавляя ко второй строке первую, умноженную на -2, получим матрицу

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{pmatrix}.$$

Применяя то же преобразование к единичной матрице второго порядка, получим матрицу

$$P = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}.$$

Тогда должно выполняться равенство $PA = B$, т. е.

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

Любую матрицу A размеров $m \times n$ и ранга r можно представить в виде $A = PRQ$, где P и Q — невырожденные матрицы соответственно порядков m и n, а R — матрица размеров $m \times n$, у которой первые r элементов главной диагонали (выходящей из левого верхнего угла) равны единице, а остальные элементы равны нулю.

Верхней (*нижней*) *треугольной* матрицей называется квадратная матрица, у которой элементы, отличные от нуля, могут стоять лишь на главной диагонали и выше (ниже) этой диагонали.

Сумма, произведение, произведение на число и обратная матрица (в случае невырожденной матрицы), взятые для верхних (нижних) треугольных матриц, снова будут верхними (нижними) треугольными матрицами.

Любую квадратную матрицу A можно представить в виде $A = PR$ или $A = RQ$, где R, по желанию, верхняя или нижняя треугольная матрица, а P и Q — невырожденные матрицы.

Любую квадратную матрицу $A = (a_{ij})_1^n$ ранга r, у которой первые r угловых миноров (т. е. миноров, стоящих в левом верхнем углу) отличны от нуля,

$$d_k = A\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0; \quad k = 1, 2, \dots, r, \tag{2.10}$$

можно представить в виде произведения

$$A = BC, \tag{2.11}$$

где

$$B = \begin{pmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

— нижняя треугольная,

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & c_{nn} \end{pmatrix}$$

— верхняя треугольная матрицы. При этом первые r элементов главных диагоналей матриц B и C можно выбрать любыми, удовлетворяющими условиям

$$\left.\begin{aligned} b_{11}\, c_{11} &= d_1, \\ b_{22}\, c_{22} &= \frac{d_2}{d_1}, \\ &\ldots\ldots \\ b_{rr}\, c_{rr} &= \frac{d_r}{d_{r-1}}. \end{aligned}\right\} \tag{2.12}$$

Задание этих элементов однозначно определяет все элементы первых r столбцов матрицы B и первых r строк матрицы C. Эти элементы можно найти по формулам

$$\left.\begin{aligned} b_{ik} &= \frac{b_{kk}}{d_k} A \begin{pmatrix} 1 & 2 & \ldots & k-1 & i \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix}, \\ c_{ki} &= \frac{c_{kk}}{d_k} A \begin{pmatrix} 1 & 2 & \ldots & k-1 & k \\ 1 & 2 & \ldots & k-1 & i \end{pmatrix}; \\ i &= k,\ k+1,\ \ldots,\ n; \quad k = 1,\ 2,\ \ldots,\ r, \end{aligned}\right\} \tag{2.13}$$

или по рекуррентным формулам

$$\left.\begin{aligned} & b_{ik} = \frac{a_{ik} - \sum_{j=1}^{k-1} b_{ij} c_{jk}}{c_{kk}}, \quad c_{ki} = \frac{a_{ki} - \sum_{j=1}^{k-1} b_{kj} c_{ji}}{b_{kk}}; \\ & i = k,\ k+1,\ \ldots,\ n; \quad k = 1,\ 2,\ \ldots,\ r. \end{aligned}\right\} \tag{2.14}$$

В случае $r < n$ все элементы последних $n-r$ столбцов матрицы B можно взять равными нулю, а элементы последних $n-r$ строк матрицы C можно выбрать произвольными или, наоборот, элементы последних $n-r$ столбцов матрицы B взять произвольными, а элементы последних $n-r$ строк матрицы C взять равными нулю ([6], стр. 37—38).

Если $A = (a_{ij})_1^n$ — симметрическая матрица ранга r, угловые миноры которой удовлетворяют условиям (2.10), то A представляется в виде

$$A = BB', \tag{2.15}$$

где $B = (b_{ij})_1^n$ — нижняя треугольная матрица и B' — матрица, транспонированная для B. Элементы матрицы B находятся по формулам

$$\left.\begin{aligned} b_{ik} &= \frac{1}{\sqrt{d_k d_{k-1}}} A\begin{pmatrix} 1 & 2 & \dots & k-1 & i \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}; \quad \begin{matrix} i=k, k+1, \dots, n, \\ k=1, 2, \dots, r; \end{matrix} \\ b_{ik} &= 0; \quad k=r+1, r+2, \dots, n \end{aligned}\right\} \qquad (2.16)$$

([6], стр. 39).

5. Многочленные матрицы, инвариантные множители и элементарные делители. *Многочленной матрицей* или *λ-матрицей* называется прямоугольная (в частности, квадратная) матрица $A(\lambda) = (a_{ij}(\lambda))_{m,n}$, элементы которой являются многочленами от одного переменного λ с числовыми коэффициентами. Если все эти коэффициенты принадлежат к данному числовому полю K (см. введение), то будем матрицу $A(\lambda)$ называть *λ-матрицей над полем K*.

Элементарными преобразованиями λ-матрицы $A(\lambda)$ называются преобразования следующих типов:

(I) перестановка двух строк,

(II) умножение строки на число c, отличное от нуля,

(III) прибавление к одной строке другой строки, умноженной на любой многочлен $f(\lambda)$, и аналогичные преобразования столбцов.

Две λ-матрицы $A(\lambda)$ и $B(\lambda)$ одинаковых размеров $m \times n$ называются *эквивалентными*, в символах $A(\lambda) \sim B(\lambda)$, если от матрицы $A(\lambda)$ к $B(\lambda)$ можно перейти при помощи цепочки из конечного числа элементарных преобразований. Отношение эквивалентности обладает тремя основными свойствами:

1) рефлективность: каждая матрица эквивалентна сама себе $A(\lambda) \sim A(\lambda)$;

2) симметрия: если $A(\lambda) \sim B(\lambda)$, то $B(\lambda) \sim A(\lambda)$;

3) транзитивность: если $A(\lambda) \sim B(\lambda)$ и $B(\lambda) \sim C(\lambda)$, то $A(\lambda) \sim C(\lambda)$.

Отсюда следует, что совокупность всех λ-матриц данных размеров $m \times n$ разбивается на непересекающиеся классы эквивалентных матриц, т. е. на такие классы, что любые две матрицы из одного класса эквивалентны, а из разных классов — не эквивалентны между собой. Возникает вопрос о канонической форме λ-матрицы, характеризующей данный класс эквивалентных λ-матриц.

Канонической или *нормальной* λ-матрицей размеров $m \times n$ называется λ-матрица, у которой на главной диагонали стоят многочлены $E_1(\lambda)$, $E_2(\lambda)$, ..., $E_p(\lambda)$, где p — меньшее из чисел m и n, причем не равные нулю среди этих многочленов имеют старшие коэффициенты, равные единице, и каждый следующий многочлен делится на предыдущий, все же элементы вне главной диагонали равны нулю.

Из этих свойств следует, что если среди многочленов $E_i(\lambda)$ имеются многочлены нулевой степени, то они равны единице и стоят в начале диагонали, если же имеются нули, то они стоят в конце главной диагонали. Так, матрица

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & \lambda(\lambda-3) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

является канонической.

Каждый класс эквивалентных λ-матриц размеров $m \times n$ содержит единственную каноническую матрицу, т. е. каждая λ-матрица $A(\lambda)$ эквивалентна единственной канонической λ-матрице $C(\lambda)$, называемой *канонической формой* или *нормальной формой* данной матрицы $A(\lambda)$, а многочлены $E_1(\lambda)$, $E_2(\lambda)$, ... , $E_p(\lambda)$, стоящие на главной диагонали $C(\lambda)$, называются *инвариантными множителями* матрицы $A(\lambda)$. Здесь p — меньший из размеров матрицы $A(\lambda)$ ([14], стр. 107—112).

П е р в ы й м е т о д в ы ч и с л е н и я и н в а р и а н т н ы х м н о ж и т е л е й λ - м а т р и ц ы состоит в приведении ее к канонической форме путем элементарных преобразований.

П р и м е р 1. Найти каноническую форму и инвариантные множители λ-матрицы

$$A(\lambda) = \begin{pmatrix} \lambda & \lambda-1 & \lambda-1 & \lambda-2 \\ \lambda & \lambda^2+\lambda & \lambda-1 & \lambda^2+\lambda-1 \\ \lambda+1 & \lambda^2+\lambda+2 & \lambda+1 & \lambda^2+2\lambda+3 \\ \lambda-1 & \lambda-3 & \lambda-3 & -6 \end{pmatrix}.$$

Третий столбец вычитаем из остальных:

$$A(\lambda) \sim \begin{pmatrix} 1 & 0 & \lambda-1 & -1 \\ 1 & \lambda^2+1 & \lambda-1 & \lambda^2 \\ 0 & \lambda^2+1 & \lambda+1 & \lambda^2+\lambda+2 \\ 2 & 0 & \lambda-3 & -\lambda-3 \end{pmatrix} \sim$$

(из второй и четвертой строк вычитаем первую, умноженную соответственно на 1 и 2)

$$\sim\begin{pmatrix} 1 & 0 & \lambda-1 & -1 \\ 0 & \lambda^3+1 & 0 & \lambda^3+1 \\ 0 & \lambda^3+1 & \lambda+1 & \lambda^3+\lambda+2 \\ 0 & 0 & -\lambda-1 & -\lambda-1 \end{pmatrix}\sim$$

(из третьего и четвертого столбцов вычитаем первый, умноженный соответственно на $\lambda-1$ и -1)

$$\sim\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda^3+1 & 0 & \lambda^3+1 \\ 0 & \lambda^3+1 & \lambda+1 & \lambda^3+\lambda+2 \\ 0 & 0 & -\lambda-1 & -\lambda-1 \end{pmatrix}\sim$$

(второй и третий столбцы вычитаем из четвертого)

$$\sim\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda^3+1 & 0 & 0 \\ 0 & \lambda^3+1 & \lambda+1 & 0 \\ 0 & 0 & -\lambda-1 & 0 \end{pmatrix}\sim$$

(из третьей строки вычитаем вторую и полученную строку прибавляем к четвертой)

$$\sim\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda^3+1 & 0 & 0 \\ 0 & 0 & \lambda+1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\sim$$

(переставляем вторую и третью строки, а также второй и третий столбцы)

$$\sim\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & \lambda^3+1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Это каноническая матрица; значит, матрица $A(\lambda)$ имеет инвариантные множители:

$$E_1(\lambda)=1,\quad E_2(\lambda)=\lambda+1,\quad E_3(\lambda)=\lambda^3+1,\quad E_4(\lambda)=0.$$

Многочлены от λ с коэффициентами из числового (или любого) поля K сами поля не образуют, но они содержатся в поле рациональных функций от λ с коэффициентами из поля K. Поэтому для λ-матриц верны основные свойства ранга, связанные с понятиями минора и линейной зависимости строк или столбцов. Так, при элементарных преобразованиях λ-матрицы ее ранг не изменяется и для λ-матриц верна теорема о ранге матрицы (стр. 56).

Пусть дана λ-матрица $A(\lambda)=(a_{ij}(\lambda))_{m,n}$ ранга r и $p=\min(m, n)$. *Делителями миноров* этой матрицы называются многочлены

$$D_1(\lambda),\ D_2(\lambda),\ \ldots,\ D_p(\lambda),$$

где при $k=1, 2, \ldots, r$ многочлен $D_k(\lambda)$ определяется как наибольший общий делитель (гл. III) всех миноров k-го порядка матрицы $A(\lambda)$, причем коэффициент его старшего члена приведен к единице, и при $k=r+1, r+2, \ldots, p$, по определению, $D_k(\lambda)=0$. При элементарных преобразованиях матрицы $A(\lambda)$ ее делители миноров не изменяются ([14], стр. 109).

Второй метод вычисления инвариантных множителей λ-матрицы основан на следующей связи их с делителями миноров:

$$D_k(\lambda)=E_1(\lambda)E_2(\lambda)\ldots E_k(\lambda),\quad k=1, 2, \ldots, p. \tag{2.17}$$

Отсюда находим

$$\left.\begin{array}{c} E_k(\lambda)=\dfrac{D_k(\lambda)}{D_{k-1}(\lambda)},\quad k=1, 2, \ldots, r \\ (\text{при } k=1 \text{ принято } D_0(\lambda)=1), \\ E_k(\lambda)=0,\quad k=r+1, r+2, \ldots, p \end{array}\right\} \tag{2.18}$$

(см. [14], стр. 112). Этот метод удобен в случаях, когда лишь немногие элементы λ-матрицы зависят от λ, как, например, в характеристической матрице $\lambda E-A$.

Пример 2. Найти инвариантные множители и каноническую форму λ-матрицы

$$A(\lambda)=\begin{pmatrix} \lambda-3 & -1 & 1 \\ 0 & \lambda-2 & 0 \\ -1 & -1 & \lambda-1 \end{pmatrix}.$$

Так как матрица содержит элементы, не зависящие от λ и отличные от нуля, то $D_1(\lambda)=1$. Все миноры второго порядка, содержащие вторую строку матрицы, делятся на $\lambda-2$. Вычисляем остальные три минора.

$$M_{21}=\begin{vmatrix}-1 & 1\\ -1 & \lambda-1\end{vmatrix}=-\lambda+2,\quad M_{22}=\begin{vmatrix}\lambda-3 & 1\\ -1 & \lambda-1\end{vmatrix}=\lambda^2-4\lambda+4,$$

$$M_{23}=\begin{vmatrix}\lambda-3 & -1\\ -1 & -1\end{vmatrix}=-\lambda+2.$$

Поэтому $D_2(\lambda)=\lambda-2$. Разлагая определитель матрицы $A(\lambda)$ по второй строке, найдем

$$D_3(\lambda)=(\lambda-2)M_{22}=(\lambda-2)^3.$$

По формулам (2.18) находим инвариантные множители

$$E_1(\lambda)=1,\quad E_2(\lambda)=\lambda-2,\quad E_3(\lambda)=(\lambda-2)^2.$$

Каноническая форма матрицы $A(\lambda)$ имеет вид

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda-2 & 0\\ 0 & 0 & (\lambda-2)^2\end{pmatrix}.$$

Аналогично случаю числовых матриц (стр. 79) каждое элементарное преобразование строк (или столбцов) λ-матрицы $A(\lambda)$ размеров $m\times n$ можно заменить умножением этой матрицы слева (соответственно справа) на так называемую *элементарную* λ-матрицу, полученную путем применения данного элементарного преобразования к единичной матрице порядка m (соответственно порядка n). Элементарные матрицы являются частным случаем обратимых λ-матриц.

Обратная матрица для невырожденной квадратной λ-матрицы в общем случае не будет λ-матрицей, так как может иметь дробные элементы.

Квадратная λ-матрица называется *обратимой* (или *унимодулярной*), если обратная для нее матрица также является λ-матрицей. Для того чтобы квадратная λ-матрица была обратимой, необходимо и достаточно, чтобы ее определитель не содержал λ и был отличен от нуля, например:

$$\begin{pmatrix}\lambda^2 & \lambda+2\\ \lambda-2 & 1\end{pmatrix}.$$

Произведение нескольких обратимых λ-матриц снова обратимо. Канонической формой обратимой λ-матрицы является единичная матрица того же порядка. Любая обратимая λ-матрица разлагается в произведение конечного числа элементарных λ-матриц ([14], стр. 113—116).

Признаки эквивалентности λ-матриц. Для того чтобы две λ-матрицы $A(\lambda)$ и $B(\lambda)$ одинаковых размеров были эквивалентны, необходимо и достаточно каждое из следующих условий:

1) эти матрицы имеют одну и ту же совокупность делителей миноров;

2) эти матрицы имеют одну и ту же совокупность инвариантных множителей;

3) эти матрицы связаны соотношением

$$B(\lambda) = P(\lambda) A(\lambda) Q(\lambda), \tag{2.19}$$

где $P(\lambda)$ и $Q(\lambda)$ — некоторые обратимые λ-матрицы ([14], стр. 113).

Алгоритм для вычисления матриц $P(\lambda)$ и $Q(\lambda)$. Приводим λ-матрицы $A(\lambda)$ и $B(\lambda)$ одинаковых размеров $m \times n$ к общей канонической форме $C(\lambda)$ путем элементарных преобразований (если канонические формы данных матриц окажутся различными, то эти матрицы не эквивалентны). Совершая в обратном порядке преобразования, обратные тем, которые переводят $B(\lambda)$ в $C(\lambda)$, мы получим цепочку преобразований, переводящих $C(\lambda)$ в $B(\lambda)$, и, поставив перед этой цепочкой преобразования, переводящие $A(\lambda)$ в $C(\lambda)$, получим цепочку преобразований, переводящих $A(\lambda)$ в $B(\lambda)$. Такую цепочку можно иногда построить, минуя каноническую форму $C(\lambda)$. Матрица $P(\lambda)$ получится из единичной матрицы E_m порядка m путем применения к ней в том же порядке всех преобразований строк цепочки, переводящей $A(\lambda)$ в $B(\lambda)$, а матрица $Q(\lambda)$ получится из единичной матрицы E_n порядка n путем применения к ней в том же порядке всех преобразований столбцов той же цепочки. Как цепочка, переводящая $A(\lambda)$ в $B(\lambda)$, так и матрицы $P(\lambda)$ и $Q(\lambda)$ не определены однозначно.

Элементарные делители. Пусть $A(\lambda)$ есть λ-матрица над числовым полем K (см. введение). Элементарными делителями над полем K матрицы $A(\lambda)$ называются максимальные степени неприводимых над полем K множителей (гл. III), входящие в разложения инвариантных множителей этой матрицы. При этом в совокупность элементарных делителей матрицы $A(\lambda)$ каждый из них входит столько раз, в разложении скольких инвариантных множителей он встречается.

Таким образом, если $E_{s+1}(\lambda)$, $E_{s+2}(\lambda)$, ..., $E_r(\lambda)$ — все инвариантные множители матрицы $A(\lambda)$, имеющие положительные степени, и если разложения их на степени попарно различных неприводимых над K множителей имеют вид

$$\left.\begin{aligned} E_i(\lambda) &= [p_{i1}(\lambda)]^{k_{i1}}[p_{i2}(\lambda)]^{k_{i2}}\dots[(p_{it_i}(\lambda)]^{k_{it_i}}, \\ i &= s+1,\ s+2,\ \dots,\ r, \end{aligned}\right\} \tag{2.20}$$

то совокупность элементарных делителей этой матрицы состоит из многочленов

$$e_1(\lambda) = [p_{s+1,1}(\lambda)]^{k_{s+1,1}},$$

$$e_2(\lambda) = [p_{s+1,2}(\lambda)]^{k_{s+1,2}},\ \dots,\ e_t(\lambda) = [p_{rt_r}(\lambda)]^{k_{rt_r}}, \tag{2.21}$$

где $t = t_{s+1} + t_{s+2} + \dots + t_r$ — число элементарных делителей.

При расширении поля K (см. гл. III) делители миноров $D_k(\lambda)$ и инвариантные множители $E_k(\lambda)$ матрицы $A(\lambda)$ не изменяются, но элементарные делители могут измениться.

Пример 3. Найдем элементарные делители λ-матрицы

$$A(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda^2+2 & 0 & 0 & 0 \\ 0 & 0 & (\lambda^4-4)^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

над полями рациональных, действительных и комплексных чисел. Матрица $A(\lambda)$ имеет каноническую форму. Поэтому на ее главной диагонали стоят инвариантные множители

$$E_1(\lambda) = 1,\quad E_2(\lambda) = \lambda^2+2,\quad E_3(\lambda) = (\lambda^4-4)^2,\quad E_4(\lambda) = 0.$$

Разлагая $E_2(\lambda)$ и $E_3(\lambda)$ на неприводимые множители над каждым из указанных полей, получим системы элементарных делителей:

1) над полем рациональных чисел:

$$e_1(\lambda) = \lambda^2+2,\quad e_2(\lambda) = (\lambda^2+2)^2,\quad e_3(\lambda) = (\lambda^2-2)^2;$$

2) над полем действительных чисел:

$$e_1(\lambda) = \lambda^2+2,\quad e_2(\lambda) = (\lambda^2+2)^2,\quad e_3(\lambda) = (\lambda+\sqrt{2})^2,$$
$$e_4(\lambda) = (\lambda-\sqrt{2})^2;$$

3) над полем комплексных чисел:

$$e_1(\lambda) = \lambda+2i,\quad e_2(\lambda) = \lambda-2i,\quad e_3(\lambda) = (\lambda+2i)^2,\quad e_4(\lambda) = (\lambda-2i)^2,$$
$$e_5(\lambda) = (\lambda+\sqrt{2})^2,\quad e_6(\lambda) = (\lambda-\sqrt{2})^2.$$

Так как над полем комплексных чисел любой многочлен положительной степени разлагается на линейные множители (гл. III), то элементарные делители любой λ-матрицы над полем комплексных чисел являются степенями многочленов первой степени, т. е. имеют вид $(\lambda-\lambda_0)^k$, где λ_0 — некоторое число. Обычно под элементарными делителями (если не оговорено противное) понимают элементарные делители над полем комплексных чисел.

Элементарными делителями λ-матрицы $A(\lambda)$ однозначно определяются ее инвариантные множители положительной степени. Последний из этих инвариантных множителей равен произведению различных неприводимых множителей в наивысших степенях, в которых они встречаются среди всех элементарных делителей. Предыдущий инвариантный множитель составляется по тому же правилу из оставшихся элементарных делителей, и т. д. Если еще даны размеры и ранг матрицы $A(\lambda)$, то этим однозначно определяются все ее инвариантные множители и, значит, ее каноническая форма.

Пример 4. Найти инвариантные множители λ-матрицы $A(\lambda)$ размеров 6×8 и ранга 5, если даны ее элементарные делители $\lambda+1$, $\lambda+1$, $(\lambda+1)^3$, $(\lambda-1)^2$, $(\lambda-1)^2$.

Так как число всех инвариантных множителей равно меньшему из размеров матрицы, т. е. 6, а число инвариантных множителей, отличных от нуля, равно рангу, т. е. 5, то $E_6(\lambda)=0$. Далее,

$$E_5(\lambda)=(\lambda+1)^3(\lambda-1)^2,\quad E_4(\lambda)=(\lambda+1)(\lambda-1)^2,\quad E_3(\lambda)=\lambda+1.$$

Так как мы уже использовали все элементарные делители, то других инвариантных множителей положительной степени не может быть. Поэтому $E_2(\lambda)=E_1(\lambda)=1$.

Указанный выше метод вычисления инвариантных множителей по элементарным делителям оказывается полезным в случае диагональных и клеточно диагональных матриц. Квадратная матрица A называется *диагональной*, если все ее элементы, не лежащие на главной диагонали, равны нулю. Диагональная матрица A обозначается через $A=\{a_1, a_2, \ldots, a_n\}$, где в скобках указаны элементы главной диагонали. Аналогично квадратная матрица A называется *клеточно диагональной*, если на ее главной диагонали стоят квадратные матрицы $A_1, A_2, \ldots, A_s$, а все элементы вне этой цепочки матриц равны нулю. Она обозначается через

$$A=\{A_1, A_2, \ldots, A_s\}.$$

Совокупность элементарных делителей клеточно диагональной (в частности, диагональной) λ-матрицы является объединением совокупностей элементарных делителей ее диагональных клеток ([14], стр. 118—119). В этом преимущество элементарных делителей по сравнению с инвариантными множителями.

Пример 5. Найти каноническую форму диагональной λ-матрицы $A(\lambda)=\{\lambda^2, 0, \lambda^2+\lambda, \lambda^2-1, \lambda^4-\lambda^2\}$. Здесь порядок матрицы равен 5, а ранг равен 4. Поэтому $E_5(\lambda)=0$. Элементарные делители диагональных элементов:

$$\lambda^2, \lambda, \lambda+1, \lambda+1, \lambda-1, \lambda^2, \lambda+1, \lambda-1.$$

Значит,

$$E_4(\lambda)=\lambda^2(\lambda+1)(\lambda-1)=\lambda^4-\lambda^2, \quad E_3(\lambda)=\lambda^2(\lambda+1)(\lambda-1)=\lambda^4-\lambda^2,$$
$$E_2(\lambda)=\lambda(\lambda+1)=\lambda^2+\lambda, \quad E_1(\lambda)=1.$$

Каноническая форма имеет вид $\{1, \lambda^2+\lambda, \lambda^4-\lambda^2, \lambda^4-\lambda^2, 0\}$.

Пусть A — квадратная матрица с элементами из данного поля K. Инвариантные множители и элементарные делители ее характеристической матрицы $\lambda E-A$ (или $A-\lambda E$) называются соответственно инвариантными множителями и элементарными делителями матрицы A. Все инвариантные множители матрицы A отличны от нуля, и сумма их степеней равна порядку n матрицы A. Поэтому сумма степеней элементарных делителей матрицы A также равна ее порядку. Так как ранг матрицы $\lambda E-A$ равен n, то по элементарным делителям матрицы A однозначно определяются ее инвариантные множители. О связи элементарных делителей матриц A и $f(A)$, где $f(\lambda)$ — многочлен (или любая функция, для которой $f(A)$ имеет смысл), см. [6], гл. 6, § 7.

6. Матричные многочлены. *Матричным многочленом* называется выражение вида

$$F(\lambda)=A_0\lambda^s+A_1\lambda^{s-1}+\ldots+A_s,$$

где λ — переменное и $A_0, A_1, \ldots, A_s$ — квадратные матрицы с числовыми элементами одного и того же порядка n. Число n называется *порядком многочлена* $F(\lambda)$. Если $A_0\neq 0$, то число s называется *степенью матричного многочлена* $F(\lambda)$. Если матрица A_0 не вырождена, т. е. $|A_0|\neq 0$, то матричный многочлен $F(\lambda)$ называется *регулярным*. Любую квадратную λ-матрицу можно представить в виде матричного многочлена, и обратно. Например,

$$\begin{pmatrix} 2\lambda^2+\lambda+3 & \lambda-5 \\ 3\lambda+1 & \lambda^2-2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\lambda^2 + \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix}\lambda + \begin{pmatrix} 3 & -5 \\ 1 & -2 \end{pmatrix}.$$

Два матричных многочлена одинакового порядка можно складывать, вычитать и умножать аналогично обычным многочленам с числовыми коэффициентами, с той разницей, что умножение числовых матриц, а потому и матричных многочленов не обязательно коммутативно. Степень суммы или разности двух матричных многочленов не выше большей из их степеней. Степень произведения двух матричных многочленов не выше суммы степеней сомножителей, а если хотя бы один из сомножителей регулярен, то равна этой сумме.

Ввиду некоммутативности умножения для матричных многочленов определяются два деления с остатком — правое и левое. Пусть даны два матричных многочлена

$$F(\lambda) = A_0\lambda^s + A_1\lambda^{s-1} + \ldots + A_s,$$
$$G(\lambda) = B_0\lambda^t + B_1\lambda^{t-1} + \ldots + B_t$$

одинакового порядка n, причем многочлен $G(\lambda)$ регулярен. Тогда существуют однозначно определенные *правое частное* $P(\lambda)$ и *правый остаток* $R(\lambda)$, удовлетворяющие условиям

$$\left.\begin{array}{c} F(\lambda) = P(\lambda)\,G(\lambda) + R(\lambda), \\ \text{причем или степень } R(\lambda) \text{ меньше степени } G(\lambda), \\ \text{или } R(\lambda) = 0, \end{array}\right\} \tag{2.22}$$

а также *левое частное* $Q(\lambda)$ и *левый остаток* $S(\lambda)$, удовлетворяющие условиям

$$\left.\begin{array}{c} F(\lambda) = G(\lambda)\,Q(\lambda) + S(\lambda), \\ \text{причем или степень } S(\lambda) \text{ меньше степени } G(\lambda), \\ \text{или } S(\lambda) = 0 \end{array}\right\} \tag{2.23}$$

([6], стр. 70).

При правом делении делитель стоит справа, а при левом — слева от частного.

Обобщенная теорема Безу. При делении матричного многочлена $F(\lambda)$ на регулярный многочлен $\lambda E - A$ (или на $A - \lambda E$) правый остаток равен *правому значению* делимого $F(\lambda)$ при $\lambda = A$, т. е. матрице

$$F^{(1)}(A) = A_0A^s + A_1A^{s-1} + \ldots + A_s,$$

а лѐвый остаток — *левому значению* делимого $F(\lambda)$ при $\lambda = A$ т. е. матрице $F^{(2)}(A) = A^s A_0 + A^{s-1} A_1 + \ldots + A_s$.

Отсюда следует, что $F(\lambda)$ тогда и только тогда делится справа (слева) на $\lambda E - A$ или на $A - \lambda E$, когда $F^{(1)}(A) = 0$ (соответственно $F^{(2)}(A) = 0$).

7. Характеристические числа и собственные векторы матрицы *). *Характеристической матрицей* для квадратной матрицы $A = (a_{ij})_1^n$ с числовыми элементами называется λ-матрица

$$\lambda E - A = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ldots & -a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ -a_{n1} & -a_{n2} & \ldots & \lambda - a_{nn} \end{pmatrix},$$

где λ — независимое переменное. *Характеристическим многочленом* матрицы A называется многочлен от λ, равный определителю характеристической матрицы, т. е.

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ldots & -a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ -a_{n1} & -a_{n2} & \ldots & \lambda - a_{nn} \end{vmatrix}.$$

Корни характеристического многочлена называются *характеристическими числами* (или *характеристическими корнями*, или *собственными значениями*) матрицы A. Совокупность всех характеристических чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$, где каждое число выписано столько раз, какова его кратность как корня характеристического многочлена, называется *спектром* матрицы A. Уравнение $|\lambda E - A| = 0$ называется *характеристическим* (или *вековым*) *уравнением* матрицы A.

Собственным вектором матрицы A, принадлежащим собственному значению λ_i, называется ненулевой вектор $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, для которого столбец X, составленный из его координат, удовлетворяет матричному уравнению

---

*) Здесь даны понятия характеристических чисел и собственных векторов матрицы. Свойства собственных векторов линейных преобразований (или линейных операторов) можно найти в любом учебнике высшей алгебры (см., например, [6], [11]).

$AX=\lambda_i X$. Это уравнение можно записать в виде $(\lambda_i E - A)X=0$. Отсюда ясно, что координаты вектора a являются решением однородной системы линейных уравнений с матрицей $\lambda_i E - A$, полученной из характеристической матрицы $\lambda E - A$ при $\lambda=\lambda_i$, т. е. системы

$$-a_{j1}x_1 - a_{j2}x_2 - \ldots + (\lambda_i - a_{jj})x_j - \ldots - a_{jn}x_n = 0,$$
$$j=1, 2, \ldots, n.$$

Часто в литературе встречается другое определение характеристической матрицы — как матрицы $A-\lambda E$. В этом случае характеристический многочлен определяется как $|A-\lambda E|$, характеристическое уравнение принимает вид $|A-\lambda E|=0$ и в системе уравнений, служащей для вычисления собственных векторов, все коэффициенты изменяют знак. Но все существенные свойства характеристических чисел и собственных векторов остаются прежними. В частности, не изменяются спектр матрицы A и совокупность собственных векторов, принадлежащих данному характеристическому числу λ_i.

Характеристический многочлен матрицы A порядка n является многочленом n-й степени от λ со старшим коэффициентом, равным единице, и потому всегда отличен от нуля. Если записать его в виде

$$|\lambda E - A| = \lambda^n + c_1\lambda^{n-1} + \ldots + c_k\lambda^{n-k} + \ldots + c_n, \quad (2.24)$$

то его коэффициенты следующим образом выражаются через элементы матрицы A:

$$c_k = (-1)^k S_k, \quad (2.25)$$

где S_k — сумма всех главных миноров (т. е. миноров, симметрично расположенных относительно главной диагонали) порядка k матрицы A. В частности,

$$c_1 = -(a_{11} + a_{22} + \ldots + a_{nn}), \quad c_n = (-1)^n |A|. \quad (2.26)$$

Сумма диагональных элементов $s_A = a_{11} + a_{22} + \ldots + a_{nn}$ называется *следом* матрицы A. Так как по формулам Виета корни характеристического многочлена (2.24) удовлетворяют равенствам

$$\lambda_1 + \lambda_2 + \ldots + \lambda_n = -c_1,$$
$$\lambda_1\lambda_2 \ldots \lambda_n = (-1)^n c_n,$$

то из (2.26) следует

$$\lambda_1+\lambda_2+\ldots+\lambda_n=s_A, \quad \lambda_1\lambda_2\ldots\lambda_n=|A|, \tag{2.27}$$

т. е. сумма характеристических чисел матрицы A равна ее следу, а произведение — определителю. Аналогично из (2.25) следует, что, вообще, значение элементарной симметрической функции σ_k (см. гл. III) от характеристических чисел матрицы A равно сумме главных миноров порядка k матрицы A.

Если под характеристическим многочленом понимать $|A-\lambda E|$, то формулы (2.24) — (2.26) изменятся так:

$$|A-\lambda E|=(-1)^n\lambda^n+c_1\lambda^{n-1}+\ldots+c_k\lambda^{n-k}+\ldots+c_n;$$
$$c_k=(-1)^{n-k}S_k; \quad c_1=(-1)^{n-1}(a_{11}+a_{22}+\ldots+a_{nn}),$$
$$c_n=|A|$$

([8], стр. 108). Формулы (2.27) не изменятся.

Из (2.27) следует, что матрица A тогда и только тогда имеет хотя бы одно характеристическое число, равное нулю, когда она вырождена, $|A|=0$. Квадратная матрица A называется *нильпотентной*, если $A^k=0$ при некотором целом положительном k. Матрица A тогда и только тогда нильпотентна, когда все ее характеристические числа равны нулю.

Теорема Гамильтона — Кэли. Любая квадратная числовая матрица A является корнем своего характеристического многочлена, т. е. аннулирует его при подстановке $\lambda\to A$: если $\varphi(\lambda)=|\lambda E-A|$, то $\varphi(A)=0$ ([14], стр. 26 или [6], стр. 74).

Если $\lambda_1, \lambda_2, \ldots, \lambda_n$ — спектр матрицы A, то для невырожденной матрицы A спектр обратной матрицы A^{-1} образуют числа $\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}$ и, вообще, для любого целого k спектр матрицы A^k имеет вид $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$; если же матрица A вырождена, то последнее утверждение остается верным для целых положительных k. Если $g(\lambda)$ — любой многочлен с числовыми коэффициентами и A — любая квадратная матрица порядка n со спектром $\lambda_1, \lambda_2, \ldots, \lambda_n$, то матрица $g(A)$ имеет спектр $g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_n)$ ([6], стр. 75). Отсюда в силу (2.27) находим

$$|g(A)|=g(\lambda_1)\,g(\lambda_2)\ldots g(\lambda_n)=R(\varphi, g), \tag{2.28}$$

где $R(\varphi, g)$ — результант характеристического многочлена $\varphi(\lambda)=|\lambda E-A|$ и данного многочлена $g(\lambda)$ (гл. III).

Присоединенной матрицей для характеристической матрицы $\lambda E - A$, где $A = (a_{ij})_1^n$, называется матрица $B(\lambda) = (b_{ij}(\lambda))_1^n$, где элемент $b_{ij}(\lambda)$ есть алгебраическое дополнение элемента $a_{ji}(\lambda)$ из j-й строки и i-го столбца матрицы $\lambda E - A$. Матрица $B(\lambda)$ есть матричный многочлен от λ степени $n-1$. Справедливо равенство

$$(\lambda E - A) B(\lambda) = B(\lambda)(\lambda E - A) = |\lambda E - A| E, \tag{2.29}$$

где E — единичная матрица. Отсюда связь присоединенной и обратной матриц:

$$B(\lambda) = |\lambda E - A| \; (\lambda E - A)^{-1}. \tag{2.30}$$

Зная коэффициенты характеристического многочлена (2.24), можно вычислить присоединенную матрицу $B(\lambda)$. Пусть $\varphi(\lambda) = |\lambda E - A|$. Частное $h(\lambda, \mu) = \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu}$ есть многочлен от двух переменных λ и μ. Тогда $B(\lambda) = h(\lambda E, A)$ ([6], стр. 75—76).

Если λ_0 — любое характеристическое число матрицы A и $B(\lambda_0) \neq 0$, то каждый ненулевой столбец матрицы $B(\lambda_0)$ будет собственным вектором матрицы A, принадлежащим значению λ_0.

Д. К. Фаддеев указал метод одновременного вычисления коэффициентов характеристического многочлена и присоединенной матрицы (см. [6], стр. 77).

8. Минимальный многочлен матрицы. *Аннулирующим многочленом* для квадратной матрицы A порядка n называется многочлен $f(\lambda)$ с числовыми коэффициентами, для которого $f(A) = 0$. Нулевой многочлен является аннулирующим для любой матрицы A. Любая матрица A порядка n имеет аннулирующий многочлен n-й степени. Таким многочленом по теореме Гамильтона — Кэли (стр. 94) будет характеристический многочлен матрицы A, т. е. $\varphi(\lambda) = |\lambda E - A|$. Любая матрица A имеет бесконечно много аннулирующих многочленов.

Минимальным многочленом квадратной матрицы A порядка n называется отличный от нуля аннулирующий многочлен для матрицы A минимальной степени, взятый со старшим коэффициентом, равным единице. Для любой матрицы A

порядка n минимальный многочлен существует, определен однозначно и имеет степень не выше n.

Пример 1. Минимальный многочлен нулевой матрицы равен λ, а единичной матрицы E равен $\lambda - 1$.

Многочлен $f(\lambda)$ тогда и только тогда является аннулирующим для матрицы A, когда $f(\lambda)$ делится на минимальный многочлен $\psi(\lambda)$ этой матрицы. В частности, характеристический многочлен $\varphi(\lambda)$ матрицы A всегда делится на минимальный многочлен $\psi(\lambda)$ этой матрицы. Некоторая степень минимального многочлена, например $[\psi(\lambda)]^n$, где n — порядок матрицы A, делится на характеристический многочлен $\varphi(\lambda)$ той же матрицы.

Пример 2. Найти минимальный многочлен инволютивной матрицы A, т. е. матрицы, для которой $A^2 = E$.

В данном случае $A^2 - E = 0$. Поэтому аннулирующим многочленом будет $f(\lambda) = \lambda^2 - 1$. Так как минимальный многочлен делит любой аннулирующий, то возможны случаи:

$$\psi(\lambda) = \lambda - 1, \quad \psi(\lambda) = \lambda + 1, \quad \psi(\lambda) = \lambda^2 - 1.$$

В первом случае $A - E = 0$, $A = E$. Во втором $A + E = 0$, $A = -E$. В третьем $A \neq \pm E$. Итак, если $A \neq \pm E$ и $A^2 = E$, то $\psi(\lambda) = \lambda^2 - 1$.

Если $D_{n-1}(\lambda)$ — наибольший общий делитель миноров $(n-1)$-го порядка характеристической матрицы $\lambda E - A$, а $D_n(\lambda) = |\lambda E - A|$, то минимальный многочлен $\psi(\lambda)$ матрицы A можно найти по формуле

$$\psi(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)} = E_n(\lambda), \tag{2.31}$$

где $E_n(\lambda)$ — последний инвариантный множитель матрицы $\lambda E - A$ (стр. 83) ([6], стр. 81).

Отсюда ясно, что минимальный многочлен тогда и только тогда совпадает с характеристическим многочленом, когда $D_{n-1}(\lambda) = 1$, т. е. когда матрица $\lambda E - A$ имеет лишь один инвариантный множитель, отличный от единицы, именно $E_n(\lambda)$.

Минимальный многочлен клеточно диагональной матрицы

$$A = \{A_1, A_2, \ldots, A_s\}$$

равен наименьшему общему кратному минимальных многочленов диагональных клеток A_i.

Пример 3. Найти минимальный многочлен матрицы

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

Ищем минимальные многочлены диагональных клеток. Для $A_1=(1)$

$$\psi_1(\lambda)=\lambda-1,$$

для $A_2=(2)$

$$\psi_2(\lambda)=\lambda-2,$$

для $A_3=\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ ищем делители миноров: $D_1(\lambda)=1$, $D_2(\lambda)=(\lambda-2)^2$. По формуле (2.31)

$$\psi_3(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)}=(\lambda-2)^2.$$

Минимальный многочлен матрицы A равен наименьшему общему кратному многочленов $\psi_1(\lambda)$, $\psi_2(\lambda)$, $\psi_3(\lambda)$, т. е. $\psi(\lambda)=(\lambda-1)(\lambda-2)^2$.

Приведенная присоединенная матрица. Пусть $D_{n-1}(\lambda)$ — наибольший общий делитель миноров $(n-1)$-го порядка матрицы $\lambda E-A$. Так как все элементы присоединенной матрицы $B(\lambda)$ (стр. 95) отличаются от этих миноров лишь, может быть, знаком, то $D_{n-1}(\lambda)$ будет наибольшим общим делителем всех элементов матрицы $B(\lambda)$. Поэтому

$$B(\lambda)=D_{n-1}(\lambda)C(\lambda), \tag{2.32}$$

где $C(\lambda)$ является λ-матрицей и называется *приведенной присоединенной матрицей* для характеристической матрицы $\lambda E-A$.

Из равенств (2.31), (2.32) и (2.29) на стр. 95 находим

$$(\lambda E-A)C(\lambda)=C(\lambda)(\lambda E-A)=\psi(\lambda)E. \tag{2.33}$$

Отсюда связь матрицы $C(\lambda)$ с обратной матрицей:

$$C(\lambda)=\psi(\lambda)(\lambda E-A)^{-1}. \tag{2.34}$$

Зная минимальный многочлен $\psi(\lambda)$ матрицы A, можно найти приведенную присоединенную матрицу $C(\lambda)$ матрицы $\lambda E-A$. Именно, пусть $g(\lambda, \mu)=\dfrac{\psi(\lambda)-\psi(\mu)}{\lambda-\mu}$ — многочлен от двух переменных λ, μ. Тогда

$$C(\lambda)=g(\lambda E,\ A) \tag{2.35}$$

([6], стр. 81).

Переходя в равенстве (2.33) к определителям, получим

$$|\lambda E-A|\,|C(\lambda)|=[\psi(\lambda)]^n. \tag{2.36}$$

Если λ_0 — любой корень характеристического многочлена матрицы A, то (в отличие от $B(\lambda)$) $C(\lambda_0)\neq 0$ и любой ненулевой столбец матрицы $C(\lambda_0)$ будет собственным вектором матрицы A, принадлежащим значению λ_0 ([6], стр. 81).

9. Подобные матрицы. Квадратная матрица A называется *подобной* матрице B, если существует невырожденная матрица T, для которой

$$B=T^{-1}AT. \tag{2.37}$$

При этом если элементы матриц A и B принадлежат числовому полю K (см. Введение), то преобразующую матрицу T всегда можно взять также с элементами из поля K. Говорят также, что матрица A *трансформируется* в матрицу B при помощи матрицы T.

Отношение подобия обладает тремя основными свойствами: рефлективность: A подобна A; симметрия: если A подобна B, то и B подобна A; транзитивность: если A подобна B и B подобна C, то A подобна C. В силу симметрии можно говорить о подобии двух матриц без указания их последовательности. Из этих свойств вытекает, что множество всех матриц данного порядка с элементами из поля K разбивается на классы подобных матриц так, что две любые матрицы из одного класса подобны, а из разных классов не подобны между собой. Возникает вопрос о канонической форме матрицы, характеризующей данный класс подобных матриц. Этот вопрос рассматривается ниже в пп. 10—12.

Критерий подобия матриц. Для того чтобы две матрицы A и B были подобны, необходимо и достаточно, чтобы их характеристические матрицы $\lambda E-A$ и $\lambda E-B$ (рассматриваемые как λ-матрицы) были эквивалентны (стр. 82) ([14], стр. 123 или [6], стр. 124).

В силу критериев эквивалентности λ-матриц (стр. 87) и свойств элементарных делителей характеристических матриц (стр. 90) для подобия матриц A и B (одинакового порядка) необходимо и достаточно каждое из следующих условий:

1) матрицы A и B имеют одинаковые инвариантные множители;

2) матрицы A и B имеют одинаковые элементарные делители;

3) существуют обратимые λ-матрицы $P(\lambda)$ и $Q(\lambda)$, удовлетворяющие равенству

$$\lambda E - B = P(\lambda)(\lambda E - A)\,Q(\lambda). \tag{2.38}$$

Таким образом, чтобы узнать, будут ли матрицы A и B подобны, достаточно найти инвариантные множители матриц $\lambda E - A$ и $\lambda E - B$ любым из указанных в п. 5 методов.

Более сложным является отыскание преобразующей матрицы T, удовлетворяющей равенству (2.37). Таких матриц бесконечно много. Если T_0 — любая такая матрица, то все преобразующие матрицы найдутся из равенства $T = CT_0$, где C — любая невырожденная матрица, перестановочная с A. Отыскание всех матриц, перестановочных с A, дано на стр. 78.

Первый метод построения преобразующей матрицы. Из (2.37) найдем: $AT = TB$. Приравняв элементы, получим n^2 однородных линейных уравнений для элементов t_{ij} матрицы T. Надо взять любое решение этой системы, для которой матрица T не вырождена. Для этого можно сначала найти общее решение системы, заменить в определителе $|t_{ij}|_1^n$ все главные неизвестные их выражениями через свободные неизвестные. Получится многочлен, не равный нулю тождественно. Поэтому можно найти такие значения свободных неизвестных, для которых этот многочлен отличен от нуля. Этими значениями однозначно определятся все элементы матрицы T.

Пример 1. Доказать, что матрицы

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

подобны, и найти преобразующую матрицу T, для которой $B = T^{-1}AT$. Инвариантные множители характеристических матриц $\lambda E - A$ и $\lambda E - B$ одинаковы:

$$E_1(\lambda) = 1, \quad E_2(\lambda) = \lambda^2 - 3\lambda + 5.$$

Поэтому матрицы A и B подобны. Полагая

$$T = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

и приравнивая элементы матриц в равенстве $AT = TB$, получим

$$\begin{aligned} x-3z &= 4x+3y, \\ x+2z &= 4z+3t, \\ y-3t &= -3x-y, \\ y+2t &= -3z-t. \end{aligned}$$

Решая методом исключения, найдем два линейно независимых уравнения:

$$\begin{aligned} x+y+z &= 0, \\ y+3z+3t &= 0. \end{aligned}$$

Общее решение: $x = 2z+3t$, $y = -3z-3t$. Подставляя эти значения x и y в определитель, найдем

$$|T| = \begin{vmatrix} 2z+3t & -3z-3t \\ z & t \end{vmatrix} = 3z^2+5zt+3t^2.$$

Беря, например, $z=-1$, $t=1$, получим $|T|=1$, $x=1$, $y=0$, откуда

$$T = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Второй метод построения преобразующей матрицы. Если матрицы A и B подобны, то их характеристические матрицы $\lambda E - A$ и $\lambda E - B$ эквивалентны. Поэтому существуют обратимые λ-матрицы $P(\lambda)$ и $Q(\lambda)$, удовлетворяющие равенству (2.38). Метод их построения указан на стр. 87. В данном случае достаточно найти лишь матрицу $Q(\lambda)$. За искомую матрицу T можно взять остаток от деления матрицы $Q(\lambda)$ справа на матрицу $\lambda E - B$ (стр. 91). Если $Q(\lambda) = C_0\lambda^s + C_1\lambda^{s-1} + \ldots + C_s$, то по обобщенной теореме Безу (стр. 91)

$$T = Q^{(1)}(B) = C_0B^s + C_1B^{s-1} + \ldots + C_s.$$

Однако практически удобнее матрицу T искать по схеме, аналогичной схеме Горнера для деления многочлена на двучлен $x-a$ (гл. III). Именно, ищем матрицы D_i, определенные рекуррентным соотношением:

$$\left.\begin{aligned} D_0 &= C_0, \\ D_i &= C_i + D_{i-1}B, \qquad i = 1, 2, \ldots, s, \\ T &= D_s. \end{aligned}\right\} \qquad (2.39)$$

Из критерия подобия следует, что если матрицы A и B подобны, то равны их характеристические и минимальные многочлены. Обратное утверждение неверно.

Каждая матрица A подобна своей транспонированной матрице A', так как матрицы $\lambda E - A$ и $\lambda E - A'$ имеют одинаковые делители миноров и потому эквивалентны.

Если хотя бы одна из двух матриц A и B, например B, не вырождена, то матрицы AB и BA подобны, так как $AB = B^{-1}(BA)B$. Это утверждение неверно, если обе матрицы A и B вырождены.

Если матрицы A и B подобны и $f(\lambda)$ — любой многочлен, то и матрицы $f(A)$ и $f(B)$ подобны, и притом с той же преобразующей матрицей, т. е. если $B = T^{-1}AT$, то $f(B) = T^{-1}f(A)T$.

Матрица A подобна матрице B, полученной из A любой перестановкой строк и такой же перестановкой столбцов. При этом $B = T^{-1}AT$, где T получается применением той же перестановки к столбцам единичной матрицы.

Любая числовая матрица A подобна над полем комплексных чисел треугольной (по желанию, верхней или нижней) матрице. Вообще, матрица A, все элементы и характеристические числа которой лежат в числовом поле K, подобна над K треугольной матрице (п. 11).

10. Диагональная форма матрицы. Каноническая форма квадратной матрицы A порядка n, т. е. матрица B специального вида, характеризующая класс матриц, подобных данной матрице A, определяется неоднозначно, если ставить задачу отыскания канонической формы B возможно более простого вида.

Наиболее простым и практически важным является случай, когда данная матрица A подобна диагональной матрице $B = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ (стр. 89). При этом числа λ_i, стоящие на диагонали матрицы B, являются характеристическими числами матрицы A с учетом их кратности. Однако, не любая матрица A подобна диагональной матрице. Именно, имеет место

Первый критерий существования диагональной формы. Для того чтобы матрица A с элементами из числового поля K (см. введение) была подобна над K диагональной матрице B, т. е. чтобы существовала невырожденная матрица T с элементами из K, удовлетворяющая равенству

$$B = T^{-1}AT, \tag{2.40}$$

необходимо и достаточно выполнение двух условий: а) последний инвариантный множитель $E_n(\lambda)$ (стр. 83) или, что то же самое, минимальный многочлен $\psi(\lambda)$ (стр. 96, (2.31)) матрицы A не имеет кратных корней; б) все характеристические числа матрицы A принадлежат полю K.

Условия а) и б) можно заменить эквивалентным требованием: все элементарные делители матрицы A над полем K должны быть многочленами первой степени ([14], стр. 127).

Если K — поле комплексных чисел, то условие б) всегда выполняется.

Более простым является

Достаточное условие существования диагональной формы. Если все характеристические числа матрицы A с элементами из числового поля K также лежат в поле K и характеристический многочлен $|\lambda E - A|$ не имеет кратных корней, то матрица A подобна над K диагональной матрице B. При этом на диагонали матрицы B должны стоять характеристические числа матрицы A.

Любую числовую матрицу сколь угодно малым изменением элементов можно превратить в матрицу, подобную диагональной матрице над полем комплексных чисел.

Второй критерий существования диагональной формы. Для того чтобы матрица A с элементами из числового поля K была подобна над K диагональной матрице B, необходимо и достаточно, чтобы все характеристические числа матрицы A лежали в поле K и чтобы матрица A имела n линейно независимых собственных векторов.

Правило построения преобразующей матрицы T, приводящей матрицу A к диагональной форме B. 1) Если характеристический многочлен не имеет кратных корней, то, беря для каждого корня λ_i любой собственный вектор e_i, $i = 1, 2, \ldots, n$, получим линейно независимую систему векторов (столбцов). Невырожденная квадратная матрица T, образованная из этих столбцов, и будет искомой преобразующей матрицей.

2) Если характеристический многочлен имеет кратные корни, то для каждого корня λ_i кратности k_i ищем фундаментальную систему решений (стр. 63) однородной системы линейных уравнений

$$a_{j1}x_1 + \ldots + (a_{jj} - \lambda_i)x_j + \ldots + a_{jn}x_n = 0, \quad j = 1, 2, \ldots, n. \tag{2.41}$$

Если для каждого корня λ_i эта фундаментальная система содержит ровно k_i решений, то, ставя координаты решений всех этих фундаментальных систем по столбцам матрицы T, мы получим преобразующую матрицу.

3) Если хотя бы для одного корня λ_i фундаментальная система решений для системы уравнений (2.41) содержит менее чем k_i решений (можно доказать, что более k_i решений она содержать не может), то матрица A не будет подобна диагональной матрице, и преобразующей матрицы T не существует.

Пример 1. Показать, что матрица

$$A=\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$$

подобна диагональной матрице B, и найти эту матрицу B и преобразующую матрицу T.

Вычисляем характеристический многочлен:

$$\begin{vmatrix} \lambda-7 & 12 & -6 \\ -10 & \lambda+19 & -10 \\ -12 & 24 & \lambda-13 \end{vmatrix}=$$

(ко второму столбцу прибавляем удвоенный первый)

$$=\begin{vmatrix} \lambda-7 & 2\lambda-2 & -6 \\ -10 & \lambda-1 & -10 \\ -12 & 0 & \lambda-13 \end{vmatrix}=$$

(из второго столбца выносим $\lambda-1$ и из первой строки вычитаем удвоенную вторую)

$$=(\lambda-1)\begin{vmatrix} \lambda+13 & 0 & 14 \\ -10 & 1 & -10 \\ -12 & 0 & \lambda-13 \end{vmatrix}=$$

(разлагаем по второму столбцу)

$$=(\lambda-1)\begin{vmatrix} \lambda+13 & 14 \\ -12 & \lambda-13 \end{vmatrix}=$$

(к первой строке прибавляем вторую, затем из второго столбца вычитаем первый)

$$=(\lambda-1)\begin{vmatrix} \lambda+1 & 0 \\ -12 & \lambda-1 \end{vmatrix}=(\lambda-1)^2(\lambda+1);$$

$$\lambda_1=\lambda_2=1, \quad \lambda_3=-1.$$

Для корня 1 система (2.41) принимает вид

$$\begin{aligned} 6x_1-12x_2+\ 6x_3&=0,\\ 10x_1-20x_2+10x_3&=0,\\ 12x_1-24x_2+12x_3&=0. \end{aligned}$$

Ее ранг равен единице и фундаментальная система состоит из двух решений, например:

$$e_1=(2,\ 1,\ 0),\quad e_2=(1,\ 0,\ -1).$$

Для корня -1 получаем систему

$$\begin{aligned} 8x_1-12x_2+\ 6x_3&=0,\\ 10x_1-18x_2+10x_3&=0,\\ 12x_1-24x_2+14x_3&=0 \end{aligned}$$

ранга 2. Фундаментальная система состоит из одного решения, например: $e_3=(3,\ 5,\ 6)$. Так как в обоих случаях число решений равно кратности корня, то матрица A подобна диагональной матрице

$$B=\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1 \end{pmatrix}.$$

Записывая координаты векторов e_1, e_2, e_3 по столбцам, получим преобразующую матрицу

$$T=\begin{pmatrix} 2 & 1 & 3\\ 1 & 0 & 5\\ 0 & -1 & 6 \end{pmatrix}.$$

Диагональная матрица B для данной матрицы A определена однозначно с точностью до порядка диагональных элементов. Матрица T определена неоднозначно. Так, в предыдущем примере за e_1 и e_2 можно принять любую фундаментальную систему решений соответствующей системы уравнений. Порядок столбцов в матрице T соответствует порядку диагональных элементов матрицы B.

Пример 2. Выяснить, подобна ли диагональной матрице матрица

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

Вычисляя характеристический многочлен, например, разложением по третьей строке, найдем

$$|\lambda E-A|=(\lambda+1)^3,\quad \lambda_1=\lambda_2=\lambda_3=-1.$$

Собственные векторы ищем из системы уравнений

$$\begin{aligned} 3x_1 - x_2 + 2x_3 &= 0, \\ 5x_1 - 2x_2 + 3x_3 &= 0, \\ -x_1 \qquad - x_3 &= 0. \end{aligned}$$

Ее ранг равен двум. Фундаментальная система состоит из одного решения, например: $e = (1, 1, -1)$. Так как трехкратному корню -1 соответствует один линейно независимый собственный вектор, то матрица A не будет подобна диагональной матрице.

Если действительная матрица A порядка n подобна над полем комплексных чисел диагональной матрице

$$B = \{\lambda_1, \lambda_2, \ldots, \lambda_n\},$$

не все диагональные элементы которой действительны, то она над полем действительных чисел подобна действительной клеточно диагональной матрице B с клетками первого и второго порядка. Именно если

$$\lambda_1 = a_1 + ib_1, \quad \lambda_2 = a_1 - ib_1, \quad \lambda_3 = a_2 + ib_2,$$
$$\lambda_4 = a_2 - ib_2, \ \ldots, \ \lambda_{2q-1} = a_q + ib_q, \quad \lambda_{2q} = a_q - ib_q,$$
$$\lambda_{2q+1} = c_{2q+1}, \ \ldots, \ \lambda_n = c_n,$$

где числа a_k, b_k, c_k действительны и b_k положительны, то

$$B = \left\{ \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}, \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix}, \ldots, \begin{pmatrix} a_q & -b_q \\ b_q & a_q \end{pmatrix}, c_{2q+1}, \ldots, c_n \right\}.$$

Таким образом, клетки первого порядка соответствуют действительным корням, а клетки второго порядка — парам сопряженных комплексных корней характеристического многочлена матрицы A. Преобразующая матрица T строится следующим образом: систему собственных векторов матрицы A можно выбрать так, что если для комплексного корня $\lambda_{2k-1} = a_k + ib_k$ взят вектор $z_{2k-1} = x_k + iy_k$, где векторы x_k, y_k имеют действительные координаты, то для $\lambda_{2k} = a_k - ib_k$ берется вектор

$$z_{2k} = x_k - iy_k, \qquad k = 1, 2, \ldots, q.$$

Для действительного λ_k берется вектор z_k с действительными координатами. Тогда по столбцам матрицы T должны стоять координаты векторов

$$y_1, x_1, y_2, x_2, \ldots, y_q, x_q, z_{2q+1}, \ldots, z_n$$

([6], стр. 231—232).

11. Жорданова форма матрицы. Так как по предыдущему пункту не всякая матрица подобна диагональной, то рассматривают другие канонические формы матриц. Одной из простейших канонических форм, существующей для любой матрицы, является жорданова форма.

Клеткой Жордана называется квадратная матрица вида

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{pmatrix}, \tag{2.42}$$

где на главной диагонали стоит одно и то же число λ_0 и все элементы одного ряда выше диагонали равны единице. Иногда ряд единиц пишут не выше, а ниже диагонали, что не вносит существенных изменений в теорию.

Жордановой матрицей называется клеточно диагональная матрица

$$B = \{B_1, B_2, \dots, B_s\}, \tag{2.43}$$

в которой на главной диагонали стоят любые клетки Жордана $B_1, B_2, \dots, B_s$, а все элементы вне этих клеток равны нулю.

Пример 1. Матрица

$$B = \left(\begin{array}{cc:cc:ccc:c} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

является жордановой. Она содержит четыре клетки Жордана: две клетки второго порядка с числом 2 на диагонали, одну клетку третьего порядка с числом 0 на диагонали и одну клетку первого порядка с числом 0 на диагонали.

Любая квадратная матрица A, все элементы и характеристические числа которой лежат в числовом поле K, подобна над K жордановой матрице B. В частности, любая числовая матрица подобна над полем комплексных чисел жордановой матрице. Вообще, так как любое поле K содержится в алгебраически замкнутом поле $\overline{K}$ (гл. IV), то любая квадратная матрица A с элементами из K подобна над $\overline{K}$ жордановой матрице B. Для данной матрицы А подобная ей жорданова матрица B определена однозначно с точностью до порядка расположения клеток Жордана $B_1, B_2, \ldots, B_s$ на главной диагонали. Эта матрица B называется *жордановой формой* матрицы A. Таким образом, каждый класс подобных матриц с элементами из алгебраически замкнутого поля содержит одну или конечное число жордановых матриц, получаемых друг из друга перестановкой диагональных клеток. Две жордановы матрицы с элементами из поля K тогда и только тогда подобны между собой, когда они различаются лишь порядком следования клеток Жордана ([14], стр. 124, теорема 4).

Пример 2. Матрицы

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

подобны, как жордановы матрицы, состоящие из одних и тех же трех клеток Жордана и различающиеся лишь расположением этих клеток. Эти матрицы не подобны матрице

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

содержащей две клетки Жордана.

Правило построения жордановой формы матрицы. Если все элементы и характеристические числа матрицы A лежат в числовом поле K, то над этим полем элементарные делители матрицы A или, что то же самое, ее характеристической матрицы $\lambda E - A$ будут степенями линейных множителей, т. е. будут иметь вид

$$(\lambda-\lambda_1)^{k_1},\quad (\lambda-\lambda_2)^{k_2},\ \ldots,\ (\lambda-\lambda_s)^{k_s}.$$

Для каждого элементарного делителя $(\lambda-\lambda_i)^{k_i}$ строим клетку Жордана

$$B_i=\begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_i \end{pmatrix}$$

порядка k_i с числом λ_i на диагонали. Клеточно диагональная матрица $B=\{B_1, B_2, \ldots, B_s\}$ с этими клетками на диагонали и будет жордановой формой матрицы A.

Пример 3. Найти жорданову форму матрицы

$$A=\begin{pmatrix} 4 & -5 & 2 & 0 & 0 & 0 \\ 5 & -7 & 3 & 0 & 0 & 0 \\ 6 & -9 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -4 & 4 & 0 \\ 0 & 0 & 0 & -2 & 1 & 2 \end{pmatrix}.$$

Ищем элементарные делители этой матрицы. Так как она клеточно диагональна, то можно искать элементарные делители каждой из двух клеток (стр. 90). Для клетки

$$A_1=\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$$

делители миноров характеристической матрицы $\lambda E - A_1$ будут

$$D_1(\lambda)=D_2(\lambda)=1,\quad D_3(\lambda)=\lambda^2(\lambda-1).$$

Их можно найти тем же методом, как в примере 2 на стр. 85. Поэтому инвариантные множители равны 1, 1, $\lambda^2(\lambda-1)$ и элементарные делители равны λ^2, $\lambda-1$. Также для второй клетки

$$A_2=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

находим делители миноров: $D_1(\lambda)=1$, $D_2(\lambda)=\lambda-2$, $D_3(\lambda)=(\lambda-2)^3$. Инвариантные множители: 1, $\lambda-2$, $(\lambda-2)^2$; элементарные делители: $\lambda-2$, $(\lambda-2)^2$. Поэтому жорданова форма матрицы A содержит четыре клетки и имеет вид

$$B=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Матрицу T, преобразующую данную матрицу A к жордановой форме B, можно найти по общим методам, указанным в п. 9. Однако в данном случае существуют и более экономные пути построения матрицы T, связанные со свойствами линейных операторов и выходящие за рамки настоящей главы (см., например, [6], гл. 6, § 9 или [20], задача 1529).

Вычисление значения многочлена $f(x)$ от жордановой матрицы $B=\{B_1, B_2, \ldots, B_s\}$ производится по формуле $f(B)=\{f(B_1), f(B_2), \ldots, f(B_s)\}$, т. е. сводится к вычислению значений этого многочлена от жордановых клеток. Значение многочлена $f(x)$ от жордановой клетки

$$B_0=\begin{pmatrix} \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_0 \end{pmatrix}$$

порядка n вычисляется по формуле

$$f(B_0)=\begin{pmatrix} f(\lambda_0) & f'(\lambda_0) & \frac{f''(\lambda_0)}{2!} & \ldots & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & f(\lambda_0) & f'(\lambda_0) & \ldots & \frac{f^{(n-2)}(\lambda_0)}{(n-2)!} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & f(\lambda_0) \end{pmatrix}$$

([14], стр. 134, (3)). По той же формуле вычисляется значение любой функции $f(x)$ от жордановой клетки B_0, если только значения

$$f(\lambda_0),\ f'(\lambda_0),\ \ldots,\ f^{(n-1)}(\lambda_0)$$

имеют смысл. В частности, степени жордановой клетки B_0 вычисляются по формуле

$$B_0^k=\begin{pmatrix} \lambda_0^k & k\lambda_0^{k-1} & C_k^2\lambda_0^{k-2} & \ldots & C_k^{n-1}\lambda_0^{k-n+1} \\ 0 & \lambda_0^k & k\lambda_0^{k-1} & \ldots & C_k^{n-2}\lambda_0^{n-k+2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_0^k \end{pmatrix}.$$

Чтобы найти жорданову форму матрицы $f(A)$, где A — данная квадратная матрица и $f(x)$ — многочлен (или любая функция, для которой значение $f(A)$ имеет смысл), достаточно, по данному выше правилу построения жордановой формы, знать элементарные делители матрицы $f(A)$. Об их вычислении см. [6], гл. 6, § 7.

12. Естественная форма и другие канонические формы матрицы. Жордановы матрицы, определенные в предыдущем пункте, не находятся во взаимно однозначном соответствии с классами подобных матриц над данным числовым полем K, лишь с точностью до перестановки клеток Жордана. Укажем каноническую форму, свободную от этого недостатка, хотя и не столь детализированную, как жорданова форма.

Пусть дан многочлен

$$f(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\ldots+a_{n-1}\lambda^{n-1}+\lambda^n$$

степени $n>0$ с коэффициентами из поля K со старшим коэффициентом, равным единице. *Сопровождающей матрицей* многочлена $f(\lambda)$ называется матрица

$$B=\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

Эта матрица является квадратной порядка n, равного степени многочлена $f(\lambda)$. Последний инвариантный множитель матрицы $\lambda E-B$ равен $f(\lambda)$, а все остальные — единице.

Матрицей *естественной формы* (или матрицей, имеющей *естественную нормальную форму*) называется клеточно диагональная матрица

$$B=\{B_1, B_2, \ldots, B_s\},$$

где диагональные клетки $B_1, B_2, \ldots, B_s$ являются сопровождающими матрицами многочленов $f_1(\lambda), f_2(\lambda), \ldots, f_s(\lambda)$, каждый из которых (начиная со второго) делится на предыдущий. Эта матрица является квадратной порядка n, равного сумме степеней многочленов $f_i(\lambda)$. Последние s инвариантных множителей матрицы $\lambda E-B$ равны соответственно $f_1(\lambda)$, $f_2(\lambda), \ldots, f_s(\lambda)$, а остальные $n-s$ равны единице.

Любая квадратная матрица A с элементами из некоторого числового поля K подобна над K единственной матрице B естественной формы ([14], стр. 128, теорема 5). Две матрицы естественной формы тогда и только тогда подобны между собой, когда они совпадают. Таким образом, каждый класс подобных матриц с элементами из поля K содержит единственную матрицу естественной формы.

Правило построения естественной формы. Пусть A — квадратная матрица порядка n с элементами из поля K. Одним из методов, указанных в п. 5, строим инвариантные множители ее характеристической матрицы $\lambda E-A$. Пусть

$$f_1(\lambda), f_2(\lambda), \ldots, f_s(\lambda)$$

— все многочлены положительной степени среди них, взятые с сохранением их порядка (так что каждый из них делится на предыдущий). Для каждого из этих многочленов $f_i(\lambda)$ строим сопровождающую матрицу B_i. Заметим, что для многочленов первой степени $\lambda-a$ сопровождающей матрицей будет матрица первого порядка (a). Клеточно диагональная матрица $B=\{B_1, B_2, \ldots, B_s\}$ будет естественной формой матрицы A.

Пример 1. Найдем естественную форму матрицы A из примера 3 предыдущего пункта (стр. 108). Там были найдены элементарные делители матрицы $\lambda E-A$:

$$\lambda^2, \quad \lambda-1, \quad (\lambda-2)^2, \quad \lambda-2.$$

Поэтому инвариантные множители будут

$$E_6(\lambda)=\lambda^2(\lambda-1)(\lambda-2)^2=\lambda^5-5\lambda^4+8\lambda^3-4\lambda^2,$$
$$E_5(\lambda)=\lambda-2.$$

Откуда

$$B=\left(\begin{array}{c|ccccc}2 & 0 & 0 & 0 & 0 & 0\\ \hline 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 4 & -8 & 5\end{array}\right).$$

Однако, та же матрица A приводится к жордановой форме с рациональными элементами, как и сама матрица A.

Пример 2. Найти естественную форму матрицы

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

Делители миноров матрицы

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

будут $D_1(\lambda)=D_2(\lambda)=1$, так как минор второго порядка в правом верхнем углу равен единице, и $D_3(\lambda)=(\lambda-2)(\lambda^2+1)$. Инвариантные множители:

$$E_1(\lambda)=E_2(\lambda)=1,\quad E_3(\lambda)=\lambda^3-2\lambda^2+\lambda-2.$$

Так как характеристический многочлен не имеет кратных корней, то матрица A подобна диагональной матрице

$$\begin{pmatrix}2 & 0 & 0\\ 0 & i & 0\\ 0 & 0 & -i\end{pmatrix}$$

над полем комплексных чисел. В то же время матрица A уже над полем рациональных чисел подобна матрице естественной формы

$$B=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 2 & -1 & 2\end{pmatrix}.$$

Квазиестественная форма. Естественная форма не всегда является клеточно диагональной матрицей с наименьшими порядками диагональных клеток среди матриц, подобных данной матрице A над числовым полем K. Это видно хотя бы из примера 1. Укажем форму, удовлетворяющую условию минимальности порядков диагональных клеток для данного поля K.

Квазиестественной матрицей называется клеточно диагональная матрица $B=\{B_1, B_2, \ldots, B_s\}$, где диагональные клетки являются сопровождающими матрицами многочленов вида

$$[\varepsilon_1(\lambda)]^{p_1}, \quad [\varepsilon_2(\lambda)]^{p_2}, \ldots, [\varepsilon_s(\lambda)]^{p_s},$$

причем многочлены $\varepsilon_1(\lambda), \varepsilon_2(\lambda), \ldots, \varepsilon_s(\lambda)$ неприводимы над полем K и имеют старшие коэффициенты, равные единице.

Любая квадратная матрица A с элементами из некоторого числового поля K подобна над K квазиестественной матрице B, определенной однозначно с точностью до порядка следования диагональных клеток. Эту матрицу B будем называть *квазиестественной формой* матрицы A. Две квазиестественные матрицы с элементами из поля K тогда и только тогда подобны между собой, когда они отличаются не более, чем порядком расположения диагональных клеток ([14], стр. 129).

Правило построения квазиестественной формы. Пусть дана квадратная матрица A с элементами из поля K. Ищем элементарные делители над K характеристической матрицы $\lambda E-A$ и для каждого из них строим сопровождающую матрицу. Эти матрицы в любом порядке принимаем за диагональные клетки клеточно диагональной матрицы B.

Пример 3. Найдем квазиестественную форму матрицы A из примера 3 предыдущего пункта. Элементарные делители матрицы $\lambda E-A$ равны λ^2, $\lambda-1$, $(\lambda-2)^2$, $\lambda-2$. Поэтому искомая матрица имеет вид

$$B=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -4 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Эта матрица не является жордановой.

Обобщенная форма Жордана. Примеры 1—3 показывают, что естественная и квазиестественная формы не являются обобщением жордановой формы. Укажем такое обобщение, которое можно всегда осуществить, не выходя за пределы данного числового поля K (и вообще любого поля K характеристики нуль (гл. IV)).

Обобщенной клеткой Жордана для многочлена вида $[\varepsilon(\lambda)]^p$, где $\varepsilon(\lambda)$ — неприводимый над числовым полем K многочлен степени $k>0$ со старшим коэффициентом, равным единице, и p — целое положительное число, называется клеточная матрица

$$B_0=\begin{pmatrix} A_0 & E & 0 & \dots & 0 \\ 0 & A_0 & E & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & E \\ 0 & 0 & 0 & \dots & A_0 \end{pmatrix},$$

где A_0 — сопровождающая матрица многочлена $\varepsilon(\lambda)$ (в смысле, указанном в начале этого пункта) и E — единичная матрица порядка k. Матрица B_0 состоит из p^2 квадратных клеток порядка k и имеет порядок, равный kp. Ее характеристическая матрица $\lambda E-B_0$ имеет над полем K единственный элементарный делитель $[\varepsilon(\lambda)]^p$ ([14], стр. 129, лемма 4).

Обобщенной матрицей Жордана называется клеточно диагональная матрица $B=\{B_1, B_2, \dots, B_s\}$, где все диагональные клетки являются обобщенными клетками Жордана. Любая квадратная матрица A с элементами из числового поля K (или любого поля характеристики нуль) подобна над K обобщенной матрице Жордана B. Матрица B определена однозначно с точностью до перестановки обобщенных клеток Жордана. Она тогда и только тогда превращается в жорданову матрицу, когда все характеристические числа матрицы A лежат в поле K ([14], стр. 131).

Заметим, что элементарный делитель $[\varepsilon(\lambda)]^p$ обобщенной клетки Жордана B_0 не изменится, если за матрицу A_0 взять любую матрицу порядка k с элементарным делителем $\varepsilon(\lambda)$. Это позволяет находить матрицы другой формы, подобные данной матрице A. Так, если K — поле действительных чисел, то неприводимыми над K будут лишь многочлены первой степени и многочлены второй степени $\lambda^2+p\lambda+q$ с комплексными корнями $\alpha\pm i\beta$, $\beta\neq 0$. Для элементарного делителя $(\lambda^2+p\lambda+q)^m$ в матрице B_0. в качестве A_0 можно взять либо сопровождающую матрицу $\begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix}$, либо матрицу $C_0=\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$. В самом деле,

$$|\lambda E-C_0|=\begin{vmatrix} \lambda-\alpha & -\beta \\ \beta & \lambda-\alpha \end{vmatrix}=\lambda^2-2\alpha\lambda+\alpha^2+\beta^2=$$
$$=\lambda^2+p\lambda+q.$$

Пример 4. Найти обобщенную жорданову форму над полем действительных чисел для матрицы

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

Характеристическая матрица:

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

Разлагая ее определитель по теореме Лапласа по первым двум столбцам, найдем делители миноров: $D_4(\lambda)=(\lambda^2-4\lambda+5)^2$, $D_3(\lambda)=1$, так как минор третьего порядка в правом верхнем углу равен единице. Поэтому матрица $\lambda E-A$ над полем действительных чисел имеет единственный элементарный делитель $(\lambda^2-4\lambda+5)^2$. Его корни: $2\pm i$, оба двойные.

Обобщенная форма Жордана состоит из одной клетки

$$B=\begin{pmatrix} 0 & 1 & 1 & 0 \\ -5 & 4 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -5 & 4 \end{pmatrix}.$$

Беря вместо сопровождающей матрицы $\begin{pmatrix} 0 & 1 \\ -5 & 4 \end{pmatrix}$ матрицу $C_0=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ согласно сделанному выше замечанию, получим другую форму матрицы, подобной матрице A, именно:

$$C=\begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Матрица C (в отличие от B) не является обобщенной жордановой формой матрицы A.

13. Якобиевы матрицы. О вычислении якобиевых определителей см. гл. I, § 1, п. 8. Доказательства приведенных ниже свойств якобиевых матриц см. в [7], гл. 2, § 1.

Якобиевой матрицей называется матрица, у которой все элементы, стоящие вне главной диагонали и двух соседних с ней параллелей, равны нулю. Такую матрицу можно записать в виде

$$J_n = \begin{pmatrix} a_1 & -b_1 & 0 & 0 & \dots & 0 \\ -c_1 & a_2 & -b_2 & 0 & \dots & 0 \\ 0 & -c_2 & a_3 & -b_3 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & a_n \end{pmatrix}. \qquad (2.44)$$

Мы будем рассматривать лишь матрицы с действительными элементами, удовлетворяющими условиям

$$b_k c_k > 0, \quad k = 1, 2, \dots, n-1. \qquad (2.45)$$

Положим

$$D_0(\lambda) = 1, \quad D_k(\lambda) = |J_k - \lambda E|, \quad k = 1, 2, \dots, n. \qquad (2.46)$$

Свойства многочленов $D_k(\lambda)$ (некоторые из этих свойств сходны со свойствами ряда Штурма) (см. гл. III):

1) Если $D_k(\alpha) = 0$, $1 \leqslant k \leqslant n-1$, то $D_{k-1}(\alpha)$ и $D_{k+1}(\alpha)$ отличны от нуля и имеют противоположные знаки.

2) При переходе действительного переменного λ через корень многочлена $D_k(\lambda)$ слева направо произведение $D_k(\lambda)$ $D_{k-1}(\lambda)$, где $k > 0$, меняет знак с плюса на минус.

3) Все корни многочлена $D_k(\lambda)$, $k > 0$, действительны и различны.

4) Если a и b, $a < b$, не являются корнями многочлена $D_k(\lambda)$, то число его корней на отрезке $[a, b]$ равно приращению числа перемен знака в ряду $D_0(\lambda)$, $D_1(\lambda), \dots, D_k(\lambda)$ при переходе λ от a к b.

5) Между двумя соседними корнями многочлена $D_k(\lambda)$, где $k > 0$, лежит точно один корень многочлена $D_{k-1}(\lambda)$.

6) Если $\lambda_1 < \lambda_2 < \dots < \lambda_n$ — корни многочлена $D_n(\lambda)$, то при $\lambda = \lambda_j$ ряд $D_0(\lambda)$, $D_1(\lambda), \dots, D_{n-1}(\lambda)$ содержит точно $j-1$ перемен знака.

7) Если $u_j = (u_{1j}, u_{2j}, \dots, u_{nj})$ — собственный вектор, соответствующий j-му по величине собственному значению λ_j

якобиевой матрицы (2.44) с условиями (2.45), то

$$u_{kj} = \frac{C_j D_{k-1}(\lambda_j)}{b_1 b_2 \ldots b_{k-1}}, \tag{2.47}$$

где C_j — произвольное число, отличное от нуля, $k = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$.

Якобиева матрица называется *нормальной*, если ее элементы действительны и удовлетворяют условиям: $b_k > 0$, $c_k > 0$; $k = 1, 2, \ldots, n-1$.

Так как условия (2.45) для нормальной якобиевой матрицы выполнены, то она обладает свойствами 1) — 7). Отсюда следует:

8) Собственные значения λ_i нормальной якобиевой матрицы действительны и различны. Если $\lambda_1 < \lambda_2 < \ldots < \lambda_n$, то в ряду координат собственного вектора, соответствующего числу λ_j, имеется точно $j-1$ перемен знака.

Пусть $u = (u_1, u_2, \ldots, u_n)$ — любой n-мерный вектор с действительными координатами. Плоская ломаная с вершинами $P_1, P_2, \ldots, P_n$, где P_k имеет прямоугольные координаты $x_k = k$, $y_k = u_k$, $k = 1, 2, \ldots, n$, *называется* u*-линией*. Точки пересечения u-линии с осью Ox называются *узлами* u*-линии* или u*-вектора*.

Рассмотрим вектор

$$u(\lambda) = (u_1(\lambda), u_2(\lambda), \ldots, u_n(\lambda)),$$

координаты которого определяются равенствами

$$u_1(\lambda) \equiv 1, \quad u_k(\lambda) = \frac{D_{k-1}(\lambda)}{b_1 b_2 \ldots b_{k-1}}, \quad k = 2, 3, \ldots, n. \tag{2.48}$$

Свойства узлов $u(\lambda)$-линии:

9) Каждая точка $u(\lambda)$-линии, общая с осью Ox, при $1 < x < n$ является ее узлом (вытекает из 1)).

10) $u(\lambda)$-линия при $\lambda = \lambda_j$ имеет точно $j-1$ узлов (вытекает из 6)).

11) Число узлов $u(\lambda)$-линии равно числу корней многочлена $D_{n-1}(\lambda)$ на интервале $(-\infty, \lambda)$ (вытекает из 4)).

12) Если $\lambda < \mu$, то между каждыми двумя узлами $u(\lambda)$-линии лежит по крайней мере один узел $u(\mu)$-линии.

13) При непрерывном возрастании λ узлы $u(\lambda)$-линии непрерывно смещаются влево.

14) Узлы двух собственных векторов нормальной якобиевой матрицы, принадлежащих двум соседним собственным значениям, перемежаются.

14. Ассоциированные и взаимные матрицы. Об ассоциированных и взаимных определителях см. гл. I, § 1, п. 13. Пусть $A=(a_{ij})_1^n$ — матрица порядка $n\geqslant 2$ и p — целое число, причем $1\leqslant p\leqslant n$. Обозначим через $N \doteq C_n^p$ число сочетаний $i_1<i_2<\ldots<i_p$ из n чисел $1, 2, \ldots, n$ по p. Занумеруем все эти сочетания в произвольном, но в дальнейшем неизменном порядке:

$$s_1, s_2, \ldots, s_N. \tag{2.49}$$

Если s_i — сочетание $i_1<i_2<\ldots<i_p$ и s_j — сочетание $j_1<<j_2<\ldots<j_p$, то минор $A\begin{pmatrix} i_1 i_2 \ldots i_p \\ j_1 j_2 \ldots j_p\end{pmatrix}$, стоящий на пересечении p строк с номерами $i_1, i_2, \ldots, i_p$ и p столбцов с номерами $j_1, j_2, \ldots, j_p$, обозначим через $\tilde{a}_{ij}$.

Для матрицы A порядка n *p-й ассоциированной* матрицей называется матрица порядка $N=C_n^p$, составленная из миноров порядка p матрицы A, взятых в определенном порядке:

$$\tilde{A}_p \doteq (\tilde{a}_{ij})_1^N.$$

Например, $\tilde{A}_1=A$, $\tilde{A}_n=(|A|)$.

Свойства ассоциированных матриц:

1) $(\widetilde{AB})_p=\widetilde{A}_p\cdot\widetilde{B}_p$.

2) $(\widetilde{E}_n)_p=E_N$, где E_n и E_N — единичные матрицы соответственно порядков n и N.

3) Если $|A|\neq 0$, то $(\widetilde{A^{-1}})_p=(\tilde{A}_p)^{-1}$ ([6], стр. 26).

4) $(\widetilde{A^k})_p=(\widetilde{A}_p)^k$, причем для невырожденной матрицы A это верно при любом целом k, а для вырожденной — при целом неотрицательном k.

5) Если матрицы A и B подобны (стр. 98), то и p-е ассоциированные с ними матрицы подобны, причем если $B==T^{-1}AT$, то $\tilde{B}_p=(\tilde{T}_p)^{-1}\cdot\tilde{A}_p\cdot\tilde{T}_p$.

6) Если $A=\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ — диагональная матрица, то и p-я ассоциированная матрица также диагональна, причем $\tilde{A}_p==\{\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_N\}$, где на главной диагонали матрицы $\tilde{A}_p$ стоят всевозможные произведения по p диагональных элементов матрицы A в следующем порядке: если s_i — сочетание $i_1<i_2<\ldots<i_p$, то $\tilde{\lambda}_i=\lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_p}$.

7) При изменении нумерации сочетаний в последовательности (2.49) в p-й ассоциированной матрице $\tilde{A}_p$ происходит

одинаковая перестановка строк и столбцов, т. е. матрица $\tilde{A}_p$ заменяется подобной матрицей (стр. 101), а характеристические числа этой матрицы не изменяются. Поэтому при вычислении характеристических чисел p-й ассоциированной матрицы можно выбрать наиболее удобный порядок нумерации в последовательности (2.49).

8) Если матрица A является треугольной и сочетания в последовательности (2.49) занумерованы в лексикографическом порядке, то и p-я ассоциированная матрица $\tilde{A}_p$ будет треугольной. При этом под лексикографическим порядком понимается расположение, при котором сочетание $i_1 < i_2 < \ldots < i_p$ предшествует сочетанию $j_1 < j_2 < \ldots < j_p$, если первая из отличных от нуля разностей $j_1 - i_1, j_2 - i_2, \ldots, j_p - i_p$ положительна. Диагональные элементы треугольных матриц A и $\widetilde{A_p}$ связаны так же, как было указано в свойстве 6).

9) Характеристические числа p-й ассоциированной матрицы $\widetilde{A_p}$ (с учетом их кратности) равны всевозможным произведениям по p характеристических чисел матрицы A.

10) Если матрица A подобна диагональной матрице $B = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, то p-я ассоциированная матрица $\widetilde{A_p}$ подобна диагональной матрице $\tilde{B}_p = \{\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_N\}$, где $\tilde{\lambda}_i$ получается по правилу, указанному в свойстве 6).

Если T — матрица, преобразующая A в B, то $\tilde{T}_p$ будет матрицей, преобразующей $\tilde{A}_p$ в $\tilde{B}_p$. Иными словами, если матрица A имеет n линейно независимых собственных векторов и из их координат, записанных по столбцам, образована матрица T, то p-я ассоциированная матрица $\tilde{A}_p$ имеет N линейно независимых собственных векторов, координаты которых стоят по столбцам матрицы $\tilde{T}_p$. Это значит, что если s_i — сочетание $i_1 < i_2 < \ldots < i_p$ и s_j — сочетание $j_1 < j_2 < \ldots < j_p$ в последовательности (2.49), то i-я координата собственного вектора матрицы $\tilde{A}_p$, соответствующего собственному значению $\tilde{\lambda}_j = \lambda_{j_1} \lambda_{j_2} \ldots \lambda_{j_p}$, равна минору $T\begin{pmatrix} i_1 i_2 \ldots i_p \\ j_1 j_2 \ldots j_p \end{pmatrix}$ порядка p в матрице T ([6], стр. 67—68 или [7], стр. 81).

Ассоциированной для матрицы $A = (a_{ij})_1^n$ назовем матрицу $\tilde{A} = (\tilde{a}_{ij})_1^n$, где $\tilde{a}_{ij}$ — минор элемента a_{ij} матрицы A. Таким образом, $\tilde{A}$ — частный вид $\tilde{A}_{n-1}$. Тогда:

11) $\widetilde{(\tilde{A})}=|A|^{n-2}A$, где при $n=2$ надо считать $|A|^0=1$ ([20], задача 968).

12) Если матрица A ортогональна (или унитарна, п. 16), то и p-я ассоциированная матрица $\tilde{A}_p$ ортогональна (соответственно унитарна) ([20], задачи 973, 974).

Присоединенной (или *взаимной*) матрицей для матрицы $A=(a_{ij})_1^n$ называется матрица $\hat{A}=(\hat{a}_{ij})_1^n$, где $\hat{a}_{ij}$ — алгебраическое дополнение элемента a_{ji} в матрице A. Она обладает следующими свойствами:

13) $A\hat{A}=\hat{A}A=|A|E$.

14) $\widehat{(\hat{A})}=|A|^{n-2}A$ ([20], задача 966).

15) $\widehat{(AB)}=\hat{B}\hat{A}$ ([20], задача 967).

О матрице, присоединенной к характеристической матрице $\lambda E-A$, см. стр. 95.

15. Кронекеровское произведение матриц. О кронекеровском произведении определителей см. гл. I, § 1, п. 15. Пусть все пары (i, j), где $i=1, 2, \ldots, m$; $j=1, 2, \ldots, n$, занумерованы в произвольном, но в дальнейшем неизменном порядке:

$$s_1, s_2, \ldots, s_{mn}. \tag{2.50}$$

Кронекеровским (или *прямым*) произведением двух матриц $A=(a_{ij})_1^m$ и $B=(b_{ij})_1^n$ называется матрица $C=(c_{ij})_1^{mn}$, составленная из всевозможных произведений элементов матриц A и B в надлежащем порядке. Именно, $c_{ij}=a_{i_1j_1}b_{i_2j_2}$, где $(i_1, i_2)=s_i$, $(j_1, j_2)=s_j$, $i, j=1, 2, \ldots, mn$. Оно обозначается через $C=A\times B$.

Свойства кронекеровского произведения матриц:

1) В общем случае вопрос о коммутативности и ассоциативности кронекеровского произведения не имеет смысла, так как задание последовательности (2.50) позволяет умножать только матрицу порядка m на матрицу порядка n.

2) Кронекеровское произведение дистрибутивно:

$$(A+B)\times C=A\times C+B\times C,$$

где A и B порядка m, а C порядка n;

$$A\times(B+C)=A\times B+A\times C,$$

где A порядка m, а B и C порядка n.

3) $(AB)\times(CD)=(A\times C)(B\times D)$, где A, B порядка m и C, D порядка n ([20], задача 963).

Чаще, чем общее понятие кронекеровского произведения, применяются следующие два его частных вида.

Правым кронекеровским произведением матриц $A=(a_{ij})_1^m$ и $B=(b_{ij})_1^n$ называется клеточная матрица $A\times\cdot B=C=(C_{ij})_1^m$, где $C_{ij}=a_{ij}B$, $i, j=1, 2, \ldots, m$.

Оно получается из общего кронекеровского произведения, если пары (i, j) в последовательности (2.50) занумеровать в лексикографическом порядке:

$$(1, 1), (1, 2), \ldots, (1, n),$$
$$(2, 1), (2, 2), \ldots, (2, n), \ldots, (m, 1), (m, 2), \ldots, (m, n).$$

Левым кронекеровским произведением тех же матриц называется клеточная матрица $A\cdot\times B=D=(D_{ij})_1^n$, где $D_{ij}=A\cdot b_{ij}$, $i, j=1, 2, \ldots, n$.

Оно получается, если в последовательности (2.50) пары (i, j), читаемые справа налево, занумеровать в лексикографическом порядке:

$$(1, 1), (2, 1), \ldots, (m, 1),$$
$$(1, 2), (2, 2), \ldots, (m, 2), \ldots, (1, n), (2, n), \ldots, (m, n).$$

Эти произведения имеют смысл для матриц любых порядков и, кроме свойств 2), 3), обладают еще следующими:

4) $A\times\cdot B=B\cdot\times A$.

5) Ассоциативность:

$$A\times\cdot(B\times\cdot C)=(A\times\cdot B)\times\cdot C,\quad A\cdot\times(B\cdot\times C)=$$
$$=(A\cdot\times B)\cdot\times C.$$

6) Если E_k — единичная матрица порядка k, то

$$E_m\times\cdot E_n=E_m\cdot\times E_n=E_{mn}.$$

7) Если A и B — невырожденные матрицы, то $(A\times\cdot B)^{-1}=A^{-1}\times\cdot B^{-1}$, $(A\cdot\times B)^{-1}=A^{-1}\cdot\times B^{-1}$ ([20], задача 964).

8) Если матрицы A_1, B_1 подобны соответственно матрицам A_2, B_2, то кронекеровские произведения $A_1\times B_1$ и $A_2\times B_2$ (взятые при любых двух расположениях пар индексов в последовательности (2.50)) подобны между собой. В частности, при изменении порядка нумерации пар в последовательности (2.50) матрица $A\times B$ заменяется подобной матрицей ([20], задача 1056).

9) Кронекеровское произведение диагональных матриц диагонально.

10) Правое и левое кронекеровские произведения двух верхних (или двух нижних) треугольных матриц являются верхней (нижней) треугольной матрицей.

11) Если $\lambda_1, \lambda_2, \ldots, \lambda_m$ — спектр матрицы A и $\mu_1, \mu_2, \ldots, \mu_n$ — спектр матрицы B, то спектр их кронекеровского произведения $A \times B$ состоит из всевозможных произведений $\lambda_i \cdot \mu_j$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ ([20], задача 1146).

12) Если матрицы A и B подобны диагональным матрицам, то это верно и для их кронекеровского произведения $A \times B$.

16. Ортогональные и унитарные матрицы. *Скалярным произведением* двух n-мерных векторов $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, $b = (\beta_1, \beta_2, \ldots, \beta_n)$ в случае действительных координат называется число

$$(a, b) = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_n\beta_n, \tag{2.51}$$

а в случае комплексных координат — число

$$(a, b) = \alpha_1\overline{\beta}_1 + \alpha_2\overline{\beta}_2 + \ldots + \alpha_n\overline{\beta}_n, \tag{2.52}$$

где чертой обозначено число, комплексно сопряженное данному числу. Можно дать другое, аксиоматическое определение скалярного произведения (см. [11], стр. 211—212).

Если $(a, b) = 0$, то векторы a и b называются *ортогональными*. Если $(a, a) = 1$, то вектор a называется *единичным* или *нормированным*. Система векторов $e_1, e_2, \ldots, e_s$ называется *ортонормированной*, если

$$(e_i, e_j) = \delta_{ij} = \begin{cases} 1 \text{ при } i = j; \\ 0 \text{ при } i \neq j; \end{cases} \quad i, j = 1, 2, \ldots, s. \tag{2.53}$$

Квадратная матрица A называется *ортогональной*, если ее элементы действительны и выполняется одно из следующих пяти эквивалентных свойств: а) строки образуют ортонормированную систему, б) столбцы образуют ортонормированную систему, в) $AA' = E$, г) $A'A = E$, д) $A' = A^{-1}$. Здесь A' обозначает транспонированную матрицу A. Аналогично определяются и комплексные ортогональные матрицы, но мы ограничимся лишь действительными матрицами.

В случае комплексных чисел. близким понятием является понятие унитарной матрицы. Матрицей, сопряженной с матрицей A, называется матрица $A^* = \bar{A}'$, полученная из A транспонированием и заменой всех элементов на комплексно сопряженные.

Квадратная матрица A с любыми комплексными (в частности, действительными) элементами называется *унитарной*, если выполняется одно из пяти эквивалентных свойств: а) строки образуют ортонормированную систему (при этом скалярное произведение определяется по формуле (2.52)), б) столбцы образуют ортонормированную систему, в) $AA^* = E$, г) $A^*A = E$, д) $A^* = A^{-1}$.

Унитарная матрица тогда и только тогда ортогональна, когда она действительна. Ортогональная или унитарная матрица не вырождена, причем определитель ортогональной матрицы равен ± 1, а определитель унитарной матрицы по модулю равен 1.

Произведение двух ортогональных матриц и матрица, обратная для ортогональной, ортогональны. Иными словами, все ортогональные матрицы данного порядка образуют группу относительно умножения (гл. IV). То же верно и для унитарных матриц.

Ортогональная матрица может не иметь действительных характеристических чисел. Например, матрица $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$, где $\varphi \neq k\pi$, имеет характеристические числа $\cos\varphi \pm i\sin\varphi$, причем $\sin\varphi \neq 0$. Характеристические числа унитарной (в частности, ортогональной) матрицы по модулю равны единице. Поэтому все действительные характеристические числа (если они существуют) равны ± 1. Собственные векторы ортогональной или унитарной матрицы, принадлежащие двум различным собственным значениям, ортогональны.

Две матрицы A и B называются *ортогонально* (или *унитарно*) *подобными*, если существует ортогональная (соответственно унитарная) матрица T, удовлетворяющая равенству

$$B = T^{-1}AT. \tag{2.54}$$

Матрица A тогда и только тогда унитарна, когда она унитарно подобна диагональной матрице B, диагональные элементы которой по модулю равны единице, т. е.

$$B = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}, \quad |\lambda_i| = 1, \quad i = 1, 2, \ldots, n. \tag{2.55}$$

При этом числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, стоящие на диагонали матрицы B, образуют спектр матрицы A, а столбцы преобразующей матрицы T состоят из координат ортонормированной системы n собственных векторов матрицы A, принадлежащих собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_n$ в соответствующем порядке ([6], стр. 224, теорема 6').

Действительная матрица A порядка n тогда и только тогда ортогональна, когда она ортогонально подобна канонической ортогональной матрице

$$B=\left\{\begin{pmatrix}\cos\varphi_1 & -\sin\varphi_1\\ \sin\varphi_1 & \cos\varphi_1\end{pmatrix}, \ldots\right.$$
$$\left.\ldots, \begin{pmatrix}\cos\varphi_q & -\sin\varphi_q\\ \sin\varphi_q & \cos\varphi_q\end{pmatrix}, 1, \ldots, 1, -1, \ldots, -1\right\}. \qquad (2.56)$$

Здесь $\sin\varphi_k>0$, $k=1, 2, \ldots, q$; число элементов $+1$ равно r, число элементов -1 равно s; $2q+r+s=n$. Некоторые из чисел q, r, s могут равняться нулю, т. е. клетки некоторых из указанных типов могут отсутствовать ([6], стр. 233). Клетки второго порядка соответствуют парам комплексно сопряженных характеристических чисел, а первого — действительным характеристическим числам матрицы A, так что спектр матрицы A имеет вид

$$\lambda_1=\cos\varphi_1+i\sin\varphi_1,\quad \lambda_2=\cos\varphi_1-i\sin\varphi_1, \ldots$$
$$\ldots, \lambda_{2q-1}=\cos\varphi_q+i\sin\varphi_q,\quad \lambda_{2q}=\cos\varphi_q-i\sin\varphi_q,$$
$$\lambda_{2q+1}=\ldots=\lambda_{2q+r}=1,\quad \lambda_{2q+r+1}=\ldots=\lambda_n=-1.$$

Ортогональная матрица T, преобразующая матрицу A к каноническому виду (2.56), строится следующим образом. Для комплексного характеристического числа $\lambda_{2k-1}=\cos\varphi_k+$ $+i\sin\varphi_k$, $k=1, 2, \ldots, q$, кратности p существует p ортогональных собственных векторов $z_1, \ldots, z_p$ с комплексными координатами. Их можно найти как решения системы однородных линейных уравнений с матрицей $A-\lambda_{2k-1}E$, причем при отыскании вектора z_s нужно требовать его ортогональности к уже построенным векторам $z_1, \ldots, z_{s-1}$. Если $z_s=$ $=x_s+iy_s$ $(s=1, \ldots, p)$, где x_s, y_s — векторы с действительными координатами, то векторы x_s и y_s оказываются ортогональными и имеющими одинаковую длину ([20], задача 1569). Нормируя эти векторы, получим два вектора, координаты которых надо поставить в двух соседних столбцах матрицы T. Номера этих столбцов должны совпадать с номерами столбцов s-й из p одинаковых диагональных клеток вида $\begin{pmatrix} \cos\varphi_k & -\sin\varphi_k \\ \sin\varphi_k & \cos\varphi_k \end{pmatrix}$ канонической матрицы (2.56). При этом сначала надо написать столбец координат вектора, полученного нормированием коэффициента мнимой части, т. е. y_s, а затем — действительной части x_s. При обратном порядке клетки матрицы (2.56) имели бы вид $\begin{pmatrix} \cos\varphi_k & \sin\varphi_k \\ -\sin\varphi_k & \cos\varphi_k \end{pmatrix}$.

Для действительного характеристического числа λ_k кратности p тем же путем можно найти ортонормированную систему p действительных решений системы однородных линейных уравнений с матрицей $A-\lambda_k E$. Координаты этих векторов надо поставить в p столбцов матрицы T с теми же номерами, как столбцы матрицы B, содержащие λ_k на диагонали.

Пример. Для матрицы

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

найти ортогональную матрицу T, приводящую A к канонической ортогональной матрице B.

Вычисляя скалярные произведения строк матрицы A, легко проверить, что эти строки ортонормированы. Значит, матрица A ортогональна, и задача разрешима. Вычисляем характеристические числа матрицы A:

$$|A-\lambda E|=\begin{vmatrix} \frac{2}{3}-\lambda & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3}-\lambda & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3}-\lambda \end{vmatrix}=$$

(к первой строке прибавляем вторую и третью)

$$=\begin{vmatrix} 1-\lambda & 1-\lambda & 1-\lambda \\ \frac{2}{3} & \frac{2}{3}-\lambda & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3}-\lambda \end{vmatrix}=$$

(первый столбец вычитаем из остальных и разлагаем определитель по первой строке)

$$=(1-\lambda)\begin{vmatrix} -\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix}=(1-\lambda)(\lambda^2-\lambda+1);$$

$$\lambda_1=\frac{1}{2}+\frac{1}{2}i\sqrt{3}=\cos\frac{\pi}{3}+i\sin\frac{\pi}{3},$$

$$\lambda_2=\frac{1}{2}-\frac{1}{2}i\sqrt{3}=\cos\frac{\pi}{3}-i\sin\frac{\pi}{3},\quad \lambda_3=1.$$

Каноническая ортогональная матрица, ортогонально подобная матрице A, имеет вид

$$B=\begin{pmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} & 0 \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ищем ортогональную преобразующую матрицу T, удовлетворяющую равенству (2.54). Координаты собственного вектора матрицы A для значения $\lambda_1=\frac{1}{2}+\frac{1}{2}i\sqrt{3}$ найдем из системы уравнений

$$\left(\frac{1}{6}-\frac{1}{2}i\sqrt{3}\right)x_1-\frac{1}{3}x_2+\frac{2}{3}x_3=0,$$

$$\frac{2}{3}x_1+\left(\frac{1}{6}-\frac{1}{2}i\sqrt{3}\right)x_2-\frac{1}{3}x_3=0,$$

$$-\frac{1}{3}x_1+\frac{2}{3}x_2+\left(\frac{1}{6}-\frac{1}{2}i\sqrt{3}\right)x_3=0.$$

Так как определитель этой системы должен равняться нулю и уравнения не пропорциональны, то ранг системы равен двум и решение определено однозначно с точностью до числового множителя. Полагая $x_1=1$, из первых двух уравнений найдем

$$x_2=-\frac{1}{2}-\frac{1}{2}i\sqrt{3};\quad x_3=-\frac{1}{2}+\frac{1}{2}i\sqrt{3}.$$

Отделяя действительную часть и коэффициент при мнимой части, получим два вектора:

$$f_1=\left(0, -\frac{1}{2}\sqrt{3}, \frac{1}{2}\sqrt{3}\right), \quad f_2=\left(1, -\frac{1}{2}, -\frac{1}{2}\right),$$

ортогональных и имеющих одинаковую длину. Нормируя эти векторы, найдем

$$e_1=\left(0, -\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right), \quad e_2=\left(\frac{1}{3}\sqrt{6}, -\frac{1}{6}\sqrt{6}, -\frac{1}{6}\sqrt{6}\right).$$

Координаты собственного вектора матрицы A для значения $\lambda_3=1$ найдем из системы уравнений

$$-\frac{1}{3}x_1-\frac{1}{3}x_2+\frac{2}{3}x_3=0,$$

$$\frac{2}{3}x_1-\frac{1}{3}x_2-\frac{1}{3}x_3=0,$$

$$-\frac{1}{3}x_1+\frac{2}{3}x_2-\frac{1}{3}x_3=0.$$

Вычитая первое уравнение из остальных, найдем: $x_1-x_3=0$, $x_2-x_3=0$, откуда $x_1=x_2=x_3$. Так как решение определено с точностью до множителя, то можно положить $x_1=x_2=x_3=1$. Отсюда $f_3=(1, 1, 1)$. Нормируя, найдем

$$e_3=\left(\frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}\right).$$

Записав координаты векторов e_1, e_2, e_3 по столбцам, получим преобразующую матрицу

$$T=\begin{pmatrix} 0 & \frac{1}{3}\sqrt{6} & \frac{1}{3}\sqrt{3} \\ -\frac{1}{2}\sqrt{2} & -\frac{1}{6}\sqrt{6} & \frac{1}{3}\sqrt{3} \\ \frac{1}{2}\sqrt{2} & -\frac{1}{6}\sqrt{6} & \frac{1}{3}\sqrt{3} \end{pmatrix}.$$

Замечание. Решение предыдущей задачи позволяет выяснить геометрический смысл матрицы A. Эта матрица задает в ортонормированной базе ортогональное преобразование трехмерного евклидова пространства R ([11], стр. 221). Каноническая форма B показывает, что это преобразование заключается в повороте пространства R вокруг оси, определяемой вектором f_3, на угол $\frac{\pi}{3}$ в направлении от e_1 к e_2. Так как $|T|=1$, то ориентация тройки e_1, e_2, e_3 совпадает с ориентацией исходной базы. В случае трехмерного пространства геометрический смысл преобразования, заданного матрицей A, можно выяснить проще, не отыскивая всей матрицы T. В этом случае характеристический многочлен имеет хотя бы один действительный корень, в нашем примере $\lambda_3=1$. Соответствующий собственный вектор $f_3=(1, 1, 1)$ определяет ось вращения. Пусть $A=(a_{ij})_1^3$. Подобные матрицы A и B имеют один и тот же характеристический многочлен, в частности один и тот же след (стр. 93). Отсюда

$$2\cos\varphi+1=a_{11}+a_{22}+a_{33},\quad \cos\varphi=\frac{a_{11}+a_{22}+a_{33}-1}{2}.$$

В нашем примере это дает $\cos\varphi=\frac{1}{2}$, $\varphi=\frac{\pi}{3}$. Остается найти направление вращения. Для этого достаточно взять любой вектор a, не коллинеарный f_3, вектор b, полученный из a при данном преобразовании, и вектор f_3. Если смотреть из конца вектора f_3 к началу, то направление поворота будет от a к b. Вычислив определитель из координат векторов a, b, f_3, мы узнаем, совпадают ли ориентации исходной базы и тройки a, b, f_3, и этим определим направление вращения относительно исходной базы. В нашем примере можно взять вектор $a=(1, 0, 0)$. Столбец координат его образа b равен произведению матрицы A на столбец координат вектора a, т. е. равен первому столбцу матрицы A. Итак, $b=\left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)$. Вычисляем определитель, по столбцам которого стоят координаты векторов a, b, f_3:

$$\begin{vmatrix} 1 & \frac{2}{3} & 1 \\ 0 & \frac{2}{3} & 1 \\ 0 & -\frac{1}{3} & 1 \end{vmatrix}=1>0.$$

Значит, преобразование, заданное матрицей A, является поворотом вокруг оси, определенной вектором $f_3=(1, 1, 1)$, на угол $\frac{\pi}{3}$ в направлении, совпадающем с поворотом от первой оси исходной системы координат ко второй, если смотреть на плоскость этих осей со стороны положительной третьей полуоси. Это совпадает с полученным ранее поворотом на угол $\frac{\pi}{3}$ от e_1 к e_2.

Теорема Шура. Любая квадратная числовая матрица унитарно подобна треугольной матрице (по желанию, верхней или нижней) ([20], задача 1548).

Любая квадратная числовая матрица A представляется как в виде $A=UC$, так и в виде $A=DV$, где C, D — верхние (нижние) треугольные и U, V — унитарные матрицы. Если матрица A не вырождена, то оба эти разложения единственны с точностью до диагонального унитарного множителя $M = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$, $|\varepsilon_i| = 1$. Если матрица A действительна, то и матрицы C, D, U, V действителвны. Таким образом, любая действительная матрица разлагается в произведение ортогональной и треугольной матриц.

17. Симметрические, кососимметрические и эрмитовы матрицы. О ранге симметрических и кососимметрических матриц см. гл. I, § 2, п. 4.

Пусть A' — матрица, транспонированная для A, и $A^* = \overline{A}'$ — матрица, сопряженная для A. Если A действительна, то $A^* = A'$. (В настоящем пункте рассматриваются только квадратные матрицы.)

Матрица A называется *симметрической*, если $A' = A$, *кососимметрической*, если $A' = -A$, *эрмитовой*, если $A^* = A$, *косоэрмитовой*, если $A^* = -A$. Действительные эрмитовы матрицы совпадают с симметрическими, а косоэрмитовы — с кососимметрическими.

О комплексных ортогональных, симметрических и кососимметрических матрицах см. [6], гл. 11.

Эрмитовы и косоэрмитовы матрицы связаны так: если в соотношении $A = \pm iB$ одна из двух матриц эрмитова, то другая — косоэрмитова, и наоборот.

Любая матрица A однозначно представляется в виде суммы $A = B + C$, где B — симметрическая и C — кососимметрическая матрицы. При этом

$$B = \frac{A + A'}{2}, \quad C = \frac{A - A'}{2}.$$

Далее, любая матрица A однозначно представляется в виде $A = B + C$, где B — эрмитова и C — косоэрмитова матрицы. При этом

$$B = \frac{A + A^*}{2}, \quad C = \frac{A - A^*}{2}.$$

Наконец, любая матрица A однозначно представляется в виде

$$A = B + iC, \tag{2.57}$$

где обе матрицы B, C эрмитовы. При этом

$$B=\frac{A+A^*}{2}, \quad C=\frac{A-A^*}{2i}.$$

Разложение (2.57) является матричным аналогом представления комплексного числа a в виде $a=b+ic$ с действительными b, c.

Из правила транспонирования произведения $(AB)'=B'A'$ и правила замены чисел на сопряженные $(\overline{AB})=\overline{A}\,\overline{B}$ вытекает равенство

$$(AB)^*=B^*A^*. \tag{2.58}$$

Отсюда легко вывести, что произведение двух симметрических (или эрмитовых) матриц A и B тогда и только тогда будет симметрическим (соответственно эрмитовым), когда $AB=BA$. Далее, произведение двух кососимметрических (или косоэрмитовых) матриц A и B тогда и только тогда будет симметрическим (соответственно эрмитовым), когда $AB=BA$. Наконец, произведение двух кососимметрических (или косоэрмитовых) матриц A и B тогда и только тогда будет кососимметрическим (соответственно косоэрмитовым), когда $AB=-BA$.

Любую матрицу можно представить в виде произведения двух симметрических матриц, из которых одна не вырождена (см. [24], задача 978 или [20], задача 1144).

Характеристические числа эрмитовой (в частности, действительной симметрической) матрицы действительны. Характеристические числа косоэрмитовой (в частности, действительной кососимметрической) матрицы чисто мнимы, т. е. имеют вид bi, где b — действительное число (в частности, некоторые из этих чисел могут равняться нулю). Собственные векторы матрицы любого из указанных видов, принадлежащие различным собственным значениям, ортогональны.

Матрица A тогда и только тогда является эрмитовой, когда она унитарно подобна действительной диагональной матрице $B=\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ ([6], стр. 223, теорема 5'). При этом числа λ_i, стоящие на диагонали матрицы B, образуют спектр матрицы A, а столбцы унитарной преобразующей матрицы T, для которой $B=T^{-1}AT$, состоят из координат ортонормированной системы n собственных векторов матрицы A, порядок которых соответствует порядку чисел λ_i

на диагонали матрицы B. Аналогичное утверждение верно для действительной симметрической матрицы A с заменой унитарного подобия на ортогональное подобие, причем для получения ортогональной преобразующей матрицы T надо брать ортонормированную систему n собственных векторов с действительными координатами.

Пример. Для матрицы

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

найти подобную ей диагональную матрицу B и ортогональную матрицу T, для которой $B=T^{-1}AT$.

Так как A — действительная симметрическая матрица, то задача разрешима. Ищем характеристические числа матрицы A:

$$|A-\lambda E|=\begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix}=$$

(к первой строке прибавляем все остальные, первый столбец вычитаем из остальных и разлагаем определитель по первой строке)

$$=(1-\lambda)\begin{vmatrix} -\lambda-1 & -2 & 0 \\ -2 & -\lambda-1 & 0 \\ 2 & 2 & -\lambda+1 \end{vmatrix}=$$

$$=(1-\lambda)^2\begin{vmatrix} \lambda+1 & 2 \\ 2 & \lambda+1 \end{vmatrix}=(\lambda-1)^2(\lambda^2+2\lambda-3);$$

$$\lambda_1=\lambda_2=\lambda_3=1, \quad \lambda_4=-3.$$

Для собственных векторов с собственным значением 1 получаем систему линейных уравнений, эквивалентную одному уравнению

$$x_1-x_2-x_3+x_4=0.$$

Надо найти три решения, составляющих ортонормированную систему. Первое решение можно взять любым, например $a_1=(1, 1, 1, 1)$. Второе решение находим, добавив к прежнему уравнению условия ортогональности к a_1:

$$x_1+x_2+x_3+x_4=0.$$

Можно взять, например, $a_2=(1, 1, -1, -1)$. Для отыскания третьего решения добавляем еще условия ортогональности к a_2:

$$x_1+x_2-x_3-x_4=0,$$

что определяет решение однозначно с точностью до числового множителя. Берем $a_3=(1, -1, 1, -1)$. Собственный вектор с собственным значением $\lambda_4=-3$ ищем из системы уравнений

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

Сложив все уравнения и сократив на 4, найдем $x_1+x_2+x_3+x_4=0$. Вычтя это уравнение из первого и второго уравнений предыдущей системы, найдем: $x_1=x_4$, $x_2=x_3$, откуда $x_1+x_2=x_3+x_4=0$. Можно взять $a_4=(1, -1, -1, 1)$. Нормируя векторы и записывая их координаты по столбцам, получим ортогональную матрицу

$$T=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

преобразующую матрицу A к диагональному виду $B=\{1, 1, 1, -3\}$.

Матрица A тогда и только тогда является косоэрмитовой, когда она унитарно подобна диагональной матрице B с чисто мнимыми элементами на диагонали (среди которых могут быть и равные нулю). Действительная матрица A тогда и только тогда является кососимметрической, когда она ортогонально подобна канонической кососимметрической матрице

$$B=\left\{\begin{pmatrix} 0 & -a_1 \\ a_1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & -a_q \\ a_q & 0 \end{pmatrix}, 0, \ldots, 0\right\}. \quad (2.59)$$

Клетки первого или второго порядка могут отсутствовать. При этом клетки первого порядка соответствуют характеристическим числам матрицы A, равным нулю, а второго порядка — парам сопряженных характеристических чисел $\pm ia_k$ с $a_k>0$, $k=1, 2, \ldots, q$ ([6], стр. 232—233). Можно построить ортогональную систему собственных векторов z_1, $z_2, \ldots, z_n$, где $z_{2k-1}=x_k+iy_k$ принадлежит значению ia_k, а $z_{2k}=x_k-iy_k$ — значению $-ia_k$, причем координаты векторов x_k, y_k действительны, $k=1, 2, \ldots, q$. Векторы x_k, y_k

будут ортогональны и одинаковой длины ([20], задача 1610). Их можно пронормировать. Для нулевых характеристических чисел можно также взять векторы $z_{2q+1}, \ldots, z_n$ действительными и ортонормированными. Записав координаты векторов в порядке

$$y_1, x_1, y_2, x_2, \ldots, y_q, x_q, z_{2q+1}, \ldots, z_n$$

по столбцам, получим ортогональную матрицу T, преобразующую A к каноническому виду (2.59).

Матрицы с положительным спектром. Эрмитова (в частности, действительная симметрическая) матрица A называется *матрицей с положительным* (или *неотрицательным*) *спектром*, если все ее характеристические числа положительны (соответственно неотрицательны). Для того чтобы эрмитова матрица A порядка n была матрицей с положительным спектром, необходимо и достаточно, чтобы все ее угловые миноры были положительны, т. е.

$$A\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix} > 0, \quad k = 1, 2, \ldots, n. \tag{2.60}$$

Для того чтобы эрмитова матрица A порядка n была матрицей с неотрицательным спектром, необходимо и достаточно, чтобы все ее главные (а не только угловые) миноры были неотрицательны. Так, например, все угловые миноры матрицы $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ неотрицательны, но ее спектр $\lambda_1 = 0$, $\lambda_2 = -1$ не является неотрицательным.

Для любой матрицы H с неотрицательным спектром существует единственная матрица F также с неотрицательным спектром такая, что $F^2 = H$. Эта матрица F называется *арифметическим корнем* из матрицы H и обозначается через $\sqrt{H}$. Для любой матрицы A обе матрицы AA^* и A^*A являются матрицами с неотрицательным спектром, а для невырожденной матрицы A — матрицами с положительным спектром. Матрица AA^* называется *левой*, а матрица A^*A — *правой нормой* матрицы A. Матрица $\sqrt{AA^*}$ называется *левым*, а матрица $\sqrt{A^*A}$ — *правым модулем* матрицы A. Для любой матрицы H порядка n с неотрицательным спектром матрица $\sqrt{H}$ представляется в виде многочлена не выше

$(n-1)$-й степени от матрицы H с действительными коэффициентами ([6], стр. 225).

Полярные разложения матриц. Любая квадратная матрица A с комплексными элементами представляется в каждом из двух следующих видов:

$$A=HU, \quad A=U_1H_1, \tag{2.61}$$

где H и H_1 — эрмитовы матрицы с неотрицательным спектром и U и U_1 — унитарные матрицы. При этом матрицы H и H_1, определены однозначно матрицей A. Матрица H является левым, а матрица H_1 — правым модулем матрицы A. Если матрица A не вырождена, то матрицы U и U_1 также определены однозначно, а матрицы H и H_1 будут матрицами с положительным спектром. Если матрица A имеет действительные элементы, то H и H_1 — действительные симметрические матрицы с неотрицательным спектром, а U и U_1 — ортогональные матрицы. Представления (2.61) являются матричным аналогом тригонометрической формы комплексного числа $z=r(\cos\varphi+i\sin\varphi)$ ([6], стр. 225, теорема 8; стр. 233, теорема 9). Существует также матричный аналог представления комплексного числа в виде $z=re^{i\varphi}$ (см. [6], стр. 227, (87), (88)).

Если A и B — эрмитовы матрицы, причем A — матрица с положительным спектром, то матрица AB имеет действительный спектр. Если матрицы A и B являются матрицами с неотрицательным спектром и хотя бы одна из них не вырождена, то матрица AB имеет действительный неотрицательный спектр ([8], стр. 138 или [20], задачи 1602, 1603).

Сумма конечного числа матриц с неотрицательным спектром будет матрицей того же типа, а если хотя бы одно из слагаемых является матрицей с положительным спектром, то такой же будет и сумма ([20], задача 1604).

Матрица ранга r с неотрицательным спектром представляется в виде суммы r (но не менее чем r) матриц ранга 1 с неотрицательным спектром ([20], задача 1605).

Матрица A порядка n тогда и только тогда является матрицей с неотрицательным спектром ранга 1, когда она представляется в виде $A=X^*X$, где X — ненулевая строка из n чисел ([20], задача 1606).

Если матрицы $A=(a_{ij})_1^n$, $B=(b_{ij})_1^n$ являются матрицами с неотрицательным спектром, то и матрица $C=(a_{ij}b_{ij})_1^n$, полученная перемножением соответствующих элементов матриц A и B, будет того же типа ([20], задача 1607).

Формулы Кэли. Эти формулы устанавливают связь между эрмитовыми (или косоэрмитовыми) и унитарными матрицами. Если F — любая эрмитова матрица, то матрица

$$U = (E + iF)(E - iF)^{-1} = (F - iE)(F + iE)^{-1} \tag{2.62}$$

будет унитарной, не имеющей собственного значения, равного -1. Второе выражение для U получается из первого вынесением множителя i из первой и $-i$ из второй скобки. Если U — унитарная матрица, не имеющая собственного значения, равного -1, то матрица

$$F = i(E - U)(E + U)^{-1} \tag{2.63}$$

будет эрмитовой. Формулы (2.62) и (2.63) являются взаимно обратными и устанавливают взаимно однозначное соответствие между всеми эрмитовыми матрицами и всеми унитарными матрицами, не имеющими собственного значения, равного -1 ([6], стр. 228, (95)).

Если F — любая эрмитова матрица, то $K = -iF$ — любая косоэрмитова матрица, поэтому формулы (2.62) и (2.63) можно объединить в одну формулу:

$$U = (E - K)(E + K)^{-1} \tag{2.64}$$

([6], стр. 235, (123), (124)). Если одна из матриц U, K в равенстве (2.64) является косоэрмитовой (в частности, действительной кососимметрической), то другая будет унитарной (соответственно ортогональной), не имеющей собственного значения, равного -1, и наоборот. Формула (2.64) устанавливает взаимно однозначное соответствие между всеми косоэрмитовыми (или действительными кососимметрическими) матрицами и всеми унитарными (соответственно ортогональными) матрицами, не имеющими собственного значения, равного -1. При этом все невырожденные косоэрмитовы (действительные кососимметрические) матрицы отображаются на все унитарные (ортогональные) матрицы, не имеющие собственных значений, равных ± 1, а все вырожденные косоэрмитовы (действительные кососимметрические) матрицы — на все унитарные (ортогональные) матрицы, имеющие собственное значение, равное $+1$, но не имеющие значения, равного -1 ([20], задачи 1614, 1615).

18. Нормальные матрицы. Матрица A с комплексными элементами называется *нормальной*, если она перестановочна со своей сопряженной матрицей $A^* = \bar{A}'$, т. е. если $AA^* = A^*A$. В частности, действительная матрица A нормальна, если $AA' = A'A$. Легко проверить, что унитарная, эрмитова, косоэрмитова и действительные ортогональная, симметрическая, кососимметрическая матрицы являются частными видами нормальных матриц.

Нормальная матрица A тогда и только тогда является унитарной, эрмитовой или косоэрмитовой, когда ее характеристические числа соответственно равны по модулю единице, действительны или чисто мнимы (включая и случай равенства нулю).

Матрица A является нормальной тогда и только тогда, когда в представлении (2.57) (стр. 129) эрмитовы матрицы B и C перестановочны или в представлении (2.61) (стр. 134) матрицы H, U, а также H_1, U_1 перестановочны ([6], стр. 219, теорема 3 и стр. 225, теорема 8).

Матрица A является нормальной тогда и только тогда, когда она унитарно подобна диагональной матрице B ([6], стр. 222, теорема 4). При этом столбцы унитарной матрицы T, для которой $B = T^{-1}AT$, состоят из координат ортонормированной системы собственных векторов матрицы A.

Матрица A порядка n с действительными элементами является нормальной тогда и только тогда, когда она ортогонально подобна канонической нормальной матрице вида

$$B = \left\{\begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}, \ldots, \begin{pmatrix} a_q & -b_q \\ b_q & a_q \end{pmatrix}, \; c_{2q+1}, \ldots, c_n\right\}, \qquad (2.65)$$

где все числа a_k, b_k, c_k действительны и $b_k > 0$, $k = 1, 2, \ldots$ $\ldots, q$. Клетки первого или второго порядка могут отсутствовать ([6], стр. 232, (107)). Клетки первого порядка соответствуют действительным корням c_k, $k = 2q+1, \ldots, n$, а второго порядка — парам комплексно сопряженных корней $a_k \pm ib_k$, $k = 1, 2, \ldots, q$, характеристического многочлена матрицы A. Преобразующая ортогональная матрица T строится аналогично случаю ортогональной матрицы (стр. 124—125). Из канонического вида (2.65) для действительной нормальной матрицы получаются, в частности, канонические виды:

1) для ортогональной матрицы, если

$a_k^2 + b_k^2 = 1$, $k = 1, 2, \ldots, q$; $c_k = \pm 1$, $k = 2q+1, \ldots, n$;

2) для действительной симметрической матрицы, если $q=0$;
3) для действительной кососимметрической матрицы, если $d_k=0,\ k=1,\ 2,\dots,q;\quad c_k=0,\quad k=2q+1,\dots,\ n.$

Если матрица A нормальна, то каждая из сопряженных матриц A и A^* представляется в виде многочлена от другой из этих матриц ([6], стр. 222, 2°).

Собственные векторы нормальной матрицы, принадлежащие различным собственным значениям, ортогональны ([20], задача 1628).

Матрица A тогда и только тогда нормальна, когда матрицы A и A^* имеют одни и те же собственные векторы ([20], задача 1630). При этом рассматриваются как действительные, так и комплексные собственные значения и собственные векторы. Если x — собственный вектор нормальной матрицы A со значением λ_0, то тот же вектор x будет собственным для матрицы A^* с сопряженным значением $\bar{\lambda}_0$ ([20], задача 1627).

Если две нормальные (или действительные нормальные) матрицы A и B подобны, то они и унитарно (соответственно ортогонально) подобны ([6], стр. 236, теорема 10).

Любая (конечная или бесконечная) совокупность попарно перестановочных нормальных матриц имеет общую ортонормированную базу собственных векторов, и значит, при помощи одной и той же унитарной матрицы T все матрицы этой совокупности можно привести к диагональному виду ([6], стр. 237, теорема 11'). При этом если все матрицы данной совокупности действительны, то существует ортогональная матрица T, приводящая каждую из матриц этой совокупности к своему каноническому виду (2.65), причем ортогональные матрицы — к виду (2.56), симметрические — к действительному диагональному виду и кососимметрические — к виду (2.59) ([6], стр. 238, теорема 12').

§ 2. Билинейные и квадратичные формы

1. Эквивалентность. Канонический и нормальный вид. Методы Лагранжа и Якоби приведения к каноническому виду. *Билинейной формой* от двух систем по n переменных

$$x_1,\ x_2,\dots,\ x_n \text{ и } y_1,\ y_2,\dots,\ y_n$$

называется многочлен с числовыми коэффициентами, однородный и имеющий первую степень относительно переменных каждого из этих рядов. Если ввести векторы

$$x=(x_1, x_2, \ldots, x_n),$$
$$y=(y_1, y_2, \ldots, y_n),$$

то билинейная форма будет функцией от этих векторов. Поэтому для нее применяется обозначение $A(x; y)$:

$$A(x; y)=\sum_{i, j=1}^{n} a_{ij}x_i y_j. \tag{2.66}$$

Матрица $A=(a_{ij})_1^n$ из коэффициентов билинейной формы называется *матрицей* этой *формы*, а ранг этой матрицы — *рангом билинейной формы*. Если положить $y_i=x_i$, $i=1, 2, \ldots, n$, то билинейная форма обратится в однородный многочлен второй степени от n переменных $x_1, x_2, \ldots$ $\ldots, x_n$, называемый *квадратичной формой* от этих переменных.

Итак, каждой билинейной форме соответствует единственная квадратичная форма. Двум различным билинейным формам может соответствовать одна и та же квадратичная форма. Например, билинейным формам $x_1y_1+x_1y_2+x_2y_1-x_2y_2$ и $x_1y_1+3x_1y_2-x_2y_1-x_2y_2$ соответствует квадратичная форма $x_1^2+2x_1x_2-x_2^2$.

Для установления взаимно однозначного соответствия между билинейными и квадратичными формами выделяют особый класс билинейных форм.

Билинейная форма называется *симметрической*, если ее матрица симметрическая (стр. 129). Симметрическая билинейная форма не изменяется при перемене местами двух рядов ее переменных, т. е. $A(x; y)=A(y; x)$. Любая квадратичная форма получается из единственной симметрической билинейной формы, которая называется *полярной билинейной формой* данной квадратичной формы. Матрица полярной билинейной формы называется *матрицей* данной *квадратичной формы*. Таким образом, квадратичную форму можно записать в виде

$$A(x; x)=\sum_{i, j=1}^{n} a_{ij}\, x_i x_j, \quad a_{ij}=a_{ji}, \quad i, j=1, 2, \ldots, n. \tag{2.67}$$

Выражение в правой части этого равенства можно принять за определение квадратичной формы. Матрица квадратичной формы $A=(a_{ij})_1^n$, по определению, является симметрической.

Ранг матрицы A квадратичной формы называется *рангом* этой формы, определитель $|A|$ — *дискриминантом* квадратичной формы. Квадратичная форма называется *невырожденной*, если $|A|\neq 0$, и *вырожденной* (или *сингулярной*), если $|A|=0$.

Полярная билинейная форма данной квадратичной формы следующим образом выражается через эту квадратичную форму:

$$A(x;\ y)=\frac{1}{2}[A(x+y;\ x+y)-A(x;\ x)-A(y;\ y)] \tag{2.68}$$

([8], стр. 53).

Если коэффициенты a_{ij} действительны, то билинейная и квадратичная формы называются *действительными*, а если комплексны, — *комплексными*. Для действительных форм значения переменных x_i, y_i обычно предполагаются действительными, а для комплексных — комплексными. Ряд свойств сохраняется также для форм с коэффициентами из любого числового поля (введение) или любого поля характеристики, не равной двум (гл. IV).

Можно определять билинейную форму $A(x;\ y)$ как функцию от двух векторов x, y аксиоматически определенного линейного пространства, линейную по каждому аргументу, и квадратичную форму $A(x;\ x)$ как функцию вектора x, полученную из билинейной формы при $y=x$ ([8], § 4). Но мы будем пользоваться данными выше определениями билинейной и квадратичной формы как многочленов вида (2.66) и (2.67).

Пусть X и Y — столбцы переменных $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$. Транспонированные матрицы X'; Y' будут строками тех же переменных. При помощи умножения матриц билинейную форму (2.66) и квадратичную форму (2.67) можно записать в виде

$$A(x;\ y)=X'AY, \tag{2.69}$$

$$A(x;\ x)=X'AX. \tag{2.70}$$

Если к билинейной форме (2.66) применить невырожденные линейные преобразования переменных

$$x_i = \sum_{j=1}^{n} s_{ij}u_j;\ y_i = \sum_{j=1}^{n} t_{ij}v_j;\ i = 1, 2, \ldots, n;\ |s_{ij}|_1^n \neq 0 \neq |t_{ij}|_1^n,$$

или в матричной форме

$$X = SU,\ Y = TV, \tag{2.71}$$

то, подставляя выражения (2.71) в (2.69), получим новую билинейную форму

$$B(u;\ v) = U'BV = \sum_{i,j=1}^{n} B_{ij}u_i v_j,$$

матрица B которой связана с матрицей A исходной формы равенством

$$B = S'AT \tag{2.72}$$

([14], стр. 209). В частности, если применить одно и то же преобразование к переменным x_i и y_i, то $S = T$ и (2.72) обращается в

$$B = T'AT,\quad |T| \neq 0. \tag{2.73}$$

Из этого равенства легко вывести, что если билинейная форма $A(x;\ y)$ была симметрической, то и новая форма $B(u,\ v)$ будет симметрической. Поэтому если к квадратичной форме (2.67) применить невырожденное линейное преобразование переменных

$$x_i = \sum_{j=1}^{n} t_{ij}y_j,\ i = 1, 2, \ldots, n,\ |T| = |t_{ij}|_1^n \neq 0, \tag{2.74}$$

то получится новая квадратичная форма

$$B(y;\ y) = \sum_{i,j=1}^{n} b_{ij}y_i y_j, \tag{2.75}$$

матрица B которой связана с матрицей A исходной формы (2.67) равенством (2.73) ([6], стр. 239, (2) и (4)). Отсюда следует, что при невырожденном линейном преобразовании переменных дискриминант квадратичной формы умножается на квадрат определителя преобразования, $|B|=|A|\cdot|T|^2$, в частности, при действительных невырожденных преобразованиях действительной формы знак ее дискриминанта не изменяется.

Две квадратичные формы, получающиеся одна из другой невырожденным линейным преобразованием переменных, называются *эквивалентными* (или *конгруэнтными*). Матрицы A и B, связанные равенством (2.73), также называются *конгруэнтными*. Если коэффициенты квадратичных форм и коэффициенты линейного преобразования действительны, то формы называются *действительно-эквивалентными*, если комплексны, — *комплексно эквивалентными*, если принадлежат некоторому числовому полю K, — *эквивалентными над полем* K. Эквивалентность квадратичных форм обладает следующими свойствами: 1) рефлективность: каждая форма эквивалентна сама себе; 2) симметрия: если первая форма эквивалентна второй, то и вторая эквивалентна первой; 3) транзитивность: если первая форма эквивалентна второй, а вторая третьей, то первая эквивалентна третьей. Отсюда следует, что все квадратичные формы с коэффициентами из некоторого числового поля K от одних и тех же переменных разбиваются на непересекающиеся классы эквивалентных над K форм. В частности, можно рассматривать классы действительно или комплексно эквивалентных форм.

Из равенства (2.73) следует, что эквивалентные квадратичные формы имеют одинаковый ранг. В случае любого числового поля K равенства рангов недостаточно для эквивалентности квадратичных форм.

Квадратичная форма называется *канонической*, если она не содержит произведений переменных, т. е. имеет вид

$$C(y;\ y)=\lambda_1 y_1^2+\lambda_2 y_2^2+\ldots+\lambda_n y_n^2. \tag{2.76}$$

Некоторые из коэффициентов λ_i могут равняться нулю. Если формы (2.67) и (2.76) эквивалентны, то форма (2.76) называется *каноническим видом* формы (2.67).

Любую квадратичную форму с коэффициентами из любого числового поля K (или любого поля характеристики, не равной двум) можно над этим полем привести к каноническому виду при помощи невырожденного линейного преобразования переменных. Укажем два метода отыскания таких преобразований.

Метод Лагранжа. Пусть надо привести к каноническому виду квадратичную форму (2.67). Если форма нулевая, т. е. все ее коэффициенты a_{ij} равны нулю, то она уже является канонической. Если форма не нулевая, но все коэффициенты a_{ii} при квадратах переменных равны нулю, то существует коэффициент $a_{ij} \neq 0$, $i \neq j$. Преобразование

$$x_i = y_1, \quad x_j = y_2, \quad x_1 = y_i, \quad x_2 = y_j, \quad x_k = y_k, \\ k \neq i, j, 1, 2 \tag{2.77}$$

является невырожденным и приводит форму к виду, где $a_{12} \neq 0$. Итак, можно считать, что $a_{12} \neq 0$ уже в исходной форме. Тогда преобразование

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_k = y_k, \quad k = 3, 4, \ldots, n, \tag{2.78}$$

приведет форму к виду, содержащему квадрат первого переменного. Если исходная форма не содержит x_1^2, но содержит x_i^2, то преобразование

$$x_1 = y_i, \quad x_i = y_1, \quad x_k = y_k, \quad k \neq 1, i \tag{2.79}$$

приведет форму к виду, содержащему квадрат первого переменного. Таким образом, без ограничения общности можно считать $a_{11} \neq 0$. Тогда, выделяя все члены, содержащие x_1, получим

$$A(x;\ x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + A_1(x;\ x) = \\ = \frac{1}{a_{11}}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)^2 + A_2(x;\ x),$$

где $A_2(x;\ x) = A_1(x;\ x) - \dfrac{1}{a_{11}}(a_{12}x_2 + \ldots + a_{1n}x_n)^2$ — форма, не содержащая первого переменного. Применив преобразование

$$\left.\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 &= x_2, \\ &\ldots\ldots\ldots\ldots\ldots \\ y_n &= x_n, \end{aligned}\right\} \tag{2.80}$$

выделим из данной формы один квадрат, т. е. приведем ее к виду $\dfrac{1}{a_{11}}y_1^2 + A_2(y;\ y)$, где $A_2(y;\ y)$ не содержит y_1.

С формой $A_2(y;\ y)$ поступаем точно так же, оставляя неизменным первое переменное (лишь меняя его обозначение). Так как все преобразования (2.77)—(2.80) не вырождены, то после конечного числа шагов получим невырожденное линейное преобразование переменных, приводящее исходную форму к каноническому виду.

Число квадратов с ненулевыми коэффициентами λ_i в каноническом виде квадратичной формы равно рангу исходной формы и не зависит от выбора невырожденного линейного преобразования, приводящего данную форму к каноническому виду.

Пример 1. Найти канонический вид и невырожденное линейное преобразование переменных, приводящее к этому виду квадратичную форму

$$A(x;\ x) = x_1x_2 + x_1x_3 + x_2x_3.$$

Так как в этой форме отсутствуют квадраты, то совершаем преобразование

$$x_1 = y_1 - y_2,\quad x_2 = y_1 + y_2,\quad x_3 = y_3. \tag{2.81}$$

Форма примет вид

$$A_1(y;\ y) = y_1^2 - y_2^2 + 2y_1y_3 = (y_1 + y_3)^2 - y_2^2 - y_3^2.$$

Применив преобразование

$$z_1 = y_1 + y_3,\quad z_2 = y_2,\quad z_3 = y_3, \tag{2.82}$$

получим канонический вид формы

$$A_2(z;\ z) = z_1^2 - z_2^2 - z_3^2.$$

Из (2.82) найдем

$$y_1 = z_1 - z_3,\quad y_2 = z_2,\quad y_3 = z_3.$$

Подставив эти выражения в (2.81), получим искомое преобразование

$$x_1 = z_1 - z_2 - z_3,$$
$$x_2 = z_1 + z_2 - z_3,$$
$$x_3 = z_3.$$

Метод Якоби. Для применения этого метода квадратичная форма (2.67) должна удовлетворять некоторым специальным условиям. Изложение метода Якоби в геометрической форме можно найти в [8], § 6. Мы дадим изложение в алгебраической форме. Пусть квадратичная форма (2.67) имеет ранг r и все угловые миноры, т. е. миноры, стоящие в левом верхнем углу матрицы формы, от первого до r-го порядка отличны от нуля:

$$D_k = A\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0, \qquad k = 1, 2, \dots, r. \tag{2.83}$$

Тогда канонический вид формы (2.67) можно найти по формуле Якоби

$$C(y;\, y) = \frac{1}{D_1}y_1^2 + \frac{D_1}{D_2}y_2^2 + \dots + \frac{D_{r-1}}{D_r}y_r^2 \tag{2.84}$$

([6], стр. 245, (26)). Коэффициенты этого канонического вида легче вычислять не непосредственно, а иным путем, дающим также и выражение новых переменных y_j через старые x_i. Именно, в силу (2.83) $a_{11} \neq 0$, и, вычитая первую строку матрицы A, умноженную на подходящие множители, из остальных, можно привести эту матрицу к виду

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a'_{22} & \dots & a'_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & a'_{n2} & \dots & a'_{nn} \end{pmatrix}, \quad a'_{22} \neq 0.$$

Теперь также обращаем в нуль все элементы второго столбца, стоящие ниже a'_{22}, и т. д., пока не приведем матрицу к треугольному виду G, в котором все строки, начиная с $(r+1)$-й, состоят лишь из нулей. Заменим нули в этих строках любыми элементами так, чтобы получилась невырожденная треугольная матрица. Например, в последних $n-r$ строках на главной диагонали поставим единицы, а остальные элементы оставим равными нулю. Получим матрицу вида

$$T = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1r} & t_{1,\,r+1} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2r} & t_{2,\,r+1} & \dots & t_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & t_{rr} & t_{r,\,r+1} & \dots & t_{rn} \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Эта матрица T и будет матрицей коэффициентов в выражениях y_j через x_i. Преобразование, обратное для построенного, приводит исходную форму (2.67) к каноническому виду (2.84).

Элементы матриц T и A связаны равенствами

$$t_{ij}=\frac{A\begin{pmatrix}1 & 2 & \dots & i-1 & i\\ 1 & 2 & \dots & i-1 & j\end{pmatrix}}{D_{i-1}};$$

$$j=i,\ i+1,\ \dots,\ n;\quad i=1,\ 2,\ \dots,\ r;\quad D_0=1. \tag{2.85}$$

Однако практически их удобнее вычислять указанным выше методом приведения матрицы A к треугольному виду. При этом попутно определяется ранг матрицы A и проверяются условия (2.83) применимости метода Якоби. Из (2.85) находим

$$t_{ii}=\frac{D_i}{D_{i-1}},\quad i=1,\ 2,\ \dots,\ r,\quad D_0=1. \tag{2.86}$$

Поэтому канонический вид (2.84) можно переписать так:

$$C(y;\ y)=\sum_{i=1}^{r}\frac{1}{t_{ii}}y_i^2, \tag{2.87}$$

что позволяет его найти по матрице T, не вычисляя миноров D_i.

Пример 2. Найти канонический вид и соответствующее невырожденное преобразование переменных для квадратичной формы $A(x;\ x)=x_1^2+x_2^2-2x_3^2+\frac{2}{3}x_4^2-$

$$-4x_1x_2+2x_1x_3-2x_1x_4+2x_2x_3+2x_2x_4.$$

Приводим матрицу формы к треугольному виду, обозначая знаком $\sim$ эквивалентность матриц:

$$\begin{pmatrix}1 & -2 & 1 & -1\\ -2 & 1 & 1 & 1\\ 1 & 1 & -2 & 0\\ -1 & 1 & 0 & \frac{2}{3}\end{pmatrix}\sim$$

(первую строку, умноженную соответственно на 2, -1, 1, прибавляем к следующим)

$$\sim\begin{pmatrix}1 & -2 & 1 & -1\\ 0 & -3 & 3 & -1\\ 0 & 3 & -3 & 1\\ 0 & -1 & 1 & -\frac{1}{3}\end{pmatrix}$$

(вторую строку, умноженную на 1 и $-\frac{1}{3}$, прибавляем к третьей и четвертой)

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

Из (2.87) находим канонический вид $C(y;\, y)=y_1^2-\frac{1}{3}y_2^2$

$$y_1=x_1-2x_2+x_3-x_4,$$
$$y_2=-3x_2+3x_3-x_4,$$
$$y_3=x_3, \qquad y_4=x_4.$$

Отсюда находим искомое преобразование переменных

$$x_1=y_1-\frac{2}{3}y_2+y_3+\frac{1}{3}y_4,$$
$$x_2=-\frac{1}{3}y_2+y_3-\frac{1}{3}y_4,$$
$$x_3=y_3, \qquad x_4=y_4.$$

Другой вывод метода Якоби см. в [7], стр. 43—44.

При выполнении условий (2.83) для формы ранга r метод Лагранжа проводится лишь путем преобразований вида (2.80) и дает тот же результат, что и метод Якоби. Однако как канонический вид данной квадратичной формы, так и преобразование, приводящее ее к данному каноническому виду, *не определены однозначно*. Так, если $a_{11}\neq 0$, то при выделении квадрата первого неизвестного можно выносить за скобку не $\frac{1}{a_{11}}$, а любой ненулевой множитель. Этот множитель и будет коэффициентом при квадрате первого переменного в каноническом виде. Если, например, вынести множитель a_{11}, то вместо (2.80) получится преобразование

$$y_1=x_1+\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n,$$
$$y_i=x_i \qquad \text{при } i>1.$$

При условиях (2.83) форма ранга r приводится к каноническому виду (2.76) путем нескольких преобразований этого вида, что дает в итоге треугольное преобразование специального вида

$$\left.\begin{aligned} y_1 &= x_1 + c_{12}x_2 + \ldots + c_{1n}x_n, \\ y_2 &= x_2 + c_{23}x_3 + \ldots + c_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ y_n &= x_n \end{aligned}\right\} \tag{2.88}$$

с определителем, равным единице. Обратное преобразование будет того же специального вида

$$\left.\begin{aligned} x_1 &= y_1 + d_{12}y_2 + \ldots + d_{1n}y_n, \\ x_2 &= y_2 + d_{23}y_3 + \ldots + d_{2n}y_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ x_n &= y_n. \end{aligned}\right\} \tag{2.89}$$

Можно доказать, что при преобразовании вида (2.89) угловые миноры матрицы A квадратичной формы не изменяются, и поэтому канонический вид, к которому приводится форма ранга r при условиях (2.83) преобразованием вида (2.89), должен выглядеть так:

$$C(y;\ y) = D_1y_1^2 + \frac{D_2}{D_1}y_2^2 + \ldots + \frac{D_r}{D_{r-1}}y_r^2 \tag{2.90}$$

([20], задачи 1221, 1222).

Итак, канонический вид не определен однозначно для данного класса эквивалентных квадратичных форм. Мы укажем здесь нормальный вид для комплексных форм (а в следующем пункте — для действительных форм), однозначно определенный для каждого класса эквивалентных форм. За счет изменения нумерации переменных y_i в каноническом виде (2.76) можно считать, что для формы ранга r выполняются условия

$$\lambda_i \neq 0 \quad \text{при} \quad i = 1,\ 2,\ \ldots,\ r;$$
$$\lambda_i = 0 \quad \text{при} \quad i = r+1,\ \ldots,\ n.$$

Тогда преобразование

$$z_i = \sqrt{\lambda_i}\,y_i \quad \text{при} \quad i = 1,\ 2,\ \ldots,\ r,$$
$$z_i = y_i \quad \text{при} \quad i = r+1,\ \ldots,\ n$$

будет невырожденным и приведет (2.76) к виду

$$D(z;\ z) = z_1^2 + z_2^2 + \ldots + z_r^2. \tag{2.91}$$

Этот вид называется *нормальным видом* исходной формы.

Две комплексные квадратичные формы тогда и только тогда комплексно эквивалентны, когда они имеют один и тот же нормальный вид, т. е. когда они имеют одинаковый ранг ([11], стр. 175). Так как ранг формы от n переменных может принимать любое значение от нуля до n, то число классов комплексно эквивалентных квадратичных форм от n переменных равно $n+1$.

Канонический вид билинейных форм. Так как матрица квадратичной формы (2.67) и матрица ее полярной билинейной формы при одинаковом преобразовании переменных преобразуются по одной и той же формуле (2.73), то любая симметрическая билинейная форма при помощи невырожденного линейного преобразования переменных (а именно преобразования, приводящего соответствующую квадратичную форму к каноническому виду (2.76), примененного к обоим рядам переменных x_i и y_i) приводится к каноническому виду

$$C(u;\ v)=\lambda_1 u_1 v_1+\lambda_2 u_2 v_2+\ldots+\lambda_n u_n v_n. \qquad (2.92)$$

Несимметрическую билинейную форму одинаковым преобразованием обоих рядов переменных нельзя привести к каноническому виду.

Билинейная форма (2.66) называется *кососимметрической*, если ее матрица A является кососимметрической (стр. 129). Ранг кососимметрической билинейной формы — всегда число четное. Любая кососимметрическая билинейная форма с коэффициентами из некоторого числового поля K посредством невырожденного линейного преобразования с коэффициентами из K, одинакового для обоих рядов переменных, приводится к нормальному виду

$$D(u;\ v)=u_1 v_2-u_2 v_1+u_3 v_4-u_4 v_3+\ldots$$
$$\ldots+u_{r-1} v_r-u_r v_{r-1}, \qquad (2.93)$$

где r — ранг билинейной формы ([14], стр. 217).

Из возможности приведения форм к виду (2.76), (2.91) и (2.93) вытекают соответствующие свойства матриц:

1) Любая симметрическая матрица с элементами из числового поля K конгруэнтна над K диагональной матрице.

2) Любая комплексная симметрическая матрица комплексно конгруэнтна диагональной матрице вида $\{1, \ldots, 1, 0, \ldots, 0\}$, где число единиц равно рангу данной матрицы.

3) Любая кососимметрическая матрица ранга r с элементами из числового поля K конгруэнтна над K матрице

$$\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \ldots, 0\right\},$$

где число клеток второго порядка равно $\frac{r}{2}$.

О приведении к каноническому виду некоторых квадратичных форм частного вида см. [20], задачи 1193—1198.

2. Действительные квадратичные формы. Закон инерции. Индексы инерции и сигнатура. Все формы в этом пункте предполагаются действительными и под эквивалентностью форм понимается их эквивалентность над полем действительных чисел.

Методы Лагранжа и Якоби приведения квадратичной формы к каноническому виду, изложенные в предыдущем пункте, в применении к действительной форме не выводят из области действительных чисел. Поэтому любая действительная квадратичная форма эквивалентна над полем действительных чисел форме канонического вида. За счет изменения нумерации переменных форму ранга r можно привести к виду

$$C(y;\ y) = \lambda_1 y_1^2 + \ldots + \lambda_p y_p^2 - \lambda_{p+1} y_{p+1}^2 - \ldots - \lambda_{p+q} y_{p+q}^2,$$
$$p + q = r, \quad \lambda_i > 0, \quad i = 1, 2, \ldots, r. \tag{2.94}$$

При этом числа p и q могут равняться нулю. Действительное невырожденное преобразование переменных

$$z_i = \sqrt{\lambda_i}\, y_i, \quad i = 1, 2, \ldots, r; \qquad z_i = y_i, \quad i = r+1, \ldots, n,$$

приводит канонический вид (2.94) к виду

$$D(z;\ z) = z_1^2 + \ldots + z_p^2 - z_{p+1}^2 - \ldots - z_{p+q}^2,$$
$$p + q = r, \tag{2.95}$$

называемому *нормальным видом* действительной квадратичной формы.

Закон инерции квадратичных форм. Число p положительных и число q отрицательных коэффициентов в каноническом виде (2.94) или в нормальном виде (2.95) данной квадратичной формы не зависит от выбора того действительного невырожденного преобразования переменных, которое приводит исходную форму к этому виду ([11], стр. 176). Числа p и q называются соответственно *положительным и отрицательным индексами инерции*, их разность $s=p-q$ называется *сигнатурой* квадратичной формы. Индексы инерции, ранг и сигнатура связаны следующим образом:

$$\left.\begin{aligned} r&=p+q, & s&=p-q;\\ p&=\frac{1}{2}(r+s), & q&=\frac{1}{2}(r-s), \end{aligned}\right\} \tag{2.96}$$

поэтому r и s однозначно определяют p и q; и наоборот.

Из закона инерции следует, что две квадратичные формы тогда и только тогда эквивалентны (в действительной области), когда они имеют одинаковый нормальный вид, или одинаковые индексы инерции, или одинаковые ранг и сигнатуру.

Число классов действительно эквивалентных квадратичных форм от n переменных равно $\frac{(n+1)(n+2)}{2}$. Число же таких классов при заданном ранге r равно $r+1$, а при заданной сигнатуре s равно $\left[\frac{n-|s|}{2}\right]+1$, где под $[a]$ понимается целая часть числа a ([20], задачи 1203, 1205).

Из записи канонического вида в форме (2.84) или (2.90) предыдущего пункта вытекает

Теорема Якоби. Если для квадратичной формы

$$A(x;\ x)=\sum_{i,\ j=1}^{n} a_{ij}x_i x_j$$

ранга r выполняются условия (2.83) предыдущего пункта, то положительный индекс инерции p равен числу сохранений знака, а отрицательный индекс инерции q — числу перемен знака в ряду чисел

$$1,\ D_1,\ D_2,\ \ldots,\ D_r, \tag{2.97}$$

$$\left.\begin{aligned} p&=P(1,\ D_1,\ D_2,\ \ldots,\ D_r),\\ q&=V(1,\ D_1,\ D_2,\ \ldots,\ D_r). \end{aligned}\right\} \tag{2.98}$$

Поэтому для сигнатуры имеем

$$s=r-2q=r-2V(1,\ D_1,\ D_2,\ \ldots,\ D_r) \tag{2.99}$$

([6], стр. 246, теорема 2).

Формула (2.99) для сигнатуры сохраняет силу и в некоторых случаях, когда отдельные миноры в ряду (2.97) равны нулю. Именно, имеет место

Правило Гульденфингера. Если для квадратичной формы ранга r: а) $D_r \neq 0$, б) в ряду миноров (2.97) нет двух нулей подряд, то сигнатуру этой формы можно определять из равенства (2.99), где при подсчете числа перемен знака нули опускаются ([6], стр. 246).

Правило Фробениуса. Если для квадратичной формы ранга r: а) $D_r \neq 0$, б) в ряду миноров (2.97) нет трех нулей подряд, то сигнатуру можно определять из равенства (2.99), причем если $D_k = 0$, но $D_{k-1}D_{k+1} \neq 0$, то нуль опускается, если же $D_k = D_{k+1} = 0$, но $D_{k-1}D_{k+2} \neq 0$, то

$$V(D_{k-1}, D_k, D_{k+1}, D_{k+2}) = \begin{cases} 1 & \text{при} \quad D_{k-1}D_{k+2} < 0, \\ 2 & \text{при} \quad D_{k-1}D_{k+2} > 0 \end{cases}$$

([6], стр. 246).

Если в ряду миноров (2.97) имеются три нуля подряд, то знаки миноров D_k не определяют однозначно сигнатуры формы. Так, для формы

$$2ax_1x_4 + bx_2^2 + cx_3^2$$

при $a \neq 0$, $bc > 0$ имеем

$$D_1 = D_2 = D_3 = 0, \quad D_4 = -a^2bc < 0.$$

Но при $b > 0$, $c > 0$ будет $q = 1$, а при $b < 0$, $c < 0$ — $q = 3$. Так же знаки миноров (2.97) не определяют сигнатуры формы ранга r, если $D_1 \cdot D_2 \ldots D_{r-1} \neq 0$, но $D_r = 0$. Так, при $a \neq 0$, $a \neq b$ форма

$$ax_1^2 + ax_2^2 + bx_3^2 + 2ax_1x_2 + 2ax_1x_3 + 2ax_2x_3 = \\ = a(x_1 + x_2 + x_3)^2 + (b - a)x_3^2$$

имеет ранг два, но $s = 0$ при $a > b > 0$ и $s = 2$ при $b > a > 0$, хотя в обоих случаях $D_1 = a > 0$, $D_2 = D_3 = 0$ ([6], стр. 247).

Линейной формой называется однородный многочлен первой степени, т. е. многочлен вида $a_1x_1 + a_2x_2 + \ldots + a_nx_n$. Произведение двух линейных форм от n переменных является квадратичной формой от этих переменных. Однако не любая квадратичная форма разлагается в произведение линейных форм. Для того чтобы квадратичная форма $A(x, x) = \sum_{i, j=1}^{n} a_{ij}x_i x_j$ с комплексными или действительными коэффициентами разлагалась в произведение двух линейных форм, необходимо и достаточно выполнение следующих условий: 1) для комплексных линейных форм: ранг квадратичной формы $\leqslant 2$; 2) для действительных линейных форм: ранг квадратичной формы $\leqslant 2$ и при ранге 2 сигнатура равна нулю ([11], стр. 178). Из существования нормального вида (2.95) вытекает такое свойство матриц: любая действительная симметрическая матрица действительно конгруэнтна единственной (для данной матрицы) диагональной матрице вида $\{1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0\}$, где число элементов ± 1 равно рангу данной матрицы (элементы $+1$, -1 или 0 могут и отсутствовать).

3. Положительно определенные квадратичные формы. Действительная квадратичная форма

$$A(x; x) = \sum_{i, j=1}^{n} a_{ij}x_i x_j \tag{2.100}$$

называется *положительно определенной*, если $A(x; x) > 0$ при любых действительных значениях переменных $x_1, x_2, \ldots, x_n$, не все из которых равны нулю; *отрицательно определенной*, если $A(x; x) < 0$ при тех же условиях; *неотрицательной*, если $A(x; x) \geqslant 0$, и *неположительной*, если $A(x; x) \leqslant 0$ при любых действительных значениях переменных. Эти четыре свойства мы будем обозначать соответственно неравенствами

$$A(x; x) > 0; \quad A(x; x) < 0; \quad A(x; x) \geqslant 0, \quad A(x; x) \leqslant 0.$$

Формы $A(x; x) > 0$ и $A(x; x) < 0$ называются *знакоопределенными*, а формы $A(x; x) \geqslant 0$ и $A(x; x) \leqslant 0$ — *полуопределенными*. Знакоопределенные формы всегда невырождены. Полуопределенные, но не знакоопределенные формы всегда вырождены ([6], стр. 248). Если $A(x; x) > 0$ (или $\geqslant 0$), то $-A(x; x) < 0$ (соответственно $\leqslant 0$). Поэтому каждому свойству форм одного знака соответствует определенное свойство форм противоположного знака.

Критерии для положительно определенных квадратичных форм.

1) Для того чтобы действительная квадратичная форма от n неизвестных была положительно определенной, необходимо и достаточно, чтобы она приводилась к нормальному виду, состоящему из n квадратов с коэффициентами $+1$, т. е. к виду

$$D(z;\ z)=z_1^2+z_2^2+\ldots+z_n^2, \tag{2.101}$$

или чтобы выполнялось одно из равенств: $p=n$, $s=n$.

2) Критерий Сильвестра. Для того чтобы действительная квадратичная форма (2.100) была положительно определенной, необходимо и достаточно, чтобы все угловые миноры ее матрицы $D_1, D_2, \ldots, D_n$, где

$$D_k=\left|A\begin{pmatrix}1 & 2 & \ldots & k\\ 1 & 2 & \ldots & k\end{pmatrix}\right|,\qquad k=1,\ 2,\ \ldots,\ n, \tag{2.102}$$

были положительны

$$D_k>0,\qquad k=1,\ 2,\ \ldots,\ n \tag{2.103}$$

([11], стр. 181).

Отсюда можно вывести, что если $A(x;\ x)>0$, то и все главные миноры $A\begin{pmatrix}i_1 & i_2 & \ldots & i_k\\ i_1 & i_2 & \ldots & i_k\end{pmatrix}$ ее матрицы положительны ([6], стр. 249). В частности, все коэффициенты при квадратах переменных положительны: $a_{kk}>0$ $(k=1,\ 2,\ \ldots,\ n)$.

3) Для того чтобы действительная форма (2.100) была неотрицательной, необходимо и достаточно, чтобы все ее главные миноры были неотрицательны ([6], стр. 249, теорема 4).

Неотрицательности одних угловых миноров для неотрицательности формы недостаточно.

4) Для того чтобы действительная квадратичная форма была положительно определенной (неотрицательной), необходимо и достаточно, чтобы все характеристические числа ее матрицы были положительны (неотрицательны).

5) Для того чтобы действительная квадратичная форма (2.100) была положительно определенной, необходимо и достаточно, чтобы ее матрица A представлялась в виде

$$A=C'C, \tag{2.104}$$

где C — невырожденная действительная матрица порядка n; C' — транспонированная матрица.

6) Для того чтобы действительная квадратичная форма (2.100) ранга r была неотрицательной, необходимо и достаточно, чтобы ее матрица A представлялась в виде (2.104), где C — действительная матрица порядка n и ранга r. Можно считать, что первые r строк матрицы C линейно независимы, а остальные строки нулевые ([20], задача 1210г).

Из свойств 1) — 6) при переходе к форме $-A(x;\ x)$ получаются соответствующие свойства отрицательно определенных и неположительных форм. Отметим лишь свойство, соответствующее 2):

7) Для того чтобы действительная квадратичная форма (2.100) была отрицательно определенной, необходимо и достаточно, чтобы знаки угловых миноров (2.102) чередовались, причем $D_1 < 0$.

8) Если обе квадратичные формы

$$A(x;\ x) = \sum_{i,\,j=1}^{n} a_{ij}x_ix_j \quad \text{и} \quad B(x;\ x) = \sum_{i,\,j=1}^{n} b_{ij}x_ix_j$$

положительно определенны (или неотрицательны), то и форма $C(x;\ x) = \sum_{i,\,j=1}^{n} a_{ij}b_{ij}x_ix_j$ положительно определенна (неотрицательна) ([20], задача 1220).

4. Ортогональные преобразования квадратичных форм. Приведение квадратичной формы к главным осям. Линейное преобразование переменных

$$x_i = \sum_{j=1}^{n} t_{ij}y_j,\quad i = 1,\ 2,\ \ldots,\ n, \tag{2.105}$$

называется *ортогональным*, если его матрица $T = (t_{ij})_1^n$ ортогональна, т. е. $T' = T^{-1}$ (стр. 122).

Если к квадратичной форме

$$A(x;\ x) = \sum_{i,\,j=1}^{n} a_{ij}x_ix_j \tag{2.106}$$

применить ортогональное преобразование переменных (2.105), то равенство

$$B = T'AT, \tag{2.107}$$

связывающее матрицу A формы (2.106) и матрицу B формы, полученной в результате преобразования, можно переписать так:

$$B = T^{-1}AT. \tag{2.108}$$

В частности, если $A(x; x) = x_1^2 + x_2^2 + \ldots + x_n^2$, т. е. $A = E$, то и $B = E$, и $x_1^2 + x_2^2 + \ldots + x_n^2 \equiv y_1^2 + y_2^2 + \ldots + y_n^2$. Это тождество необходимо и достаточно для того, чтобы преобразование (2.105) было ортогональным.

Итак, отношение конгруэнтности (2.107) превращается в отношение подобия (2.108). Из того, что любая действительная симметрическая матрица ортогонально подобна действительной диагональной матрице (стр. 131), следует, что для любой действительной квадратичной формы (2.106) существует ортогональное преобразование переменных (2.105), приводящее ее к каноническому виду

$$B(y; y) = \sum_{i=1}^{n} \lambda_i y_i^2. \tag{2.109}$$

При любом таком преобразовании коэффициенты λ_i при квадратах переменных оказываются характеристическими числами (с учетом их кратности) матрицы A данной формы ([6], стр. 251, теорема 7). Поэтому канонический вид (2.109) определен однозначно с точностью до порядка коэффициентов λ_i.

Если форма (2.106) имеет ранг r, то можно считать $\lambda_i \neq 0$ при $i = 1, 2, \ldots, r$ и $\lambda_i = 0$ при $i = r+1, \ldots, n$. Ортогональное преобразование (2.105), приводящее форму (2.106) к каноническому виду (2.109), в общем случае определяется неоднозначно.

Для того чтобы ортогональное преобразование переменных (2.105) приводило действительную квадратичную форму (2.106) к каноническому виду (2.109), необходимо и достаточно, чтобы столбцы матрицы T составляли ортонормированную систему собственных векторов (с действительными координатами) матрицы A данной формы. Порядок столбцов в матрице T должен соответствовать порядку коэффициентов λ_i в каноническом виде (2.109). Столбцы, соответствующие простым корням, определены однозначно с точностью до знака, а соответствующие кратным корням определены не однозначно. Число столбцов, соответствующих данному корню, равно кратности этого корня.

Пример 1. Найти ортогональное преобразование переменных, приводящее к каноническому виду квадратичную форму

$$A(x;\ x)=2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4.$$

Матрица этой формы

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

была рассмотрена в примере 1 на стр. 131. Там были найдены собственные значения $\lambda_1=\lambda_2=\lambda_3=1$, $\lambda_4=-3$ и ортогональная матрица, по столбцам которой стоят координаты собственных векторов,

$$T=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Поэтому ортогональное преобразование переменных

$$x_1=\frac{1}{2}(y_1+y_2+y_3+y_4),$$

$$x_2=\frac{1}{2}(y_1+y_2-y_3-y_4),$$

$$x_3=\frac{1}{2}(y_1-y_2+y_3-y_4),$$

$$x_4=\frac{1}{2}(y_1-y_2-y_3+y_4)$$

приведет данную форму к виду

$$B(y;\ y)=y_1^2+y_2^2+y_3^2-3y_4^2.$$

Ранг квадратичной формы (2.106) равен числу отличных от нуля собственных значений ее матрицы A, индексы инерции — числу положительных и числу отрицательных среди этих значений, а сигнатура — разности этих чисел. Если при непрерывном изменении коэффициентов действительной формы остается неизменным ее ранг, то остается неизменной также сигнатура, и наоборот ([6], стр. 251).

Замечание. Приведение квадратичной формы к каноническому виду при помощи ортогонального преобразования переменных называется также приведением этой формы к *главным осям*. Это название объясняется связью с геометрией. Пусть дана центральная гиперповерхность второго порядка n-мерного евклидова пространства:

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j = c, \quad a_{ij} = a_{ji}, \quad c \neq 0. \tag{2.110}$$

Приведя квадратичную форму, стоящую в левой части уравнения к каноническому виду ортогональным преобразованием, получим уравнение

$$\sum_{i=1}^{n} \lambda_i y_i^2 = c, \tag{2.111}$$

которое показывает, что новые оси координат являются осями симметрии или, как их называют, *главными осями* гиперповерхности. Направления этих осей заданы векторами, координаты которых стоят по столбцам матрицы T преобразования (2.105). За счет изменения знака одного из столбцов можно считать $|T| = 1$. Тогда новая система координат получается из старой путем некоторого «поворота».

Из свойств перестановочных нормальных (в частности, действительных симметрических) матриц (стр. 137) следует, что совокупность двух или более действительных квадратичных форм тогда и только тогда имеет ортогональное преобразование переменных, одновременно приводящее к каноническому виду каждую форму данной совокупности, когда матрицы форм этой совокупности попарно перестановочны.

В геометрической форме это означает, что две или более центральных поверхностей второго порядка тогда и только тогда имеют параллельные главные оси, когда их матрицы попарно перестановочны.

Канонический вид неоднородного многочлена 2-й степени. По аналогии с геометрией можно рассматривать более общую алгебраическую задачу — приведение любого (необязательно однородного) многочлена 2-й степени от n переменных к каноническому виду путем ортогонального (не обязательно однородного) преобразования переменных. Пусть дан многочлен

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j + 2\sum_{i=1}^{n} b_ix_i + c. \tag{2.112}$$

Во-первых, квадратичную форму из всех членов 2-й степени приведем к каноническому виду $\sum_{i=1}^{r}\lambda_i y_i^2$, где r — ранг этой формы, при помощи ортогонального преобразования переменных

$$x_i = \sum_{j=1}^{n} p_{ij} y_j, \quad i = 1, 2, \ldots, n. \tag{2.113}$$

После этого преобразования многочлен (2.112) примет вид

$$\left.\begin{array}{l} \sum_{i=1}^{r}\lambda_i y_i^2 + 2\sum_{i=1}^{n} b'_i y_i + c, \\ \lambda_i \neq 0, \ i = 1, 2, \ldots, r. \end{array}\right\} \tag{2.114}$$

Во-вторых, из равенства

$$\lambda_i y_i^2 + 2b'_i y_i = \lambda_i \left(y_i + \frac{b'_i}{\lambda_i}\right)^2 - \frac{b'^2_i}{\lambda_i}$$

ясно, что параллельный перенос

$$\left.\begin{array}{ll} y_i = z_i - \frac{b'_i}{\lambda_i} & \text{при } i = 1, 2, \ldots, r, \\ y_i = z_i & \text{при } i = r+1, \ldots, n \end{array}\right\} \tag{2.115}$$

приведет многочлен (2.114) к виду

$$\sum_{i=1}^{r}\lambda_i z_i^2 + 2\sum_{i=r+1}^{n} b'_i z_i + c'; \tag{2.116}$$

в-третьих, если не все $b'_{r+1}, \ldots, b'_n$ равны нулю, то $(n-r)$-мерный вектор

$$\frac{1}{\sqrt{\sum_{i=r+1}^{n} b'^2_i}}(b'_{r+1}, \ldots, b'_n) = (q_{r+1,r+1}, \ldots, q_{r+1,n})$$

можно дополнить до ортонормированной базы $(n-r)$-мерного пространства векторами

$$\begin{gathered}(q_{r+2,r+1}, \ldots, q_{r+2,n}),\\ \cdots\cdots\cdots\\ (q_{n,r+1}, \ldots, q_{n,n}).\end{gathered}$$

Преобразование

$$\begin{gathered}u_i = z_i \text{ при } i = 1, 2, \ldots, r,\\ u_i = \sum_{j=r+1}^{n} q_{ij} z_j \text{ при } i = r+1, \ldots, n\end{gathered}$$

будет ортогональным. Обратное преобразование

$$\left.\begin{aligned} z_i &= u_i && \text{при } i = 1, 2, \ldots, r,\\ z_i &= \sum_{j=r+1}^{n} q_{ji} u_j && \text{при } i = r+1, \ldots, n \end{aligned}\right\} \tag{2.117}$$

(матрица которого получается в силу ортогональности транспонированием матрицы предыдущего преобразования) приведет многочлен (2.116) к виду

$$\sum_{i=1}^{r} \lambda_i u_i^2 + 2\lambda_{r+1} u_{r+1} + c', \tag{2.118}$$

где $\lambda_{r+1} = \sqrt{\sum_{i=r+1}^{n} b_i'^2} \neq 0.$

Наконец, преобразование

$$u_i = v_i \text{ при } i \neq r+1, \quad u_{r+1} = v_{r+1} - \frac{c'}{2\lambda_{r+1}} \tag{2.119}$$

приведет многочлен (2.118) к виду

$$\sum_{i=1}^{r} \lambda_i v_i^2 + 2\lambda_{r+1} v_{r+1}. \tag{2.120}$$

Последовательное выполнение преобразований (2.113), (2.115), (2.117) и (2.119) дает ортогональное преобразование (в общем случае неоднородное), приводящее исходный многочлен (2.112) к одному из трех канонических видов: или (2.120)

$(\lambda_{r+1} \neq 0)$, или

$$\sum_{i=1}^{r} \lambda_i v_i^2 \quad (c' = \lambda_{r+1} = 0), \tag{2.121}$$

или

$$\sum_{i=1}^{r} \lambda_i v_i^2 + a \quad (c' \neq 0, \ \lambda_{r+1} = 0). \tag{2.122}$$

Если рассматривать многочлен (2.112) как левую часть уравнения гиперповерхности второго порядка в n-мерном евклидовом пространстве, то в случае канонического вида (2.120) поверхность имеет ось симметрии, а в случаях (2.121) и (2.122) центр симметрии.

Пример 2. Найти ортогональное преобразование, приводящее к каноническому виду многочлен

$$2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4 + 3x_1 + \\ + 3x_2 + x_3 + x_4 + x_5 + x_6 + 1.$$

Квадратичная форма этого многочлена рассмотрена в примере 1. Так как теперь многочлен зависит от шести переменных, то преобразование (2.113) с учетом результата примера 1 принимает вид

$$\left.\begin{aligned} x_1 &= \frac{1}{2}(y_1 + y_2 + y_3 + y_4), & x_2 &= \frac{1}{2}(y_1 + y_2 - y_3 - y_4), \\ x_3 &= \frac{1}{2}(y_1 - y_2 + y_3 - y_4), & x_4 &= \frac{1}{2}(y_1 - y_2 - y_3 + y_4), \\ x_5 &= y_5, & x_6 &= y_6. \end{aligned}\right\} \tag{2.123}$$

Это преобразование приведет многочлен к виду

$$y_1^2 + y_2^2 + y_3^2 - 3y_4^2 + 4y_1 + 2y_2 + y_5 + y_6 + 1. \tag{2.124}$$

Полагаем

$$z_1 = y_1 + 2, \quad z_2 = y_2 + 1, \quad z_i = y_i \text{ при } i = 3, 4, 5, 6.$$

Отсюда найдем преобразование, аналогичное (2.115),

$$y_1 = z_1 - 2, \ y_2 = z_2 - 1, \ y_i = z_i \text{ при } i = 3, 4, 5, 6. \tag{2.125}$$

Это преобразование приведет многочлен (2.124) к виду

$$z_1^2 + z_2^2 + z_3^2 - 3z_4^2 + z_5 + z_6 - 4. \tag{2.126}$$

Далее надо вектор $a_1 = (1, 1)$ из коэффициентов при z_5, z_6 нормировать и дополнить до ортонормированной системы. Удобнее, однако, сначала строить ортогональную систему, а уже затем нормировать векторы. В данном примере $n - r = 6 - 4 = 2$, и для вектора a_1 надо найти лишь один вектор, к нему ортогональный. Этот вектор ищем из условия ортогональности: $x_1+x_2=0$. Можно взять $a_2=(1, -1)$. Нормируя, получим матрицу

$$\frac{1}{2}\sqrt{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}.$$

Обратная матрица совпадает с той же матрицей ввиду ее ортогональности и симметрии. Поэтому преобразование, аналогичное (2.117), имеет вид

$$\left.\begin{aligned} &z_i=u_i \quad \text{при} \quad i=1,\ 2,\ 3,\ 4,\\ &z_5=\frac{1}{2}\sqrt{2}\,(u_5+u_6),\quad z_6=\frac{1}{2}\sqrt{2}\,(u_5-u_6).\end{aligned}\right\} \tag{2.127}$$

Многочлен (2.126) принимает вид

$$u_1^2+u_2^2+u_3^2-3u_4^2+\sqrt{2}\,u_5-4. \tag{2.128}$$

Наконец, полагая $\sqrt{2}\,u_5-4=\sqrt{2}\,(u_5-2\sqrt{2})=\sqrt{2}\,v_5$, находим преобразование, аналогичное (2.119),

$$u_i=v_i \text{ при } i\neq 5,\quad u_5=v_5+2\sqrt{2}. \tag{2.129}$$

Это преобразование приведет многочлен (2.128) к каноническому виду

$$v_1^2+v_2^2+v_3^2-3v_4^2+\sqrt{2}\,v_5. \tag{2.130}$$

Выполнив преобразование (2.123), (2.125), (2.127) и (2.129), найдем неоднородное ортогональное преобразование, приводящее исходный многочлен к каноническому виду (2.130),

$$x_1=\frac{1}{2}(v_1+v_2+v_3+v_4)-\frac{3}{2},\qquad x_2=\frac{1}{2}(v_1+v_2-v_3-v_4)-\frac{3}{2},$$

$$x_3=\frac{1}{2}(v_1-v_2+v_3-v_4)-\frac{1}{2},\qquad x_4=\frac{1}{2}(v_1-v_2-v_3+v_4)-\frac{1}{2},$$

$$x_5=\frac{1}{2}\sqrt{2}\,(v_5+v_6)+2,\qquad x_6=\frac{1}{2}\sqrt{2}\,(v_5-v_6)+2.$$

Ортогональная эквивалентность. Две действительные квадратичные формы $A(x;\ x)$ и $B(x;\ x)$ от n переменных $x_1, x_2, \ldots, x_n$ называются *ортогонально эквивалентными*, если существует ортогональное преобразование вида (2.105), которое после его применения к форме $A(x, x)$ и замены переменных y_i снова на x_i переводит форму $A(x;\ x)$ в форму $B(x;\ x)$. Ортогональная эквивалентность форм обладает рефлективностью, симметрией и транзитивностью (стр. 141). Поэтому все действительные квадратичные формы от n неизвестных распадаются на классы ортогонально эквивалентных форм.

Две действительные квадратичные формы тогда и только тогда ортогонально эквивалентны, когда их матрицы имеют одинаковые характеристические многочлены или, что то же самое, одинаковые характеристические числа ([20], задача 1265). Отсюда следует, что если в каноническом виде (2.109) писать коэффициенты λ_i в определенном, например, невозрастающем, порядке, то каждый класс ортогонально эквивалентных квадратичных форм (а значит, и ортогонально подобных действительных симметрических матриц) однозначно характеризуется содержащейся в нем формой (соответственно матрицей) канонического вида.

5. Пара форм. Пучок форм. Все формы в этом пункте предполагаются действительными. Свойства пучка квадратичных форм применяются в теории малых колебаний механических систем (см. [6], гл. 10, § 8).

Пусть дана пара квадратичных форм

$$A(x;\ x)=\sum_{i,j=1}^{n} a_{ij}x_i x_j \quad \text{и} \quad B(x;\ x)=\sum_{i,j=1}^{n} b_{ij}x_i x_j. \tag{2.131}$$

Спрашивается, можно ли обе эти формы одним невырожденным линейным преобразованием неизвестных привести к каноническому виду? В общем случае эта задача не имеет решения. Так, формы

$$A(x,\ x)=x_1^2,\quad B(x;\ x)=x_1x_2$$

нельзя привести к каноническому виду одним невырожденным преобразованием (см. [11], стр. 231). Имеет место следующее достаточное условие разрешимости поставленной задачи.

Если хотя бы одна из двух форм (2.131), например $B(x;\ x)$, положительно определенная, то существует невырожденное линейное преобразование переменных, приводящее эту форму к нормальному виду, а другую — к каноническому. Для отыскания такого преобразования практически наиболее удобно сначала найти невырожденное преобразование

$$X=PY, \tag{2.132}$$

приводящее $B(x;\ x)$ к нормальному виду

$$B_1(y;\ y)=y_1^2+y_2^2+\ldots+y_n^2. \tag{2.133}$$

Применив то же преобразование (2.132) к форме $A(x; x)$, получим новую форму

$$A_1(y; y) = \sum_{i,j=1}^{n} a'_{ij} y_i y_j. \tag{2.134}$$

Для этой формы ищем ортогональное преобразование переменных

$$Y = QZ, \tag{2.135}$$

приводящее ее к каноническому виду

$$A_2(z; z) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2. \tag{2.136}$$

Применив преобразование (2.135) к форме (2.133), получим форму с матрицей $Q'EQ = Q'Q = E$, т. е. снова форму нормального вида

$$B_2(z; z) = z_1^2 + z_2^2 + \ldots + z_n^2. \tag{2.137}$$

Таким образом, невырожденное преобразование неизвестных

$$X = PQZ \tag{2.138}$$

приводит форму $A(x; x)$ к каноническому виду (2.136), а форму $B(x; x)$ — к нормальному виду (2.137) ([11], стр 232).

Пучком квадратичных форм называется совокупность форм

$$A(x, x) - \lambda B(x, x), \tag{2.139}$$

где λ — параметр, принимающий любые действительные значения. Если форма $B(x, x)$ положительно определенная, то пучок называется *регулярным*. Уравнение

$$|A - \lambda B| = 0 \tag{2.140}$$

называется *характеристическим уравнением пучка форм* (2.139) (или λ-уравнением), а его корни — характеристическими корнями или характеристическими числами этого пучка. Если λ_0 — характеристический корень пучка (2.139), то существует ненулевой вектор $z = (z_1, z_2, \ldots, z_n)$, столбец Z из координат которого удовлетворяет уравнению

$$(A - \lambda_0 B) Z = 0, \quad Z \neq 0. \tag{2.141}$$

Вектор z с этим свойством называется *главным вектором пучка* (2.139), принадлежащим характеристическому числу λ_0.

Характеристическое уравнение регулярного пучка (2.139) имеет n действительных характеристических корней (считая их кратность) $\lambda_1, \lambda_2, \ldots, \lambda_n$, которым принадлежат главные векторы

$$z_k = (z_{1k}, z_{2k}, \ldots, z_{nk}), \quad k = 1, 2, \ldots, n. \qquad (2.142)$$

Столбцы Z_k из координат этих векторов удовлетворяют условиям

$$AZ_k = \lambda_k BZ_k, \quad k = 1, 2, \ldots, n. \qquad (2.143)$$

Эти главные векторы можно выбрать так, чтобы выполнялись следующие условия ортонормированности:

$$B(z_i; z_j) = Z_i' BZ_j = \delta_{ij}, \quad i, j = 1, 2, \ldots, n. \qquad (2.144)$$

Здесь $B(x, y)$ — полярная билинейная форма квадратичной формы $B(x, x)$ ([6], стр. 252, теорема 8).

Главной матрицей регулярного пучка (2.139) называется матрица

$$Z = (z_{ij})_1^n, \qquad (2.145)$$

столбцы которой состоят из координат главных векторов, удовлетворяющих условиям (2.144). Отсюда можно вывести, что столбцы главной матрицы Z линейно независимы, и значит, эта матрица не вырождена.

Невырожденное линейное преобразование неизвестных

$$X = ZY, \qquad (2.146)$$

где X, Y — столбцы старых и новых переменных и Z — матрица преобразования, тогда и только тогда приводят положительно определенную форму $B(x, x)$ к нормальному, а другую форму $A(x, x)$ к каноническому виду, когда матрица Z является главной матрицей регулярного пучка (2.139), соответствующего данной паре форм. При этом коэффициенты λ_i канонического вида формы $A(x, x)$ являются характеристическими корнями (с учетом их кратности) пучка (2.139), следовательно, канонический вид определен однозначно до порядка нумерации коэффициентов λ_i ([6], стр. 254, теорема 9).

6. Экстремальные свойства характеристических чисел регулярного пучка форм. Пусть дан регулярный пучок квадратичных форм (его определение и свойства даны в предыдущем пункте)

$$A(x, x) - \lambda B(x, x). \tag{2.147}$$

Расположим характеристические числа этого пучка в неубывающем порядке:

$$\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n. \tag{2.148}$$

Соответствующие главные векторы пучка обозначим через

$$z_k \doteq (z_{1k}, z_{2k}, \ldots, z_{nk}); \ k = 1, 2, \ldots, n. \tag{2.149}$$

Два вектора x, y назовем *ортогональными*, если для них удовлетворяется равенство $B(x, y) = 0$. Главные векторы (2.149) можно считать попарно ортогональными и ненулевыми (даже ортонормированными, т. е. удовлетворяющими равенствам (2.144) предыдущего пункта).

Наименьшее характеристическое число λ_1 регулярного пучка форм равно минимуму отношения этих форм:

$$\lambda_1 = \min \frac{A(x, x)}{B(x, x)}, \quad x \neq 0, \tag{2.150}$$

причем этот минимум достигается тогда и только тогда, когда x является главным вектором пучка, принадлежащим значению λ_1 ([6], стр. 258, теорема 10). p-е по величине характеристическое число λ_p, $2 \leqslant p \leqslant n$, равно минимуму того же отношения форм, причем минимум берется при условии, что вектор $x \neq 0$ ортогонален к первым $p-1$ попарно ортогональным главным векторам

$$\left.\begin{aligned} &\lambda_p = \min \frac{A(x, x)}{B(x, x)}, \\ &x \neq 0, \ B(z_i, x) = 0, \ i = 1, 2, \ldots, p-1, \\ &B(z_i; z_j) = 0, \ i \neq j, \ i, j = 1, 2, \ldots, p-1. \end{aligned}\right\} \tag{2.151}$$

Этот минимум достигается тогда и только тогда, когда вектор x является главным вектором пучка, принадлежащим значению λ_p ([6], стр. 259, теорема 11).

В этой характеристике числа λ_p нужно знать первые $p-1$ главных векторов, выбор которых в общем случае не однозначен. Можно дать другую характеристику для λ_p.

Говорят, что на вектор x (или на переменные $x_1, x_2, \ldots$ $\ldots, x_n$) наложены h связей $L_1, L_2, \ldots; L_h$, если его координаты удовлетворяют системе h линейных уравнений

$$L_k(x) = l_{1k}x_1 + l_{2k}x_2 + \ldots + l_{nk}x_n = 0, \; k = 1, 2, \ldots, h. \tag{2.152}$$

Мы будем рассматривать также специализированные связи $\tilde{L}_k$, $k = 1, 2, \ldots, n$, заданные условиями ортогональности вектора x к главным векторам:

$$\tilde{L}_k(x) = B(z_k, x) = 0, \quad k = 1, 2, \ldots, n. \tag{2.153}$$

Минимум отношения форм

$$\frac{A(x, x)}{B(x, x)}$$

при условии, что на вектор x наложены h связей $L_1, L_2, \ldots$ $\ldots, L_h$, обозначим через $\min\left(\frac{A}{B}; L_1, L_2, \ldots, L_h\right)$. Тогда для p-го характеристического числа λ_p регулярного пучка форм (2.147) имеет место выражение

$$\lambda_p = \max_{L_1, L_2, \ldots, L_{p-1}} \left[\min\left(\frac{A}{B}; L_1, L_2, \ldots, L_{p-1}\right)\right], \tag{2.154}$$

где минимум берется при данных $p-1$ связях, а максимум — при варьировании этих связей ([6], стр. 260, теорема 12).

Формулу (2.151) можно переписать так:

$$\lambda_p = \min\left(\frac{A}{B}; \tilde{L}_1, \tilde{L}_2, \ldots, \tilde{L}_{p-1}\right). \tag{2.155}$$

Применяя формулы (2.150), (2.155) и (2.154) с заменой $A(x, x)$ на $-A(x, x)$, можно получить формулы

$$\lambda_n = \max \frac{A(x, x)}{B(x, x)}, \quad x \neq 0; \tag{2.156}$$

$$\lambda_{n-p+1} = \max\left(\frac{A}{B}; \tilde{L}_n, \tilde{L}_{n-1}, \ldots, \tilde{L}_{n-p+2}\right), \tag{2.157}$$

$$\lambda_{n-p+1} = \min_{L_1, L_2, \ldots, L_{p-1}} \left[\max\left(\frac{A}{B}; L_1, L_2, \ldots, L_{p-1}\right)\right] \tag{2.158}$$

([6], стр. 261, теорема 13).

Связи $L_1, L_2, \ldots, L_h$ называются *независимыми*, если ранг матрицы из их коэффициентов равен h.

Если на регулярный пучок форм (2.147) с характеристическими числами (2.148) наложить h независимых связей

$$L_1^0(x)=0,\ L_2^0(x)=0, \ldots,\ L_h^0(x)=0, \qquad (2.159)$$

то из этих уравнений h неизвестных из $x_1, x_2, \ldots, x_n$ можно выразить через остальные $n-h$. Подставляя эти выражения в пучок (2.147), получим новый пучок

$$A^0(v;\ v)-\lambda B^0(v;\ v), \qquad (2.160)$$

где через $v_1, v_2, \ldots, v_{n-h}$ обозначены оставшиеся $n-h$ из переменных $x_1, x_2, \ldots, x_n$. Пучок (2.160) снова будет регулярным. Его характеристические числа

$$\lambda_1^0 \leqslant \lambda_2^0 \leqslant \ldots \leqslant \lambda_{n-h}^0 \qquad (2.161)$$

не зависят от того, какие из h прежних неизвестных x_i были исключены при помощи данных связей (2.159). При этом справедливы неравенства

$$\lambda_p \leqslant \lambda_p^0 \leqslant \lambda_{p+h};\ p=1,\ 2, \ldots,\ n-h \qquad (2.162)$$

([6], стр. 262, теорема 14).

Пусть даны два регулярных пучка форм: пучок (2.147) с характеристическими числами (2.148) и пучок

$$\tilde{A}(x;\ x)-\lambda\tilde{B}(x;\ x) \qquad (2.163)$$

с характеристическими числами

$$\tilde{\lambda}_1 \leqslant \tilde{\lambda}_2 \leqslant \ldots \leqslant \tilde{\lambda}_n. \qquad (2.164)$$

Если для любого $x \neq 0$ выполняется неравенство

$$\frac{A(x;\ x)}{B(x;\ x)} \leqslant \frac{\tilde{A}(x;\ x)}{\tilde{B}(x;\ x)},$$

то $\lambda_p \leqslant \tilde{\lambda}_p$, $p=1,\ 2, \ldots,\ n$ ([6], стр. 263, теорема 15).

Если вторые формы пучков (2.147) и (2.163) совпадают и

$$\tilde{A}(x;\ x)=A(x;\ x)+\sum_{i=1}^{r}[X_i(x)]^2,$$

где $X_i(x)$, $i=1,\ 2, \ldots,\ r$, — линейно независимые линейные формы, то

$$\lambda_p \leqslant \tilde{\lambda}_p \leqslant \lambda_{p+r},\quad p=1,\ 2, \ldots,\ n, \qquad (2.165)$$

причем вторые части неравенств имеют место лишь при $p \leqslant n-r$ ([6], стр. 264, теорема 16).

Если первые формы пучков (2.147) и (2.163) совпадают и

$$\tilde{B}(x;\ x) = B(x;\ x) + \sum_{i=1}^{r} [X_i(x)]^2,$$

где $X_i(x)$, $i = 1, 2, \ldots, r$, — линейно независимые линейные формы, то имеют место неравенства

$$\lambda_{p-r} \leqslant \tilde{\lambda}_p \leqslant \lambda_p; \quad p = 1, 2, \ldots, n, \tag{2.166}$$

причем первые части неравенств имеют место лишь при $p > r$ ([6], стр. 264, теорема 17).

Беря в пучке (2.147) в качестве второй формы сумму квадратов

$$B(x;\ x) = x_1^2 + x_2^2 + \ldots + x_n^2,$$

из приведенных в этом пункте свойств регулярного пучка форм получим экстремальные свойства характеристических чисел одной квадратичной формы $A(x;\ x)$ (см. [8], § 17).

7. Комплексные формы второго рода. Эрмитовы формы. В случае поля комплексных чисел, кроме билинейных и квадратичных форм, определенных в п. 1, рассматриваются билинейные и квадратичные формы в другом смысле. Комплексные формы в смысле п. 1 будут называться *формами первого рода*, а в новом смысле — *формами второго рода*. Эта терминология не является общепринятой, хотя она встречается в применении к линейным формам (см., например, [8], стр. 79).

Билинейной формой второго рода называется функция от $2n$ комплексных переменных $x_1, x_2, \ldots, x_n;\ y_1, y_2, \ldots, y_n$, заданная при помощи выражения

$$A(x;\ y) = \sum_{i,j=1}^{n} a_{ij} x_i \bar{y}_j, \tag{2.167}$$

где a_{ij} $(i, j = 1, 2, \ldots, n)$ — любые комплексные числа, а черта над y_j означает, что при подстановке значений аргументов вместо $\bar{y}_j$ подставляется число, комплексно сопряженное с y_j. Коэффициенты a_{ij}, а следовательно, и запись билинейной формы в виде суммы (2.167) определены для данной формы однозначно.

Матрица из коэффициентов a_{ij} —

$$A=(a_{ij})_1^n \tag{2.168}$$

называется *матрицей билинейной формы* (2.167), а ее ранг — *рангом* этой *формы*. Можно определить билинейную форму второго рода как функцию двух векторов x, y комплексного линейного пространства (см. [8], стр. 80). Мы будем считать векторы x, y заданными их координатами $x=(x_1, x_2, \ldots, x_n)$, $y=(y_1, y_2, \ldots, y_n)$.

Квадратичной формой второго рода называется функция от n комплексных переменных $x_1, x_2, \ldots, x_n$, полученная из билинейной формы второго рода при $x=y$ или при $x_i=y_i$, $i=1, 2, \ldots, n$, т. е. функция, заданная при помощи выражения

$$A(x;\ x)=\sum_{i,j=1}^{n} a_{ij}x_i\bar{x}_j. \tag{2.169}$$

В отличие от квадратичных форм первого рода, коэффициенты a_{ij}, а следовательно, и представление (2.169), однозначно определяются квадратичной формой второго рода.

Матрица, ранг, дискриминант, понятия невырожденной и вырожденной (или сингулярной) квадратичной формы для форм второго рода определяются аналогично случаю форм первого рода (стр. 138—139).

Так как билинейные и квадратичные формы второго рода находятся во взаимно однозначном соответствии с их матрицами, то, в отличие от форм первого рода, между билинейными и квадратичными формами второго рода установлено взаимно однозначное соответствие, причем соответствующие друг другу формы имеют одинаковые матрицы. Поэтому значения билинейной формы второго рода должны определяться значениями соответствующей квадратичной формы. Существует следующая связь между этими формами:

$$\begin{aligned} A(x;\ y)=\frac{1}{4}[A(x+y;\ x+y)+iA(x+iy;\ x+iy)&-\\ -A(x-y;\ x-y)-iA(x-iy;\ x-iy)&], \end{aligned}\tag{2.170}$$

$$\begin{aligned} A(y;\ x)=\frac{1}{4}[A(x+y;\ x+y)-iA(x+iy;\ x+iy)&-\\ -A(x-y;\ x-y)+iA(x-iy;\ x-iy)&] \end{aligned}\tag{2.171}$$

([8], стр. 81—82, (1), (2)).

С помощью умножения матриц билинейная форма второго рода (2.167) записывается в виде

$$A(x;,y)=X'A\overline{Y}, \tag{2.172}$$

а квадратичная форма второго рода (2.169) — в виде

$$A(x;\ x)=X'A\overline{X}, \tag{2.173}$$

где X, Y — столбцы переменных x_i, y_i. Если к форме (2.172) применить преобразования переменных

$$X=SU, \quad Y=TV, \tag{2.174}$$

то ее матрица преобразуется по формуле

$$B=S'A\overline{T}. \tag{2.175}$$

Матрица квадратичной формы (2.173) при преобразовании переменных

$$X=TY \tag{2.176}$$

изменяется по формуле

$$B=T'A\overline{T}. \tag{2.177}$$

Эта формула после замены $Q=\overline{T}$ принимает вид

$$B=Q^*AQ, \quad Q^*=\overline{Q}'. \tag{2.178}$$

Из формул (2.175) и (2.177) следует, что при невырожденных линейных преобразованиях переменных ранг билинейной и квадратичной форм второго рода не изменяется.

Среди форм второго рода аналогичную роль с симметрическими действительными билинейными и действительными квадратичными формами играют так называемые эрмитовы формы. На эти формы переносятся все основные результаты пп. 1—5 (причем всюду скалярное умножение векторов надо определять по формуле (2.52), стр. 122) (см. [6], гл. 10, § 9).

Эрмитовой билинейной (или квадратичной) формой называется билинейная (соответственно квадратичная) форма второго рода, матрица которой является эрмитовой (стр. 129), т. е. обладает свойством

$$A^*=A, \quad a_{ij}=\overline{a}_{ji}, \quad i,j=1, 2,\dots, n. \tag{2.179}$$

Из равенства (2.178) следует, что эрмитовы билинейная и квадратичная формы при линейном преобразовании переменных (2.176) переходят снова в эрмитовы формы.

Билинейная форма второго рода $A(x; y)$ тогда и только тогда будет эрмитовой, когда выполняется тождество

$$A(x; y) = A\overline{(y; x)}. \tag{2.180}$$

Отсюда и из равенств (2.170) и (2.171) следует, что квадратичная форма второго рода $A(x; x)$ тогда и только тогда будет эрмитовой, когда она имеет действительные значения при любых комплексных значениях $x_1, x_2, \ldots, x_n$.

Из условий (2.179) эрмитовости матрицы A следует, что ее элементы, стоящие на главной диагонали, и вообще все главные миноры $A\begin{pmatrix} i_1 i_2 \ldots i_k \\ i_1 i_2 \ldots i_k \end{pmatrix}$ действительны.

На эрмитовы квадратичные формы переносятся методы Лагранжа и Якоби приведения к каноническому виду (п. 1). При выделении первого члена по методу Лагранжа при $a_{11} \neq 0$ собираем все слагаемые, содержащие x_1 и $\bar{x}_1$, и применяем тождество

$$a_{11}x_1\bar{x}_1 + \sum_{k=2}^{n} a_{1k}x_1\bar{x}_k + \sum_{k=2}^{n} a_{k1}x_k\bar{x}_1 =$$
$$= \frac{1}{a_{11}}\left(\sum_{k=1}^{n} a_{1k}x_k\right)\cdot\left(\sum_{k=1}^{n} \bar{a}_{1k}\bar{x}_k\right) -$$
$$- \frac{1}{a_{11}}\left(\sum_{k=2}^{n} a_{1k}x_k\right)\cdot\left(\sum_{k=2}^{n} \bar{a}_{1k}\bar{x}_k\right).$$

Таким образом, любая эрмитова квадратичная форма при помощи невырожденного комплексного линейного преобразования переменных приводится к каноническому виду

$$\lambda_1 y_1\bar{y}_1 + \lambda_2 y_2\bar{y}_2 + \ldots + \lambda_n y_n\bar{y}_n, \tag{2.181}$$

где коэффициенты λ_i действительны; для формы ранга r можно считать $\lambda_i \neq 0$ при $i = 1, 2, \ldots, r$; $\lambda_i = 0$ при $i = r+1, \ldots, n$.

Для эрмитовых квадратичных форм сохраняют силу закон инерции, понятия индексов инерции и сигнатуры, их определение по знакам угловых миноров, правила Фробениуса и Гульденфингера (п. 2), а также определение и критерии положительно определенных форм (п. 3).

Из того, что любая эрмитова матрица унитарно подобна действительной диагональной матрице с характеристическими числами на диагонали (стр. 130), вытекает

Приведение эрмитовой квадратичной формы к главным осям. Любую эрмитову квадратичную форму (2.169) с матрицей A при помощи унитарного преобразования переменных вида (2.176) можно привести к каноническому виду (2.181). Коэффициенты λ_i этого вида будут характеристическими числами матрицы A и, следовательно, действительны и определены однозначно. Столбцы матрицы преобразования T состоят из чисел, комплексно сопряженных координатам ортонормированной системы собственных векторов матрицы A. Напоминаем, что ортонормированную систему векторов надо понимать в смысле скалярного умножения по формуле (2.52) (стр. 122).

Свойства пары форм (п. 5) и экстремальные свойства характеристических чисел регулярного пучка (п. 6) также переносятся на любые эрмитовы формы (см. [6], гл. 10, § 9).

8. Ганкелевы формы. Пусть дана последовательность $2n-1$ чисел $s_0, s_1, s_2, \ldots, s_{2n-2}$. *Ганкелевой формой* называется квадратичная форма от n переменных $x_0, x_1, x_2, \ldots, x_{n-1}$ вида

$$S(x;\ x) = \sum_{i,\ j=0}^{n-1} s_{i+j} x_i x_j. \qquad (2.182)$$

Симметрическая матрица ганкелевой формы

$$S = \begin{pmatrix} s_0 & s_1 & s_2 & \ldots & s_{n-1} \\ s_1 & s_2 & s_3 & \ldots & s_n \\ s_2 & s_3 & s_4 & \ldots & s_{n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{n-1} & s_n & s_{n+1} & \ldots & s_{2n-2} \end{pmatrix} \qquad (2.183)$$

называется *ганкелевой матрицей*.

Для действительных ганкелевых форм индексы инерции и сигнатура полностью определяются значениями главных миноров их матриц. Следующие три предложения справедливы для ганкелевых форм с коэффициентами из любого поля.

1) Если в ганкелевой матрице (2.183) первые h строк линейно независимы, а первые $h+1$ строк линейно зависимы

$(h<n)$, то

$$D_h = S\begin{pmatrix}1 & 2 & \ldots & h\\ 1 & 2 & \ldots & h\end{pmatrix} \neq 0 \tag{2.184}$$

([6], стр. 274, лемма 1).

2) Если для ганкелевой матрицы (2.183) при некотором h выполняются условия

$$D_h \neq 0, \quad D_{h+1} = \ldots = D_n = 0, \quad h<n, \tag{2.185}$$

и если положить

$$t_{ij} = \frac{1}{D_h} S\begin{pmatrix}1 & 2 & \ldots & h & h+i+1\\ 1 & 2 & \ldots & h & h+j+1\end{pmatrix},$$
$$i, j = 0, 1, \ldots, n-h-1, \tag{2.186}$$

то матрица $T = (t_{ij})_0^{n-h-1}$ также ганкелева, и все ее элементы, лежащие выше второй диагонали, равны нулю ([6], стр. 274, лемма 2).

3) Если ганкелева матрица (2.183) имеет ранг r и при некотором h выполняются условия

$$D_h \neq 0, \quad D_{h+1} = \ldots = D_r = 0, \quad h<r, \tag{2.187}$$

то главный минор $D^{(r)}$ порядка r, стоящий на пересечении первых h и последних $r-h$ строк и таких же столбцов, отличен от нуля, т. е.

$$D^{(r)} = S\begin{pmatrix}1 & 2 & \ldots & h & n-r+h+1 & \ldots & n\\ 1 & 2 & \ldots & h & n-r+h+1 & \ldots & n\end{pmatrix} \neq 0 \tag{2.188}$$

([6], стр. 276, теорема 23).

4) Теорема Фробениуса. Для действительной ганкелевой формы (2.182) ранга r индексы инерции p, q и сигнатура s могут определяться по знакам угловых миноров (ср. п. 2) по формулам

$$\left.\begin{aligned} p &= P(1, D_1, D_2, \ldots, D_r),\\ q &= V(1, D_1, D_2, \ldots, D_r),\\ s &= r - 2V(1, D_1, D_2, \ldots, D_r),\end{aligned}\right\} \tag{2.189}$$

где P и V обозначают соответственно число сохранений и перемен знака в ряду чисел, стоящем в скобках.

При этом предполагается, что $D_r \neq 0$ и что в любой группе из m промежуточных миноров

$$D_h \neq 0, \quad D_{h+1} = \ldots = D_{h+m} = 0, \quad D_{h+m+1} \neq 0 \qquad (2.190)$$

нулевым минорам приписаны знаки по формуле

$$\operatorname{sign} D_{h+j} = (-1)^{\frac{j(j-1)}{2}} \operatorname{sign} D_h. \qquad (2.191)$$

При таком выборе знаков величины P, V, $P - V$, соответствующие группе миноров (2.190), получат значения:

$$P = \begin{cases} \dfrac{m+1}{2} & \text{при } m \text{ нечетном}, \\ \dfrac{m+1+\varepsilon}{2} & \text{при } m \text{ чётном}, \end{cases}$$

$$V = \begin{cases} \dfrac{m+1}{2} & \text{при } m \text{ нечетном}, \\ \dfrac{m+1-\varepsilon}{2} & \text{при } m \text{ четном}, \end{cases}$$

$$P - V = \begin{cases} 0 & \text{при } m \text{ нечетном}, \\ \varepsilon & \text{при } m \text{ четном}, \end{cases}$$

$$\varepsilon = (-1)^{\frac{m}{2}} \operatorname{sign} \frac{D_{h+m+1}}{D_h}.$$

В случае $D_r = 0$ сказанное остается верным, если заменить D_r на $D^{(r)} \neq 0$ согласно (2.188) ([6], стр. 277, теорема 24).

ГЛАВА III

АЛГЕБРА МНОГОЧЛЕНОВ [11]

§ 1. Общие свойства многочленов

1. Определения. Примеры. *Многочленом* (или *полиномом*) *степени n* (n — неотрицательное целое число) от переменного x над числовым полем K (см. введение) называется выражение вида

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n,$$

где $a_0, \ldots, a_n$ — числа из K, причем $a_0 \neq 0$ *). Многочлены нулевой степени — это просто отличные от нуля числа. Число нуль также считается многочленом; его степень не определена.

Обычные обозначения для многочленов: $f(x)$, $g(x)$, $\varphi(x)$ и т. д. Два многочлена считаются *равными*, если у них равны коэффициенты при одинаковых степенях переменного.

Многочлены можно складывать, вычитать, умножать друг на друга. *Суммой* (*разностью*) двух многочленов $f(x)$ и $g(x)$ называется многочлен, у которого коэффициент при каждой степени переменного x равен сумме (разности) коэффициентов при той же степени x в многочленах $f(x)$ и $g(x)$.

Пример 1. $f(x) = x^3 - 4x^2 + x + 1$, $g(x) = x^3 + 1$;

$$f(x) + g(x) = x^3 - 3x^2 + x + 2; \quad f(x) - g(x) = x^3 - 5x^2 + x.$$

Чтобы *умножить* многочлен $f(x)$ на многочлен $g(x)$, нужно каждый член многочлена $f(x)$ умножить на каждый член $g(x)$, сложить полученные произведения и привести подобные члены. Степень произведения двух многочленов всегда равна сумме степеней сомножителей.

*) Аналогично определяются многочлены над произвольным полем P, не обязательно с числовыми элементами (гл. IV, § 3).

Пример 2. $f(x)=x^2-3x+2$, $g(x)=x^2+2x-3$;

$$f(x)g(x)=x^4+2x^3-3x^2-3x^3-6x^2+9x+2x^2+4x-6=$$
$$=x^4-x^3-7x^2+13x-6.$$

Деление (нацело) одного многочлена на другой не всегда возможно.

2. Деление с остатком. Любой многочлен $f(x)$ можно *разделить* на любой другой многочлен $g(x)$ ($g(x)\neq 0$) *с остатком*, т. е. можно единственным образом представить многочлен $f(x)$ в виде $f(x)=g(x)q(x)+r(x)$, где $q(x)$ и $r(x)$ — какие-то многочлены, причем степень $r(x)$ меньше степени $g(x)$ или же $r(x)=0$; $q(x)$ называется *частным* от деления многочлена $f(x)$ на $g(x)$, $r(x)$ называется *остатком* от этого деления. Если $r(x)=0$, то говорят, что $f(x)$ *делится на* $g(x)$ *без остатка* (нацело).

Если $g(x)$ — многочлен нулевой степени, т. е. число, то на него делится без остатка любой многочлен $f(x)$.

Тогда и только тогда $f(x)$ делится без остатка на $g(x)$ и одновременно $g(x)$ делится без остатка на $f(x)$, когда $g(x)=cf(x)$, где $c\neq 0$ — некоторое число.

Пример 1. $f(x)=x^3-1$, $g(x)=x^2+x+1$.
Здесь $x^3-1=(x^2+x+1)(x-1)$, т. е. $f(x)$ делится на $g(x)$ без остатка ($q(x)=x-1$, $r(x)=0$).
Пример 2. $f(x)=2x^2-3x+1$, $g(x)=x^2-4x-1$.
Здесь $f(x)=g(x)q(x)+r(x)$, где $q(x)=2$, $r(x)=5x+3$.

Если степень $f(x)$ меньше степени $g(x)$, то частным от деления $f(x)$ на $g(x)$ является многочлен $q(x)=0$, а остатком — сам многочлен $f(x)$. В общем случае для нахождения частного и остатка можно воспользоваться известным из курса средней школы *правилом деления «уголком»* (оно применимо к многочленам с коэффициентами из любого поля).

3. Схема Горнера. Если многочлен $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ нужно разделить с остатком на многочлен $g(x)=x-b$, то это проще всего сделать с помощью *схемы Горнера*, которая составляется так: выписываются в строку коэффициенты многочлена $f(x)$, начиная с a_0; слева пишется число b; затем под a_0 пишется еще раз число a_0, под a_1 пишется число $a_0b+a_1=b_1$, под a_2 — число $b_1b+a_2=b_2$ и т. д.; под a_n пишется число $b_{n-1}b+a_n=b_n$.

b_n является остатком от деления $f(x)$ на $x-b$, a_0, $b_1, \ldots, b_{n-1}$ — коэффициенты частного.

Пример 1. $f(x)=2x^3-3x+5$, $g(x)=x-4$.
Составляем схему Горнера:

	2	0	−3	5
4	2	8	29	121

$$f(x)=(x-4)(2x^2+8x+29)+121.$$

Если нужно разделить с остатком многочлен $f(x)$ на многочлен $g(x)=ax+b$, $a\neq0$, то можно сначала разделить его по методу Горнера на многочлен $g_1(x)=x-\left(-\frac{b}{a}\right)$; если $q(x)$ и r — частное и остаток от деления $f(x)$ на $g(x)$, а $q_1(x)$, r_1 — частное и остаток от деления $f(x)$ на $g_1(x)$, то $r=r_1$, а $q(x)=\frac{1}{a}q_1(x)$.

Пример 2. $f(x)=x^3-2x^2-10x+3$, $g(x)=2x+5$.

Берем $g_1(x)=x+\frac{5}{2}$ и выписываем схему Горнера деления $f(x)$ на $g_1(x)$:

	1	−2	−10	3
$-\frac{5}{2}$	1	$-\frac{9}{2}$	$\frac{5}{4}$	$-\frac{1}{8}$

Получаем $f(x)=g(x)\,q(x)+r$, где

$$q(x)=\frac{1}{2}\left(x^2-\frac{9}{2}x+\frac{5}{4}\right)=\frac{1}{2}x^2-\frac{9}{4}x+\frac{5}{8},\quad r=-\frac{1}{8}.$$

Остаток от деления произвольного многочлена $f(x)$ на многочлен $g(x)=x-b$ является числом, равным значению многочлена $f(x)$ при $x=b$ (теорема Безу). В силу этого метод Горнера удобен также для быстрого вычисления значения многочлена при данном значении неизвестного.

Пример 3. $f(x)=x^4-2x^3+x^2+x+1$. Найти значение $f(x)$ при $x=-3$.
Составляем схему Горнера:

	1	−2	1	1	1
−3	1	−5	16	−47	142

$$f(-3)=142.$$

4. Разложение многочлена по степеням разности ([23], стр. 97—98; [31], стр. 77—79). Для любого многочлена

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

и любого числа c можно написать *разложение* $f(x)$ *по степеням разности* $x-c$:

$$f(x)=b_0(x-c)^n+b_1(x-c)^{n-1}+\ldots+b_{n-1}(x-c)+b_n.$$

Чтобы найти коэффициенты этого разложения, нужно сначала разделить с остатком $f(x)$ на $x-c$. В остатке получится b_n, частное будет каким-то многочленом $q(x)$. Затем нужно $q(x)$ разделить на $x-c$; в остатке получится b_{n-1}, частное — $q_1(x)$. Затем делим $q_1(x)$ на $x-c$, в остатке получается b_{n-2} и т. д. Вычисления удобно производить по схеме Горнера (п. 3).

Пример 1. $f(x)=x^4-5x^3-3x^2+9$, $c=3$.

Составляем схему Горнера, где в первой строке выписываем коэффициенты многочлена $f(x)$, во второй получаем коэффициенты частного $q(x)$ и остаток b_4 от деления $f(x)$ на $x-3$, в третьей строке получаем коэффициенты частного $q_1(x)$ и остаток b_3 от деления $q(x)$ на $x-3$ и т. д.:

	1	—5	—3	0	9
3	1	—2	—9	—27	—72
	1	1	—6	—45	
	1	4	6		
	1	7			
	1				

$$f(x)=(x-3)^4+7(x-3)^3+6(x-3)^2-45(x-3)-72.$$

Коэффициенты разложения многочлена $f(x)$ по степеням разности $x-c$ связаны со значениями этого многочлена и его производных*) при $x=c$:

$$b_n=f(c),\quad b_{n-1}=\frac{f'(c)}{1!},\quad b_{n-2}=\frac{f''(c)}{2!},\ \ldots,\ b_0=\frac{f^{(n)}(c)}{n!}.$$

Поэтому схема Горнера, выписанная в предыдущем примере, позволяет также найти значения производных многочлена $f(x)$

*) *Производной* многочлена $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ с коэффициентами из любого поля называется многочлен

$$f'(x)=na_0x^{n-1}+(n-1)a_1x^{n-2}+\ldots+a_{n-1}.$$

при $x=c$ (в указанном примере эти значения: $f'(3)=-45$, $f''(3)=6\cdot 2!=12$, $f'''(3)=7\cdot 3!=42$, $f^{\mathrm{IV}}(3)=4!=24$).

Разложение многочлена по степеням разности $x-c$ может быть использовано при разложении дроби, знаменатель которой есть степень линейного двучлена, на простейшие (§ 3, п. 9).

Пример 2. $\Phi(x)=\dfrac{x^2+x+1}{(x+2)^4}$.

Берем многочлен $f(x)=x^2+x+1$ и разлагаем его по степеням разности $x-(-2)=x+2$:

$$\begin{array}{r|rrr}
 & 1 & 1 & 1 \\ \hline
-2 & 1 & -1 & 3 \\
 & 1 & -3 & \\
 & 1 & &
\end{array}$$

$$f(x)=(x+2)^2-3(x+2)+3.$$

Получаем $\Phi(x)=\dfrac{f(x)}{(x+2)^4}=\dfrac{1}{(x+2)^2}-\dfrac{3}{(x+2)^3}+\dfrac{3}{(x+2)^4}$.

5. Наибольший общий делитель двух многочленов. Если каждый из многочленов $f(x)$ и $g(x)$ делится без остатка на многочлен $\varphi(x)$, то $\varphi(x)$ называется *общим делителем* $f(x)$ и $g(x)$. *Наибольшим общим делителем* (н. о. д.) многочленов $f(x)$ и $g(x)$ называется такой их общий делитель $d(x)$, который делится на все другие общие делители многочленов $f(x)$ и $g(x)$ (обозначение: $d(x)=(f(x),\ g(x))$). Для любых двух многочленов, не равных одновременно нулю, н. о. д. существует и определен однозначно с точностью до постоянных отличных от нуля множителей. Из всех наибольших общих делителей многочленов $f(x)$ и $g(x)$ обычно выбирается тот, у которого старший коэффициент равен 1.

Пример 1. $f(x)=x^2-ix$, $g(x)=x^2+1$, $d(x)=x-i$.
Пример 2. $f(x)=6x+12$, $g(x)=4x-2$, $d(x)=1$.

Если известны разложения многочленов $f(x)$ и $g(x)$ на линейные множители (§ 3, п. 6), причем

$$f(x)=a_0(x-\alpha_1)^{k_1}\ldots(x-\alpha_p)^{k_p}(x-\beta_1)^{u_1}\ldots(x-\beta_t)^{u_t};$$
$$g(x)=b_0(x-\alpha_1)^{l_1}\ldots(x-\alpha_p)^{l_p}(x-\gamma_1)^{v_1}\ldots(x-\gamma_s)^{v_s}$$

(числа α_j, β_q, γ_r все различны между собой), то $d(x) = (x-\alpha_1)^{m_1} \ldots (x-\alpha_p)^{m_p}$, где для каждого j ($j = 1, \ldots, p$) m_j — наименьшее из чисел k_j, l_j.

Пример 3. $f(x) = (x-1)^3 (x+2)^2 (x-5)^2$,
$g(x) = (x-1)(x+2)^4 (x+7)(x+1)^2$; $d(x) = (x-1)(x+2)^2$.

Если разложения многочленов $f(x)$ и $g(x)$ на линейные множители не известны, то для нахождения их наибольшего общего делителя применяется **алгоритм Евклида**. Многочлен $f(x)$ делится с остатком на $g(x)$ (п. 2); получается остаток $r_1(x)$ (быть может, равный нулю); затем, если $r_1(x) \neq 0$, $g(x)$ делится с остатком на $r_1(x)$; получается остаток $r_2(x)$; если $r_2(x) \neq 0$, то $r_1(x)$ делится с остатком на $r_2(x)$, и т. д. до получения остатка, равного нулю. Последний не равный нулю остаток $r_k(x)$ является н. о. д. многочленов $f(x)$ и $g(x)$. Если уже $r_1(x) = 0$, то наибольшим общим делителем многочленов $f(x)$ и $g(x)$ является $g(x)$.

Пример 4. $f(x) = 3x^3 - 2x^2 + x + 4$, $g(x) = x^2 - x - 1$.
Делим с остатком $f(x)$ на $g(x)$:

$$f(x) = g(x)(3x+1) + 5(x+1).$$

Получаем $r_1(x) = 5(x+1) \neq 0$. Делим $g(x)$ на $r_1(x)$:

$$g(x) = r_1(x)\frac{1}{5}(x-2) + 1; \quad r_2(x) = 1 \neq 0.$$

$r_1(x)$ делится на $r_2(x)$ без остатка. Итак, $(f(x), g(x)) = 1$.

Применяя алгоритм Евклида, можно умножать делимое и делитель на любые не равные нулю числа, причем это можно делать даже в процессе самого деления. Этим пользуются при действиях с многочленами с целыми коэффициентами, чтобы не иметь дела с дробями.

Пример 5. $f(x) = 2x^4 + 2x^3 + x^2 - x - 1$, $g(x) = 3x^4 + 2x^2 - x + 2$.

Делим с остатком (обычным образом) многочлен $3f(x)$ на $g(x)$ и получаем в остатке многочлен $r_1(x) = 6x^3 - x^2 - x - 7$. Затем делим $2g(x)$ на $r_1(x)$:

$$\begin{array}{r|l}
6x^4 \qquad + 4x^2 - 2x + 4 & 6x^3 - x^2 - x - 7 \\ \cline{2-2}
\underline{6x^4 - x^3 - x^2 - 7x \qquad} & x \,\|\, + 1 \\
\underline{x^3 + 5x^2 + 5x + 4} & \\
\underline{6x^3 + 30x^2 + 30x + 24} & \\
\underline{6x^3 - x^2 - x - 7} & \\
31x^2 + 31x + 31 &
\end{array}$$

(Здесь в процессе деления произведено умножение на 6; символом ‖ отделена та часть частного, в которой произошло искажение.) Остаток (после сокращения на 31) равен

$$r_2(x) = x^2 + x + 1;$$

$r_1(x)$ делится на $r_2(x)$ без остатка, следовательно, $r_2(x) = (f(x), g(x))$.

6. Взаимно простые многочлены. Два многочлена, н.о.д. которых равен 1, называются *взаимно простыми*. Многочлены взаимно просты тогда и только тогда, когда у них нет ни одного общего корня (п. 10).

Пример 1. $f(x) = x^3 - 3x^2 + 5$, $g(x) = x^2 - 4x + 3$. Здесь $d(x) = 1$, многочлены $f(x)$ и $g(x)$ взаимно просты.

Если $d(x) = (f(x), g(x))$, то многочлены $f_1(x) = \frac{f(x)}{d(x)}$ и $g_1(x) = \frac{g(x)}{d(x)}$ взаимно просты. Если многочлен $\varphi(x)$ взаимно прост с каждым из многочленов $f(x)$ и $g(x)$, то он взаимно прост и с их произведением. Если произведение многочленов $f(x)$ и $g(x)$ делится на $\varphi(x)$, причем $f(x)$ и $\varphi(x)$ взаимно просты, то $g(x)$ делится на $\varphi(x)$. Если многочлен $f(x)$ делится на каждый из многочленов $\varphi(x)$ и $\psi(x)$, которые между собой взаимно просты, то $f(x)$ делится и на их произведение.

7. Представление $d(x) = (f(x), g(x))$ в виде $d(x) = f(x)u(x) + g(x)v(x)$. Наибольший общий делитель $d(x)$ многочленов $f(x)$ и $g(x)$ всегда можно представить в виде

$$d(x) = f(x)u(x) + g(x)v(x), \tag{3.1}$$

где $u(x)$ и $v(x)$ — некоторые многочлены, причем это можно сделать даже бесчисленным множеством различных способов. Если, в частности, многочлены $f(x)$ и $g(x)$ взаимно просты (п. 6), то получаем $1 = f(x)u(x) + g(x)v(x)$.

Если степени многочленов $f(x)$ и $g(x)$ больше нуля, то многочлены $u(x)$ и $v(x)$ в равенстве (3.1) можно подобрать так, что степень $u(x)$ будет меньше степени $g(x)$, а степень $v(x)$ — меньше степени $f(x)$.

При практическом отыскании многочленов $u(x)$ и $v(x)$ удобнее вместо $f(x)$ и $g(x)$ взять многочлены $f_1(x) = \frac{f(x)}{d(x)}$, $g_1(x) = \frac{g(x)}{d(x)}$ и сначала подобрать $u(x)$, $v(x)$ так, чтобы выполнялось равенство

$$1 = f_1(x)u(x) + g_1(x)v(x). \tag{3.2}$$

Это делается следующим образом. Если $f_1(x)$ или $g_1(x)$ — многочлен нулевой степени, то многочлены $u(x)$ и $v(x)$ просто подбираются: например, при $g_1(x)=a\neq 0$ можно положить $u(x)=0$, $v(x)=1/a$. Если же степени $f_1(x)$ и $g_1(x)$ положительны, то нужно написать $u(x)$ и $v(x)$ с неопределенными коэффициентами (считая степень $u(x)$ меньше степени $g_1(x)$, степень $v(x)$ меньше степени $f_1(x)$) и приравнять коэффициенты при одинаковых степенях x в левой и правой частях равенства (3.2) (или же можно придавать переменному x различные числовые значения и приравнивать единице получающиеся при этом значения многочлена $f_1(x)u(x)+g_1(x)v(x)$); это даст систему линейных уравнений, решая которую (гл. I, § 2, пп. 1—3) найдем неизвестные коэффициенты многочленов $u(x)$ и $v(x)$. Если теперь умножить обе части равенства (3.2) на $d(x)$, то получится равенство (3.1) (с теми же многочленами $u(x)$ и $v(x)$, что и в (3.2)). При таком способе отыскания $u(x)$ и $v(x)$ степень $u(x)$ получается меньше разности степеней $g(x)$ и $d(x)$, степень $v(x)$ — меньше разности степеней $f(x)$ и $d(x)$ (если только указанные разности $\neq 0$). При этом условии на степени многочлены $u(x)$ и $v(x)$, входящие в равенство (3.1), определены уже однозначно.

Пример 1. $f(x)=x^7-2x^6-x^5-11x^4-6x^3-4x^2+5x+14$, $g(x)=x^2+x+2$.

Здесь $d(x)=g(x)=x^2+x+2$, следовательно, можно взять $u(x)=0$, $v(x)=1$.

Пример 2. $f(x)=3x^5-4x^4+x^3-3x^2+4x-1$, $g(x)=3x^5+5x^4+x^3-x^2-3x+1$.

Здесь $d(x)=3x^3+2x^2+2x-1$, $f_1(x)=x^2-2x+1$, $g_1(x)=x^2+x-1$. Пишем для многочленов $f_1(x)$, $g_1(x)$ равенство (3.2) с неопределенными коэффициентами $u(x)$ и $v(x)$:

$$1=(x^2-2x+1)(ax+b)+(x^2+x-1)(a_1x+b_1).$$

Приравниваем коэффициенты при одинаковых степенях x и получаем систему уравнений

$$\begin{cases} a + a_1=0, \\ b-2a+b_1+a_1=0, \\ -2b+a+b_1-a_1=0, \\ b - b_1 = 1. \end{cases}$$

Решая эту систему, находим: $a=3$, $a_1=-3$, $b=5$, $b_1=4$, т. е.

$$d(x)=f(x)(3x+5)+g(x)(-3x+4).$$

Пример 3. $f(x)=2x^4-4x^3-3x^2+7x-2$, $g(x)=6x^3+4x^2-5x+1$.

Здесь $d(x)=2x^2+2x-1$, $f_1(x)=x^2-3x+2$, $g_1(x)=3x-1$. Пишем равенство (3.2):

$$1=(x^2-3x+2)a+(3x-1)(bx+c).$$

Придавая x последовательно значения 0, 1, 2, получаем систему уравнений

$$\begin{cases} 2a \quad -c=1, \\ \quad 2b+2c=1, \\ \quad 10b+5c=1. \end{cases}$$

Решая эту систему, находим: $a=0{,}9$, $b=-0{,}3$, $c=0{,}8$, откуда $d(x)=f(x)\,0{,}9+g(x)(-0{,}3x+0{,}8)$.

Чтобы многочлен $h(x)$, отличный от н. о. д. $d(x)$ многочленов $f(x)$ и $g(x)$, можно было представить в виде

$$h(x)=f(x)u_1(x)+g(x)v_1(x), \tag{3.3}$$

где $u_1(x)$, $v_1(x)$ — некоторые многочлены, необходимо и достаточно, чтобы $h(x)$ делился на $d(x)$. Если при этом степень $h(x)$ меньше суммы степеней $f(x)$ и $g(x)$, то многочлены $u_1(x)$ и $v_1(x)$ можно подобрать так, что степень $u_1(x)$ будет меньше степени $g(x)$, а степень $v_1(x)$ — меньше степени $f(x)$. В этом случае $u_1(x)$ и $v_1(x)$ можно найти, записав их с неопределенными коэффициентами и приравнивая затем коэффициенты при одинаковых степенях x в левой и правой частях равенства (3.3) (единственности $u_1(x)$, $v_1(x)$ здесь может и не быть). Если же степень $h(x)$ больше или равна сумме степеней $f(x)$ и $g(x)$, причем $h(x)=d(x)k(x)$, то нужно сначала представить $d(x)$ в виде (3.1), а затем положить $u_1(x)=u(x)k(x)$, $v_1(x)=v(x)k(x)$.

Пример 4. $f(x)=x^3-x^2-x+1$, $g(x)=x^2+x-2$, $h(x)=x^4-1$.

Здесь $d(x)=x-1$, $h(x)=d(x)k(x)$, где $k(x)=x^3+x^2+x+1$. Так как степень $h(x)$ меньше суммы степеней $f(x)$ и $g(x)$, то пишем равенство

$$x^4-1=(x^3-x^2-x+1)(ax+b)+(x^2+x-2)(a_1x^2+b_1x+c_1).$$

Приравниваем коэффициенты при одинаковых степенях x:

$$\begin{cases} a + a_1 = 1, \\ -a+b+a_1+b_1 = 0, \\ -a-b-2a_1+b_1+c_1=0, \\ a-b-2b_1+c_1=0, \\ b-2c_1=-1. \end{cases}$$

Полученная система уравнений имеет бесчисленное множество решений. Любое из них дает коэффициенты таких многочленов $u_1(x)$, $v_1(x)$, что выполнено равенство (3.3). Например, можно взять

$$u_1(x)=x+\frac{1}{3}, \quad v_1(x)=\frac{2}{3}(x+1).$$

8. Уничтожение иррациональности в знаменателе ([23], стр. 227—228). Представление н.о.д. многочленов $f(x)$ и $g(x)$ в виде (3.1) используется при решении следующей задачи.

Имеем выражение вида $\frac{m(\alpha)}{g(\alpha)}$, где $m(x)$ и $g(x)$ — многочлены с рациональными коэффициентами, а α — корень многочлена $f(x)$ тоже с рациональными коэффициентами, взаимно простого с $g(x)$ (п.6). Требуется найти такой многочлен $N(x)$ с рациональными коэффициентами, что $\frac{m(\alpha)}{g(\alpha)}=N(\alpha)$.

Для решения задачи представляем $(f(x), g(x))=1$ в виде (3.1) и затем в качестве $N(x)$ берем многочлен $m(x)v(x)$ (или, лучше, остаток от деления этого многочлена на $f(x)$).

Пример. Уничтожить иррациональность в знаменателе выражения $\frac{10\sqrt[3]{2}-10}{\sqrt[3]{4}-2\sqrt[3]{2}+2}$.

Здесь $\alpha=\sqrt[3]{2}$, $m(x)=10x-10$, $g(x)=x^2-2x+2$, $f(x)=$ $=x^3-2$, $(f(x), g(x))=1=f(x)(-0,1x-0,1)+g(x)(0,1x^2+0,3x+0,4)$, т. е. $v(x)=0,1x^2+0,3x+0,4$, $m(x)v(x)=x^3+2x^2+x-4=$ $=(x^3-2)\cdot 1+2x^2+x-2$. Полагаем $N(x)=2x^2+x-2$ и получаем $\frac{10\sqrt[3]{2}-10}{\sqrt[3]{4}-2\sqrt[3]{2}+2}=2\sqrt[3]{4}+\sqrt[3]{2}-2$.

Другой способ уничтожения иррациональности в знаменателе, использующий симметрические многочлены, см. в [23], стр. 227.

9. Наибольший общий делитель нескольких многочленов. Если дано несколько многочленов $f_1(x), \ldots, f_s(x)$, то их *наибольшим общим делителем* называется многочлен $D(x)$, на который делятся все данные многочлены и который сам делится на любой общий делитель многочленов $f_1(x), \ldots$ $\ldots, f_s(x)$. Чтобы найти $D(x)$, нужно сначала найти н.о.д. $d_1(x)$ многочленов $f_1(x)$ и $f_2(x)$, затем н.о.д. $d_2(x)$ многочленов $d_1(x)$ и $f_3(x)$, н.о.д. $d_3(x)$ многочленов $d_2(x)$ и $f_4(x)$ и т. д. до $d_{s-1}(x) = (d_{s-2}(x), f_s(x)) = D(x)$.

Пример. $f_1(x) = x^2 - 2x + 2$, $f_2(x) = x^2 + 2ix - 2$, $f_3(x) = 2x^3 - 3x^2 - 2x + 1$, $f_4(x) = x^3 - 2x^2 + x + 1$.

Здесь $d_1(x) = x - (1 - i)$, $d_2(x) = 1$, $d_3(x) = D(x) = 1$.

10. Корни многочлена. Отделение кратных корней. Если при значении $x = c$ многочлен $f(x)$ принимает значение $f(c) = 0$, то число c называется *корнем* этого многочлена. Всякий многочлен ненулевой степени с любыми числовыми коэффициентами имеет хотя бы один корень, действительный или комплексный (основная теорема алгебры комплексных чисел). Число c тогда и только тогда является корнем многочлена $f(x)$, когда $f(x)$ делится без остатка на $x - c$. Если при этом $f(x)$ делится без остатка на $(x - c)^k$ $(k \geqslant 1)$, но уже не делится на $(x - c)^{k+1}$, то c называется *k-кратным корнем* (или *корнем кратности k*) многочлена $f(x)$. Корни кратности $k = 1$ называются *простыми* корнями многочлена. Если каждый корень считать столько раз, какова его кратность, то получится, что каждый многочлен степени n $(n \geqslant 1)$ с любыми числовыми коэффициентами имеет n корней.

Чтобы проверить, будет ли число c корнем многочлена $f(x)$ и какой кратности, можно воспользоваться схемой Горнера (п. 3), где сначала $f(x)$ делится на $x - c$, затем, если остаток равен нулю, полученное частное делится снова на $x - c$ и т. д. до получения ненулевого остатка.

Пример 1. $f(x) = x^4 - 7x^3 + 9x^2 + 8x + 16$, $c = 4$.

Составляем схему Горнера:

	1	−7	9	8	16
4	1	−3	−3	−4	0
	1	1	1	0	
	1	5	21		

$c = 4$ — корень кратности 2.

Число c является k-кратным корнем многочлена $f(x)$ тогда и только тогда, когда многочлен $f(x)$ и его производные *) до порядка $k-1$ включительно обращаются в нуль при $x=c$, а $f^{(k)}(c) \neq 0$. Корень кратности $k>1$ многочлена является корнем кратности $k-1$ для его первой производной.

Если разделить многочлен $f(x)$ на наибольший общий делитель $d(x)$ этого многочлена и его производной $f'(x)$, то получится многочлен $\frac{f(x)}{d(x)}=\varphi(x)$, имеющий те же корни, что и многочлен $f(x)$, но только первой кратности. Это позволяет найти все корни многочлена $f(x)$ путем отыскания всех корней многочлена $\varphi(x)$, степень которого меньше степени $f(x)$, если только $d(x)$ — многочлен ненулевой степени.

Пример 2. $f(x)=x^5-5x^4-5x^3+45x^2-108$;

$$f'(x)=5x^4-20x^3-15x^2+90x; \quad d(x)=x^3-4x^2-3x+18;$$

$$\varphi(x)=\frac{f(x)}{d(x)}=x^2-x-6.$$

Корни многочлена $\varphi(x)$ равны $\alpha_1=-2$, $\alpha_2=3$. Эти же числа (и только они) являются корнями $f(x)$; их кратности определяются по схеме Горнера (стр. 185).

Укажем еще один способ, позволяющий иногда найти все корни многочлена (и заодно определить их кратности). Пусть $d_1(x)$ — н.о.д. многочлена $f(x)$ и его производной $f'(x)$; $d_2(x)$ — н.о.д. многочлена $d_1(x)$ и его производной $d_1'(x)$; $d_3(x)$ — н.о.д. $d_2(x)$ и $d_2'(x)$ и т. д., пока получится многочлен $d_s(x)$ нулевой степени. Положим

$$v_1(x)=\frac{f(x)}{d_1(x)}, \quad v_2(x)=\frac{d_1(x)}{d_2(x)}, \quad \ldots, \quad v_s(x)=\frac{d_{s-1}(x)}{d_s(x)}$$

и затем

$$F_1(x)=\frac{v_1(x)}{v_2(x)}, \quad F_2(x)=\frac{v_2(x)}{v_3(x)}, \quad \ldots, \quad F_s(x)=v_s(x).$$

Все корни многочленов $F_1(x)$, $F_2(x)$, ..., $F_s(x)$ простые, причем корнями $F_1(x)$ являются все простые корни многочлена $f(x)$ (и только они), корнями $F_2(x)$ являются все двукратные корни $f(x)$ и т. д., корнями $F_s(x)$ являются s-кратные корни $f(x)$; корней выше s-й кратности многочлен $f(x)$ не имеет ([23], стр. 101—102; [13], стр. 165).

*) См. сноску на стр. 178.

Пример 3. $f(x)=x^5+6x^4+13x^3+14x^2+12x+8$.
Здесь $d_1(x)=x^2+4x+4$, $d_2(x)=x+2$, $d_3(x)=1$;

$$v_1(x)=x^3+2x^2+x+2,\quad v_2(x)=x+2,\quad v_3(x)=x+2;$$
$$F_1(x)=x^2+1,\quad F_2(x)=1,\quad F_3(x)=x+2.$$

Многочлен $f(x)$ имеет простые корни $\pm i$ и 3-кратный корень -2.

11. Формулы Виета. Корни $\alpha_1, \ldots, \alpha_n$ многочлена

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n \quad (a_0\neq 0)$$

связаны с его коэффициентами по формулам (формулы Виета):

$$\alpha_1+\alpha_2+\ldots+\alpha_n=-\frac{a_1}{a_0},$$
$$\alpha_1\alpha_2+\alpha_1\alpha_3+\ldots+\alpha_1\alpha_n+\alpha_2\alpha_3+\ldots+\alpha_{n-1}\alpha_n=\frac{a_2}{a_0},$$
$$\alpha_1\alpha_2\alpha_3+\alpha_1\alpha_2\alpha_4+\ldots+\alpha_{n-2}\alpha_{n-1}\alpha_n=-\frac{a_3}{a_0},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_1\alpha_2\ldots\alpha_{n-1}+\alpha_1\alpha_2\ldots\alpha_{n-2}\alpha_n+\ldots+\alpha_2\alpha_3\ldots\alpha_n=(-1)^{n-1}\frac{a_{n-1}}{a_0},$$
$$\alpha_1\alpha_2\ldots\alpha_n=(-1)^n\frac{a_n}{a_0},$$

т. е. сумма произведений корней во всевозможных сочетаниях, содержащих каждое k корней, равна $(-1)^k a_k/a_0$ ($k=1, 2, \ldots, n$). Пользуясь формулами Виета, легко составить многочлен, имеющий наперед заданные корни.

Пример. Пусть корни многочлена $\alpha_1=5$, $\alpha_2=i$, $\alpha_3=i$. Многочленом третьей степени, имеющим эти корни, будет, например, многочлен с коэффициентами $a_0=1$, $a_1=-(5+i+i)=-5-2i$, $a_2=5i+5i+ii=-1+10i$, $a_3=-(5ii)=5$, т.е. $f(x)=x^3+(-5-2i)x^2+(-1+10i)x+5$.

12. Общие корни двух многочленов. Результант. Дискриминант. Два многочлена $f(x)$ и $g(x)$ тогда и только тогда имеют общие корни, когда их наибольший общий делитель $d(x)$ отличен от постоянной. В этом случае общими корнями $f(x)$ и $g(x)$ являются в точности все корни $d(x)$.

Пример 1. $f(x)=2x^5+x^4+3x^3+x-1$, $g(x)=2x^4-5x^3+4x^2-5x+2$.
Здесь $d(x)=2x^3-x^2+2x-1=(2x-1)(x^2+1)$, общими корнями $f(x)$ и $g(x)$ являются числа $\frac{1}{2}$, i, $-i$.

Укажем еще один способ, позволяющий установить наличие общих корней у многочленов

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$$

и

$$g(x)=b_0x^s+b_1x^{s-1}+\ldots+b_{s-1}x+b_s$$

($a_0\neq 0$, $b_0\neq 0$). Составим *результант* заданных многочленов; так называется определитель *)

$$R(f,\ g)=\begin{vmatrix} a_0 & a_1 & \ldots & a_n & 0 & \ldots & 0 \\ 0 & a_0 & a_1 & \ldots & a_n & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & a_0 & a_1 \ldots & a_n \\ b_0 & b_1 & \ldots & b_s & 0 & \ldots & 0 \\ 0 & b_0 & b_1 & \ldots & b_s & 0 \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & b_0 & b_1 \ldots & b_s \end{vmatrix}\begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} s \text{ строк,} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \text{ строк;} \end{matrix}$$

он равен нулю тогда и только тогда, когда $f(x)$ и $g(x)$ имеют хотя бы один общий корень.

Пример 2. $f(x)=x^3-3x^2+1$, $g(x)=2x^4-5x^2-x-1$. Здесь

$$R(f,\ g)=\begin{vmatrix} 1 & -3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -3 & 0 & 1 \\ 2 & 0 & -5 & -1 & -1 & 0 & 0 \\ 0 & 2 & 0 & -5 & -1 & -1 & 0 \\ 0 & 0 & 2 & 0 & -5 & -1 & -1 \end{vmatrix}=543\neq 0,$$

$f(x)$ и $g(x)$ общих корней не имеют.

Другие выражения для результанта многочленов $f(x)$ и $g(x)$:

$$R(f,\ g)=a_0^sb_0^n\prod_{i=1}^{n}\prod_{j=1}^{s}(\alpha_i-\beta_j)$$

*) Определитель $(n+s)$-го порядка (см. гл. I, § 1).

(α_i, β_j — корни соответственно многочленов $f(x)$ и $g(x)$);

$$R(f, g)=a_0^s\prod_{i=1}^{n} g(\alpha_i);$$

$$R(f, g)=(-1)^{ns}R(g, f)=(-1)^{ns}b_0^n\prod_{j=1}^{s} f(\beta_j).$$

Результант многочленов используется при решении систем алгебраических уравнений с несколькими неизвестными (§ 5, п. 4).

Дискриминантом многочлена $f(x)=a_0x^n+a_1x^{n-1}+\ldots$ $\ldots+a_{n-1}x+a_n$, имеющего корнями числа $\alpha_1, \alpha_2, \ldots, \alpha_n$, называется произведение

$$D(f)=a_0^{2n-2}\prod_{n\geqslant i>j\geqslant 1}(\alpha_i-\alpha_j)^2.$$

Дискриминант тогда и только тогда равен нулю, когда среди корней многочлена имеются равные (т. е. когда многочлен имеет хотя бы один кратный корень (п. 10)).

Дискриминант связан с результантом многочлена $f(x)$ и его производной $f'(x)$ по формуле

$$R(f, f')=(-1)^{\frac{n(n-1)}{2}}a_0D(f),$$

позволяющей выразить дискриминант многочлена через его коэффициенты.

Пример 3. $f(x)=x^3-3x+7$.
Здесь $f'(x)=3x^2-3$,

$$(-1)^3D(f)=R(f, f')=\begin{vmatrix} 1 & 0 & -3 & 7 & 0\\ 0 & 1 & 0 & -3 & 7\\ 3 & 0 & -3 & 0 & 0\\ 0 & 3 & 0 & -3 & 0\\ 0 & 0 & 3 & 0 & -3\end{vmatrix}=1215.$$

$D(f)=-1215$; $f(x)$ кратных корней не имеет.

Можно выразить $D(f)$ через коэффициенты многочлена $f(x)$ и другим путем, пользуясь тем, что $D(f)$ является симметрическим многочленом от корней $f(x)$ (§ 5, п. 2).

Выражение дискриминанта через степенные суммы (§ 5, п. 3):

$$D(f)=a_0^{2n-2}\begin{vmatrix} n & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & s_4 & \dots & s_{n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{n-1} & s_n & s_{n+1} & \dots & s_{2n-2} \end{vmatrix}$$

(s_i — сумма i-х степеней корней $\alpha_1, \alpha_2, \dots, \alpha_n$ многочлена $f(x)$).

Пример 4. $f(x)=x^3+3x^2-2$.

Согласно § 5 (конец п. 2 и п. 3), находим $s_1=-3$, $s_2=9$, $s_3=-21$, $s_4=57$. Отсюда

$$D(f)=\begin{vmatrix} 3 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}=\begin{vmatrix} 3 & -3 & 9 \\ -3 & 9 & -21 \\ 9 & -21 & 57 \end{vmatrix}=108.$$

Если $f(x)$ — многочлен с действительными коэффициентами, не имеющий кратных корней, то $D(f)>0$, если число пар комплексно сопряженных корней $f(x)$ четно, и $D(f)<0$, если это число нечетно.

Пример 5. Для многочлена

$$f(x)=x^4-x^3+x^2+2x+2$$

$D(f)=4360>0$, т. е. либо все четыре корня многочлена $f(x)$ комплексные (две пары комплексно сопряженных корней), либо $f(x)$ вообще не имеет комплексных корней (число пар комплексно сопряженных корней $f(x)$ равно нулю).

Пример 6. $f(x)=2x^3+9x^2+9x+7$.

Здесь $D(f)=-9126<0$, $f(x)$ имеет одну пару комплексно сопряженных корней, т. е. имеет два комплексных и один действительный корень.

§ 2. Интерполяционные формулы Лагранжа и Ньютона

Если $b_0, b_1, \dots, b_n$ — различные между собой числа, $c_0, c_1, \dots, c_n$ — какие-то числа, не все равные нулю, то существует ровно один многочлен $f(x)$ степени не выше n со свойствами: $f(b_0)=c_0$, $f(b_1)=c_1$, $\dots$, $f(b_n)=c_n$ *). Этот

*) При $c_0=c_1=\dots=c_n=0$ всякий многочлен со свойствами $f(b_i)=c_i$ ($i=0, 1, 2, \dots, n$) или имеет степень больше n, или тождественно равен нулю.

многочлен можно найти одним из следующих трех способов:

1) Составляется система уравнений:

$$\begin{aligned}
b_0^n x_0 + b_0^{n-1}x_1 + \ldots + b_0 x_{n-1} + x_n &= c_0,\\
b_1^n x_0 + b_1^{n-1}x_1 + \ldots + b_1 x_{n-1} + x_n &= c_1,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
b_n^n x_0 + b_n^{n-1}x_1 + \ldots + b_n x_{n-1} + x_n &= c_n.
\end{aligned}$$

Если $a_0, a_1, \ldots, a_{n-1}, a_n$ — решение этой системы уравнений (гл. I, § 2, пп. 1—3), то искомый многочлен:

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

Пример 1. $b_0 = -1$, $b_1 = 1$, $b_2 = 2$, $b_3 = 3$; $c_0 = 4$, $c_1 = 2$, $c_2 = -2$, $c_3 = 0$.

Составляем систему уравнений:

$$\begin{aligned}
-x_0 + x_1 - x_2 + x_3 &= 4,\\
x_0 + x_1 + x_2 + x_3 &= 2,\\
8x_0 + 4x_1 + 2x_2 + x_3 &= -2,\\
27x_0 + 9x_1 + 3x_2 + x_3 &= 0.
\end{aligned}$$

Решая эту систему, находим: $a_0 = 1$, $a_1 = -3$, $a_2 = -2$, $a_3 = 6$. Искомый многочлен: $f(x) = x^3 - 3x^2 - 2x + 6$.

2) Формула Лагранжа:

$$f(x) = \sum_{k=0}^{n} c_k \frac{(x-b_0)\ldots(x-b_{k-1})(x-b_{k+1})\ldots(x-b_n)}{(b_k-b_0)\ldots(b_k-b_{k-1})(b_k-b_{k+1})\ldots(b_k-b_n)},$$

или, в другой записи,

$$f(x) = c_0 \frac{g_0(x)}{g'(b_0)} + c_1 \frac{g_1(x)}{g'(b_1)} + \ldots + c_n \frac{g_n(x)}{g'(b_n)}, \qquad (3.4)$$

где

$$g(x) = (x-b_0)(x-b_1)\ldots(x-b_n),$$

а $g_k(x)$, $k = 0, 1, \ldots, n$, — произведение тех же скобок, но с пропуском скобки $(x-b_k)$, $g'(x)$ — производная многочлена $g(x)$ ([23], стр. 110).

Пример 2. $b_0 = 1$, $b_1 = 3$, $b_2 = 4$; $c_0 = 2$, $c_1 = -2$, $c_2 = -1$. Искомый многочлен:

$$f(x) = 2\frac{(x-3)(x-4)}{(1-3)(1-4)} - 2\frac{(x-1)(x-4)}{(3-1)(3-4)} - \frac{(x-1)(x-3)}{(4-1)(4-3)} = x^2 - 6x + 7.$$

Формулой Лагранжа удобно пользоваться, если нужно определить не коэффициенты многочлена $f(x)$ степени не выше n со свойствами $f(b_i)=c_i$, $i=0, 1, \ldots, n$, а его значение при каком-нибудь значении $x=b_{n+1}$.

Пример 3. Зная, что $f(x)$ — многочлен не выше 3-й степени и что $f(-2)=0$, $f(1)=5$, $f(3)=-1$, $f(5)=-1$, найти $f(6)$.

В силу формулы Лагранжа имеем

$$f(6)=5\frac{(6+2)(6-3)(6-5)}{(1+2)(1-3)(1-5)}-\frac{(6+2)(6-1)(6-5)}{(3+2)(3-1)(3-5)}-\\-\frac{(6+2)(6-1)(6-3)}{(5+2)(5-1)(5-3)}=4\frac{6}{7}.$$

Формула Лагранжа дает также разложение дроби на простейшие в том случае, когда знаменатель дроби не имеет кратных корней и разлагается над данным полем на линейные множители (§ 3, п. 9). Если $b_0, b_1, \ldots, b_n$ — корни знаменателя дроби $\frac{f(x)}{g(x)}$, где $g(x)$ — многочлен $(n+1)$-й степени, $f(x)$ — многочлен степени $\leqslant n$, и если $f(b_0)=c_0$, $f(b_1)=c_1, \ldots, f(b_n)=c_n$, то выражаем $f(x)$ с помощью формулы (3.4) и получаем

$$\frac{f(x)}{g(x)}=\frac{c_0}{g'(b_0)(x-b_0)}+\frac{c_1}{g'(b_1)(x-b_1)}+\ldots+\frac{c_n}{g'(b_n)(x-b_n)}.$$

Пример 4. Разложить дробь $\Phi(x)=\frac{x+1}{x^2-x-6}$ на простейшие. Здесь

$$b_0=-2,\ b_1=3;\ c_0=-1,\ c_1=4;\ g'(b_0)=-5,\ g'(b_1)=5.$$

$$\Phi(x)=\frac{1}{5(x+2)}+\frac{4}{5(x-3)}.$$

3) Формула Ньютона ([23], стр. 111):

$$f(x)=\lambda_0+\lambda_1(x-b_0)+\lambda_2(x-b_0)(x-b_1)+\ldots\\ \ldots+\lambda_n(x-b_0)\ldots(x-b_{n-1}),$$

где λ_0 находим, полагая в обеих частях равенства $x=b_0$ ($\lambda_0=f(b_0)=c_0$), затем λ_1 находим, полагая в обеих частях равенства $x=b_1$, и т. д. до λ_n, которое находим, полагая в обеих частях равенства $x=b_n$.

Пример 5. $b_0=1$, $b_1=-2$, $b_2=3$; $c_0=4$, $c_1=7$, $c_2=12$. $f(x)=4+\lambda_1(x-1)+\lambda_2(x-1)(x+2)$.

Полагая $x=-2$, получаем: $7=4+\lambda_1(-3)$, $\lambda_1=-1$. Берем $x=3$: $12=4-(3-1)+\lambda_2(3-1)(3+2)$, $10=10\lambda_2$, $\lambda_2=1$. Искомый многочлен: $f(x)=4-(x-1)+(x-1)(x+2)=x^2+3$.

Если $b_0, b_1, \ldots, b_n$ — последовательные целые числа, то вычисления удобно производить по схеме, где в первом столбце выписываются числа $c_0, c_1, \ldots, c_n$, во втором столбце — их разности $\Delta c_0=c_1-c_0, \ldots, \Delta c_{n-1}=c_n-c_{n-1}$, в третьем столбце — разности чисел, стоящих во втором столбце, $\Delta^2 c_0=$ $=\Delta c_1-\Delta c_0, \ldots, \Delta^2 c_{n-2}=\Delta c_{n-1}-\Delta c_{n-2}$ и т. д. ([5], ч. I, стр. 97):

$$\begin{array}{ccccc}
c_0 & & & & \\
 & \Delta c_0 & & & \\
c_1 & & \Delta^2 c_0 & & \\
 & \Delta c_1 & \vdots & \vdots & \Delta^n c_0 \\
c_2 & \vdots & & & \\
\vdots & \Delta c_{n-1} & & & \\
c_n & & & &
\end{array}$$

Тогда $\lambda_0=c_0$, $\lambda_1=\Delta c_0$, $\lambda_2=\dfrac{\Delta^2 c_0}{2!}, \ldots, \lambda_n=\dfrac{\Delta^n c_0}{n!}$.

Пример 6. $b_0=-2$, $b_1=-1$, $b_2=0$, $b_3=1$, $b_4=2$; $c_0=-8$, $c_1=-3$, $c_2=4$, $c_3=13$, $c_4=24$.

Выписываем схему:

$$\begin{array}{ccccc}
-8 & & & & \\
 & 5 & & & \\
-3 & & 2 & & \\
 & 7 & & 0 & \\
4 & & 2 & & 0 \\
 & 9 & & 0 & \\
13 & & 2 & & \\
 & 11 & & & \\
24 & & & &
\end{array}$$

Искомый многочлен:

$$f(x)=-8+5(x+2)+(x+2)(x+1)=x^2+8x+4.$$

Если уже построен многочлен $f_1(x)$ степени не выше $s\ (s<n)$, принимающий при $x=b_0, b_1, \ldots, b_s$ заданные значения $c_0, c_1, \ldots, c_s$, то многочлен $f(x)$ степени не выше n со свойствами $f(b_i)=c_i$, $i=0, 1, \ldots, n$, строится по формуле

$$f(x)=f_1(x)+\lambda_{s+1}(x-b_0)\ldots(x-b_s)+\ldots$$
$$\ldots+\lambda_n(x-b_0)\ldots(x-b_{n-1}),$$

где λ_{s+1} находим, полагая $x=b_{s+1}$, и т. д. ([23], стр. 112).

Пример 7. $b_0=1$, $b_1=-2$, $b_2=3$, $b_3=0$, $b_4=4$; $c_0=4$, $c_1=7$, $c_2=12$, $c_3=3$, $c_4=-5$.

Многочлен $f_1(x)=x^2+3$ обладает свойствами $f_1(1)=4$, $f_1(-2)=7$, $f_1(3)=12$, $f_1(0)=3$. Искомый многочлен:

$$f(x)=x^2+3+\lambda_4(x-1)(x+2)(x-3)x.$$

Берем $x=4$: $-5=19+\lambda_4(4-1)(4+2)(4-3)\cdot 4$, $\lambda_4=-\frac{1}{3}$. Следовательно,

$$f(x)=-\frac{1}{3}x^4+\frac{2}{3}x^3+\frac{8}{3}x^2-2x+3.$$

Общий вид всех многочленов, принимающих при $x=b_0$, $b_1, \ldots, b_n$ соответственно значения $c_0, c_1, \ldots, c_n$:

$$F(x)=f(x)+\varphi(x)(x-b_0)\ldots(x-b_n),$$

где $f(x)$ — многочлен не выше n-й степени со свойствами $f(b_0)=c_0, \ldots, f(b_n)=c_n$, а $\varphi(x)$ — произвольный многочлен ([13], стр. 159).

§ 3. Отыскание корней многочленов. Разложение многочленов на множители

1. Решение уравнений второй степени. Корни квадратного уравнения $ax^2+bx+c=0$ с любыми действительными или комплексными коэффициентами находятся по формуле $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$. Для корней «приведенного» квадратного уравнения $x^2+px+q=0$ справедлива формула

$$x=-\frac{p}{2}\pm\sqrt{\frac{p^2}{4}-q}.$$

Пример 1. Решить уравнение $(6+3i)x^2+(5+4i)x+(1+i)=0$.

Получаем:

$$x=\frac{-5-4i\pm\sqrt{9+40i-12-36i}}{12+6i}=$$

$$=\frac{-5-4i\pm\sqrt{-3+4i}}{12+6i}=\frac{-5-4i\pm(1+2i)}{12+6i},\quad x_1=\frac{-4-2i}{12+6i}=$$

$$=-\frac{1}{3},\quad x_2=\frac{-6-6i}{12+6i}=-\frac{3}{5}-\frac{1}{5}i.$$

Пример 2. Решить уравнение $x^2 + ix - (4-2i) = 0$.

Имеем: $x = -\frac{i}{2} \pm \sqrt{\frac{-1}{4} + 4 - 2i} = -\frac{i}{2} \pm \frac{4-i}{2}$, т. е. $x_1 = 2 - i$, $x_2 = -2$.

2. Решение уравнений третьей степени. Чтобы найти все корни уравнения 3-й степени

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0,$$

удобно разделить обе части этого уравнения на a_0, в результате чего получится уравнение вида

$$x^3 + b_1x^2 + b_2x + b_3 = 0, \tag{3.5}$$

имеющее те же корни, что и исходное уравнение, а затем сделать замену неизвестного $x = y - \frac{b_1}{3}$. Эту замену проще всего осуществить, разложив многочлен $f(x) = x^3 + b_1x^2 + b_2x + b_3$ по степеням $x + \frac{b_1}{3}$ с помощью схемы Горнера (§ 1, п. 4) и заменив затем $x + \frac{b_1}{3}$ через y. В результате указанной замены получится уравнение вида

$$y^3 + py + q = 0. \tag{3.6}$$

Корни этого уравнения находятся по формуле

$$y = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{3.7}$$

(**формула Кардано**). Применяя эту формулу, нужно для каждого из трех значений корня

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{3.8}$$

брать то значение корня

$$\beta = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}, \tag{3.9}$$

для которого выполняется условие $\alpha\beta=-\frac{p}{3}$ (такое значение корня β всегда существует). Если y_1, y_2, y_3 — корни уравнения (3.6), то $x_1=y_1-\frac{b_1}{3}$, $x_2=y_2-\frac{b_1}{3}$, $x_3=y_3-\frac{b_1}{3}$ — корни уравнения (3.5).

Пример 1. Решить уравнение $x^3+3x^2-6x+20=0$.

Разлагаем многочлен, стоящий в левой части уравнения, по степеням $x+1$ и полагаем $x+1=y$. Получаем уравнение $y^3-9y+28=0$. Его корни находятся по формуле

$$y=\sqrt[3]{-14+\sqrt{196-27}}+\sqrt[3]{-14-\sqrt{196-27}}.$$

Значениями корня $\alpha=\sqrt[3]{-14+\sqrt{169}}=\sqrt[3]{-1}$ являются числа $\alpha_1=-1$, $\alpha_2=\frac{1}{2}+\frac{\sqrt{3}}{2}i$, $\alpha_3=\frac{1}{2}-\frac{\sqrt{3}}{2}i$. Соответствующие им значения второго корня: $\beta_1=-\frac{-9}{3\alpha_1}=-3$, $\beta_2=-\frac{-9}{3\alpha_2}=\frac{3}{2}-\frac{3\sqrt{3}}{2}i$, $\beta_3=-\frac{-9}{3\alpha_3}=\frac{3}{2}+\frac{3\sqrt{3}}{2}i$. Отсюда $y_1=\alpha_1+\beta_1=-4$, $y_2=\alpha_2+\beta_2=2-i\sqrt{3}$, $y_3=\alpha_3+\beta_3=2+i\sqrt{3}$. Корни исходного уравнения: $x_1=-5$, $x_2=1-i\sqrt{3}$, $x_3=1+i\sqrt{3}$.

Корни уравнения (3.6) можно также находить по формулам

$$\left.\begin{aligned} y_1&=\alpha_1+\beta_1,\\ y_2&=-\frac{1}{2}(\alpha_1+\beta_1)+i\frac{\sqrt{3}}{2}(\alpha_1-\beta_1),\\ y_3&=-\frac{1}{2}(\alpha_1+\beta_1)-i\frac{\sqrt{3}}{2}(\alpha_1-\beta_1),\end{aligned}\right\} \qquad (3.10)$$

где в качестве α_1 берется любое из значений корня (3.8), а в качестве β_1 — то значение корня (3.9), для которого $\alpha_1\beta_1=-\frac{p}{3}$.

Пример 2. Решить уравнение $x^3-9x^2+21x-5=0$.

Делаем замену $x=y+3$ и получаем уравнение $y^3-6y+4=0$. Для него одно из значений корня $\alpha=\sqrt[3]{-2+\sqrt{4-8}}=\sqrt[3]{-2+2i}$ есть $\alpha_1=1+i$. Соответствующее ему значение второго корня $\beta_1=-\frac{-6}{3(1+i)}=1-i$. По формулам (3.10) находим:

$y_1=2$, $y_2=-1-\sqrt{3}$, $y_3=-1+\sqrt{3}$. Корни данного нам уравнения: $x_1=5$, $x_2=2-\sqrt{3}$, $x_3=2+\sqrt{3}$.

Если в уравнении (3.6) $p=0$, то корни этого уравнения — все значения $\sqrt[3]{-q}$; если $q=0$, то корни уравнения (3.6): $y_1=0$, $y_{2,3}=\pm\sqrt{-p}$. Если $p\neq 0$ и $q\neq 0$, но $\left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3=0$, то корни уравнения (3.6) находятся так:

$$y_1=\frac{3q}{p},\quad y_2=y_3=-\frac{3q}{2p} \tag{3.11}$$

([17], стр. 215).

Пример 3. Решить уравнение $x^3+3x-2i=0$.

Здесь член с x^2 отсутствует, так что это уже уравнение вида (3.6); $p=3$, $q=-2i$, $\left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3=0$. Формулы (3.11) дают: $x_1=-2i$, $x_2=x_3=i$.

Если все коэффициенты уравнения (3.6) — действительные числа, то его корни можно найти без явного использования формулы Кардано (3.7) следующим образом ([21], стр. 452—454):

1-й случай. Пусть $D=\left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3<0$ *). Тогда все три корня уравнения (3.6) являются действительными и находятся по формулам

$$y_k=2\sqrt{-\frac{p}{3}}\cos\frac{\varphi+2k\pi}{3}\ ^{**)}, \tag{3.12}$$

где $k=1, 2, 3$, а φ — наименьший положительный угол, удовлетворяющий условию $\cos\varphi=-\frac{q}{2}\sqrt{-\left(\frac{3}{p}\right)^3}$.

Пример 4. $y^3-3y+1=0$. Здесь $D=\frac{1}{4}-1<0$.

Применяем формулы (3.12): $\cos\varphi=-\frac{1}{2}\sqrt{1}=-\frac{1}{2}$, $\varphi=120°$; $y_k=2\cos\frac{120°+2k\pi}{3}$, $k=1, 2, 3$, т. е. $y_1=2\cos 160°\approx-1{,}880$, $y_2=2\cos 280°\approx 0{,}348$, $y_3=2\cos 40°\approx 1{,}532$.

---

*) Число $-108\,D$ является дискриминантом многочлена, стоящего в левой части уравнения (3.6) (§ 1, п. 12).

**) Здесь и в следующих формулах нужно брать положительное значение квадратного корня.

2-й случай. Пусть $D>0$. Тогда уравнение (3.6) имеет один действительный и два сопряженных комплексных корня. Если $p<0$, то эти корни находятся так:

$$y_1=-\frac{2\sqrt{-\frac{p}{3}}}{\sin 2\varphi},\quad y_{2,3}=\frac{\sqrt{-\frac{p}{3}}}{\sin 2\varphi}\pm i\sqrt{-p}\operatorname{ctg}2\varphi, \tag{3.13}$$

где $\operatorname{tg}\varphi=\sqrt[3]{\operatorname{tg}\frac{\omega}{2}}$, $\sin\omega=\frac{2}{q}\sqrt{-\left(\frac{p}{3}\right)^3}$.

Если $p>0$, то имеем формулы

$$y_1=-2\sqrt{\frac{p}{3}}\operatorname{ctg}2\varphi,\quad y_{2,3}=\sqrt{\frac{p}{3}}\operatorname{ctg}2\varphi\pm\frac{i\sqrt{p}}{\sin 2\varphi}, \tag{3.14}$$

где $\operatorname{tg}\varphi=\sqrt[3]{\operatorname{tg}\frac{\omega}{2}}$, $\operatorname{tg}\omega=\frac{2}{q}\sqrt{\left(\frac{p}{3}\right)^3}$.

Пример 5. $y^3-9y+12=0$.
Здесь $D=36-27>0$, $p=-9<0$.

Применяем формулы (3.13); $\sin\omega=\frac{2}{12}\sqrt{3^3}=\frac{\sqrt{3}}{2}$, $\omega=60°$.
$\operatorname{tg}\varphi=\sqrt[3]{\operatorname{tg}30°}\approx 0{,}833$, $\varphi\approx 39°50'$; $\sin 2\varphi\approx 0{,}984$; $\operatorname{ctg}2\varphi\approx 0{,}182$; $y_1\approx-\frac{2\sqrt{3}}{0{,}984}\approx-3{,}520$, $y_{2,3}\approx\frac{\sqrt{3}}{0{,}984}\pm i\cdot 3\cdot 0{,}182\approx 1{,}760\pm 0{,}546i$.

Пример 6. $x^3-6x^2+17x-25=0$.

Замена $y=x-2$ дает уравнение $y^3+5y-7=0$, для которого $D=\frac{49}{4}+\frac{125}{27}>0, p=5>0$. Применяем формулы (3.14), причем вычисления производим с помощью логарифмов: $\operatorname{tg}\omega=\frac{2}{-7}\sqrt{\frac{125}{27}}$, $\operatorname{tg}(180°-\omega)=-\operatorname{tg}\omega=\frac{2}{7}\sqrt{\frac{125}{27}}$, $\lg\operatorname{tg}(180°-\omega)=\bar{1}{,}7886$, $180°-\omega=31°35'$, $\omega=148°25'$; $\operatorname{tg}\varphi=\sqrt[3]{\operatorname{tg}74°13'}$, $\lg\operatorname{tg}\varphi=0{,}1829$, $\varphi=56°44'$; $y_1=-2\sqrt{\frac{5}{3}}\operatorname{ctg}113°28'=2\sqrt{\frac{5}{3}}\operatorname{ctg}66°32'$, $\lg y_1=0{,}0496$, $y_1=1{,}121$; $y_{2,3}=\sqrt{\frac{5}{3}}\operatorname{ctg}113°28'\pm\frac{i\sqrt{5}}{\sin 113°28'}=-0{,}561\pm\frac{i\sqrt{5}}{\sin 66°32'}=-0{,}561\pm 2{,}438i$; $x_1=3{,}121$, $x_{2,3}=1{,}439\pm 2{,}438i$.

В случае $D>0$ корни уравнения (3.6) можно также находить по формулам ([4], стр. 139):

при $p<0$ $y_1=-2r\operatorname{ch}\frac{\varphi}{3}$, $y_{2,3}=r\operatorname{ch}\frac{\varphi}{3}\pm i\sqrt{3}r\operatorname{sh}\frac{\varphi}{3}$, где $\operatorname{ch}\varphi = \frac{q}{2r^3}$, $r = \pm\sqrt{-\frac{p}{3}}$, причем знак $+$ или $-$ выбирается в зависимости от того, $q > 0$ или $q < 0$;

при $p > 0$ $\quad y_1 = -2r\operatorname{sh}\frac{\varphi}{3}$, $\quad y_{2,3} = r\operatorname{sh}\frac{\varphi}{3} \pm i\sqrt{3}r\operatorname{ch}\frac{\varphi}{3}$,

где $\operatorname{sh}\varphi = \frac{q}{2r^3}$, $r = \pm\sqrt{\frac{p}{3}}$, причем знак квадратного корня опять выбирается совпадающим со знаком q.

Пример 7. $y^3 - 6y + 9 = 0$.

Здесь $\frac{q^2}{4} + \frac{p^3}{27} = \frac{81}{4} - 8 > 0$, $p = -6 < 0$, $q = 9 > 0$, и мы имеем: $r = \sqrt{2} \approx 1{,}41$, $\operatorname{ch}\varphi = \frac{9}{2\cdot 2\sqrt{2}} \approx 1{,}59$, $\quad \varphi \approx 1{,}05$; $\quad y_1 \approx \approx -2\sqrt{2}\operatorname{ch}0{,}35 \approx -2{,}99$, $\quad y_{2,3} \approx \sqrt{2}\operatorname{ch}0{,}35 \pm i\sqrt{3}\sqrt{2}\operatorname{sh}0{,}35 \approx \approx 1{,}49 \pm 0{,}86i$.

Пример 8. $y^3 + 9y - 26 = 0$.

Здесь $\frac{q^2}{4} + \frac{p^3}{27} > 0$, $p = 9 > 0$, $q = -26 < 0$, и мы имеем: $r = -\sqrt{3} \approx -1{,}73$, $\operatorname{sh}\varphi = \frac{-13}{-3\sqrt{3}} \approx 2{,}51$, $\quad \varphi \approx 1{,}65$; $\quad y_1 \approx \approx 2\sqrt{3}\operatorname{sh}0{,}55 \approx 2{,}01$, $\quad y_{2,3} \approx -\sqrt{3}\operatorname{sh}0{,}55 \mp i\sqrt{3}\sqrt{3}\operatorname{ch}0{,}55 \approx \approx -1 \mp 3{,}48i$.

3-й случай. Пусть $D = 0$. Тогда все корни уравнения (3.6) действительны, причем два из них равны между собой. Корни вычисляются по формулам (3.11) (при $p \neq 0$).

3. Решение уравнений четвертой степени. Чтобы найти все корни уравнения 4-й степени, делят обе его части на коэффициент при x^4 (что не меняет корней уравнения) и затем в уравнении

$$x^4 + b_1x^3 + b_2x^2 + b_3x + b_4 = 0 \tag{3.15}$$

делают замену неизвестного $x = y - \frac{b_1}{4}$ (используя, как и в п. 2, схему Горнера). В результате получается уравнение

$$y^4 + py^2 + qy + r = 0. \tag{3.16}$$

Если y_1, y_2, y_3, y_4 — корни уравнения (3.16), то $x_i = y_i - \frac{b_1}{4}$ ($i = 1, 2, 3, 4$) — корни уравнения (3.15).

Корни уравнения (3.16) можно найти следующим образом.

1) Способ Феррари. Составляется кубическое уравнение

$$z^3+2pz^2+(p^2-4r)z-q^2=0 \tag{3.17}$$

(*кубическая резольвента* уравнения (3.16)) и берется один из его корней z_0. Затем составляются два квадратных уравнения:

$$u^2-\sqrt{z_0}u+\left(\frac{p+z_0}{2}+\frac{q}{2\sqrt{z_0}}\right)=0, \tag{3.18}$$

$$u^2+\sqrt{z_0}u+\left(\frac{p+z_0}{2}-\frac{q}{2\sqrt{z_0}}\right)=0. \tag{3.19}$$

Корни уравнений (3.18) и (3.19) служат корнями уравнения (3.16).

Пример 1. Решить уравнение $x^4-4x^3+7x^2-2x-5=0$. Замена $y=x-1$ дает $y^4+y^2+4y-3=0$. Кубическая резольвента этого уравнения $z^3+2z^2+13z-16=0$; один из ее корней $z_0=1$. Уравнения (3.18) и (3.19): $u^2-u+3=0$, $u^2+u-1=0$; их корни: $y_{1,2}=\frac{1}{2}\pm\sqrt{\frac{1}{4}-3}=\frac{1\pm i\sqrt{11}}{2}$, $y_{3,4}=-\frac{1}{2}\pm$ $\pm\sqrt{\frac{1}{4}+1}=\frac{-1\pm\sqrt{5}}{2}$. Корни данного нам уравнения: $x_{1,2}=$ $=\frac{3\pm i\sqrt{11}}{2}$, $x_{3,4}=\frac{1\pm\sqrt{5}}{2}$.

2) Способ Эйлера ([23], стр. 204—205, [17], стр. 222—225 или [31], стр. 156—158). Составляем кубическую резольвенту (3.17) и находим ее корни z_0, z_1, z_2. Корни уравнения (3.16) находим по формулам

$$y_1=\frac{1}{2}(\sqrt{z_0}+\sqrt{z_1}+\sqrt{z_2}),$$

$$y_2=\frac{1}{2}(\sqrt{z_0}-\sqrt{z_1}-\sqrt{z_2}),$$

$$y_3=\frac{1}{2}(-\sqrt{z_0}+\sqrt{z_1}-\sqrt{z_2}),$$

$$y_4=\frac{1}{2}(-\sqrt{z_0}-\sqrt{z_1}+\sqrt{z_2}),$$

где знаки $\sqrt{z_0}$, $\sqrt{z_1}$, $\sqrt{z_2}$ выбираются так, чтобы выполнялось равенство $\sqrt{z_0}\cdot\sqrt{z_1}\cdot\sqrt{z_2}=-q$.

Пример 2. $y^4-6y^2+8y-3=0$.

Здесь кубическая резольвента $z^3-12z^2+48z-64=0$ имеет корни $z_0=z_1=z_2=4$; $\sqrt{z_i}=\pm 2$, $i=0, 1, 2$. Корни данного нам уравнения 4-й степени:

$$y_1=\frac{1}{2}(2+2-2)=1,\quad y_2=\frac{1}{2}(2-2+2)=1,$$

$$y_3=\frac{1}{2}(-2+2+2)=1,\quad y_4=\frac{1}{2}(-2-2-2)=-3.$$

3) Если в уравнении (3.16) $q=0$ (*биквадратное* уравнение), то его корни находятся по формуле

$$y=\pm\sqrt{-\frac{p}{2}\pm\sqrt{\frac{p^2}{4}-r}}.$$

Пример 3. $y^4+10y^2+9=0$; $y=\pm\sqrt{-5\pm\sqrt{25-9}}$; корни уравнения: $y_{1,2}=\pm i$, $y_{3,4}=\pm 3i$.

4. Об уравнениях степени выше четвертой. Говорят, что уравнение

$$a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n=0$$

разрешимо в радикалах, если его корни можно выразить через его коэффициенты с помощью действий сложения, вычитания, умножения, деления и извлечения корней. Любое уравнение степени $n\leqslant 4$ разрешимо в радикалах, причем для каждого фиксированного $n\leqslant 4$ можно указать общую формулу, с помощью которой корни любого уравнения степени n выражаются через коэффициенты этого уравнения. Для $n>4$ такой общей формулы уже указать нельзя (теорема Руффини—Абеля) ([23], стр. 392 или [17], стр. 279). Более того, для каждого $n>4$ можно указать такое уравнение степени n даже с целыми коэффициентами, что его корни вообще нельзя выразить через его коэффициенты с помощью указанных выше действий (для $n=5$ таким будет, например, уравнение $x^5-4x-2=0$). Но наряду с этим для каждого n существуют и такие уравнения степени n, которые разрешимы в радикалах, например *двучленное уравнение* $x^n-a=0$, корнями которого являются все значения корня n-й степени из числа a ([11], стр. 126). Полностью вопрос об условиях, при которых данное уравнение разрешимо в радикалах, был исследован Галуа в тридцатых годах прошлого века ([5], ч. I, стр. 221).

5. Рациональные корни многочленов с рациональными коэффициентами. Многочлен $f(x)$ с рациональными коэффициентами имеет те же корни, что и многочлен $g(x)$ с целыми коэффициентами, получающийся из $f(x)$ умножением на общее кратное всех знаменателей коэффициентов $f(x)$.

Рациональными корнями многочлена

$$g(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n \tag{3.20}$$

с целыми коэффициентами могут быть только те несократимые дроби вида $\frac{p}{q}$, у которых p — делитель числа a_n, q — делитель числа a_0, и даже только те из этих дробей, для которых при любом целом m $g(m)$ делится на $p-mq$. Если $a_0=1$, то все рациональные корни многочлена (3.20) (если они у него вообще есть) необходимо являются целыми числами.

Чтобы найти все рациональные корни многочлена (3.20), можно сначала умножить этот многочлен на a_0^{n-1} и сделать замену $y=a_0x$, в результате чего получится многочлен

$$h(y)=y^n+b_1y^{n-1}+\ldots+b_{n-1}y+b_n \tag{3.21}$$

опять с целыми коэффициентами. Чтобы найти все рациональные корни $g(x)$, достаточно найти все целые корни $h(y)$ и разделить их на a_0. Для нахождения целых корней $h(y)$, нужно, например, с помощью схемы Горнера проверить, какие из делителей числа b_n являются корнями $h(y)$ (§ 1, п.10), причем те делители p числа b_n, для которых $h(1)$ не делится на $p-1$ или $h(-1)$ не делится на $p+1$, рассматривать не нужно; можно еще взять какое-нибудь целое число $m\neq\pm1$ и исключить из рассмотрения также те делители p числа b_n, для которых $h(m)$ не делится на $p-m$ или $h(-m)$ не делится на $p+m$.

Пример 1. $g(x)=2x^5+7x^4+3x^3-11x^2-16x-12$.

Здесь $h(y)=y^5+7y^4+6y^3-44y^2-128y-192$; делители числа 192: ±1, ±2, ±3, ±4, ±6, ±8, ±12, ±16, ±24, ±32, ±48, ±64, ±96, ±192; $h(1)=-350$, $h(-1)=-108$; делители p числа 192, для которых $h(1)$ делится на $p-1$ и $h(-1)$ делится на $p+1$: 2, 3, -4, 8. С помощью схемы Горнера проверяем, что

$h(2) \neq 0$, $h(3) = 0$, $h(-4) = 0$, $h(8) \neq 0$, причем -4 — двукратный корень $h(y)$. Рациональные корни $g(x)$: $x_1 = \frac{3}{2}$, $x_{2,3} = -2$.

Вместо того, чтобы сразу переходить от многочлена (3.20) к многочлену (3.21), можно сначала путем испытания делителей числа a_n найти все целые корни $x_1, \ldots, x_k$ многочлена $g(x)$ (каждый корень выписан столько раз, какова его кратность), а затем указанным выше образом найти все рациональные корни частного от деления $g(x)$ на $(x-x_1) \ldots \ldots (x-x_k)$.

Пример 2. $g(x) = 3x^6 - 5x^4 - 10x^3 - 8x^2 + x - 2$.

Число $x = 2$ — корень $g(x)$ кратности 1, других целых корней $g(x)$ не имеет. Частное от деления $g(x)$ на $x - 2$: $g_1(x) = 3x^5 + 6x^4 + 7x^3 + 4x^2 + 1$. Умножение $g_1(x)$ на 3^4 и замена $y = 3x$ дают: $h_1(y) = y^5 + 6y^4 + 21y^3 + 36y^2 + 81$. $h_1(1) = 145$, $h_1(-1) = 101$. Ни для какого делителя p числа 81 $h_1(1)$ не делится на $p - 1$ (и $h_1(-1)$ не делится на $p + 1$), т. е. $h_1(y)$ целых корней не имеет. Следовательно, $g_1(x)$ не имеет рациональных корней. Единственный рациональный корень многочлена $g(x)$ есть $x_1 = 2$.

Если для многочлена (3.20) числа a_0, a_n и хотя бы одно из чисел $g(1)$, $g(-1)$ нечетны или если числа a_0, a_n и оба числа $g(1)$, $g(-1)$ не делятся на 3, то многочлен $g(x)$ не имеет рациональных (в том числе и целых) корней ([11], изд. 4, стр. 352).

Пример 3. $g(x) = 2x^8 - x^7 + 8x^4 + x^3 - 5$.

Здесь числа $a_0 = 2$, $a_8 = -5$, $g(1) = 5$, $g(-1) = 7$ не делятся на 3, $g(x)$ не имеет рациональных корней.

6. Разложение многочлена на множители первой и второй степени. Если многочлен $f(x) = a_0x^n + a_1x^{n-1} + \ldots \ldots + a_{n-1}x + a_n$ имеет корни $\alpha_1, \alpha_2, \ldots, \alpha_s$ соответственно кратностей $k_1, \ldots, k_s$ $(k_1 + \ldots + k_s = n)$, то его можно следующим образом разложить на множители:

$$f(x) = a_0(x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\ldots(x-\alpha_s)^{k_s} \tag{3.22}$$

(*разложение $f(x)$ на линейные множители*). Разложение всякого многочлена на линейные множители единственно с точностью до порядка множителей.

Если $f(x)$ — многочлен с действительными коэффициентами и α — комплексный корень этого многочлена, то сопряженное с α число $\bar{\alpha}$ ([11], стр. 121) также является корнем $f(x)$, причем кратности корней α и $\bar{\alpha}$ совпадают. Если в разложении (3.22), написанном для многочлена с действительными коэффициентами, объединить попарно скобки, соответствующие комплексно сопряженным корням этого многочлена, то получится разложение $f(x)$ на множители, также имеющие действительные коэффициенты:

$$f(x)=a_0(x-\alpha_1)^{k_1}\dots(x-\alpha_l)^{k_l}\times$$
$$\times(x^2+p_1x+q_1)^{k'_1}\dots(x^2+p_tx+q_t)^{k'_t}.$$

Квадратные трехчлены, входящие в это разложение, не имеют действительных корней и уже на линейные множители с действительными коэффициентами разлагаться не могут. Такое разложение многочлена $f(x)$ с действительными коэффициентами единственно с точностью до порядка множителей.

Пример.

$$f(x)=(x-1)(x-i)(x+i)[x-(1+i)]^2[x-(1-i)]^2=$$
$$=(x-1)(x^2+1)(x^2-2x+2)^2.$$

7. Неприводимые многочлены. Многочлен $f(x)$ степени $n\geqslant 1$ с коэффициентами из поля K (стр. 175) называется *неприводимым* над этим полем, если он не может быть разложен в произведение многочленов степеней $<n$ с коэффициентами из этого же поля; в противном случае многочлен $f(x)$ называется *приводимым* над полем K. Все многочлены 1-й степени неприводимы над любым полем. Если K есть поле всех комплексных чисел, то многочленов выше 1-й степени, неприводимых над этим полем, не существует (см. предыдущий пункт). Над полем действительных чисел, кроме многочленов 1-й степени, неприводимыми являются также многочлены 2-й степени, не имеющие действительных корней, а все многочлены степени выше 2-й уже приводимы. Если K есть поле рациональных чисел, то для всякого n существует многочлен степени n, неприводимый над этим полем.

Пример 1. $f(x)=x^2+1$.

Этот многочлен неприводим над полем действительных чисел, но приводим над полем комплексных чисел: $f(x)=(x+i)(x-i)$.

Пример 2. $f(x)=x^2-2$.

Этот многочлен приводим как над полем комплексных чисел, так и над полем действительных чисел: $f(x)=(x-\sqrt{2})(x+\sqrt{2})$. Над полем рациональных чисел данный многочлен неприводим.

Многочлен выше 1-й степени, неприводимый над некоторым полем, не может иметь в этом поле корней (обратное неверно).

Если $f(x)$ — произвольный, а $p(x)$ — неприводимый многочлен, оба с коэффициентами из поля K, то либо $f(x)$ делится без остатка на $p(x)$, либо же эти многочлены взаимно просты (§ 1, п. 6).

Если произведение нескольких многочленов делится на неприводимый многочлен $p(x)$, то хотя бы один из множителей делится на $p(x)$.

Всякий многочлен $f(x)$ с коэффициентами из поля K, имеющий степень $n \geqslant 1$, разлагается в произведение многочленов, неприводимых над этим полем, причем если многочлен $f(x)$ двумя способами разложен в произведение неприводимых множителей: $f(x) = p_1(x)p_2(x)\ldots p_s(x) = q_1(x)q_2(x)\ldots$ $\ldots q_t(x)$, то $s = t$ и, при соответствующей нумерации, имеют место равенства $q_i(x) = c_i p_i(x)$, $i = 1, 2, \ldots, s$, где c_i — отличные от нуля элементы поля K.

Если в разложении многочлена $f(x)$ на неприводимые множители из каждого из этих множителей вынести за скобку старший коэффициент, то получится разложение $f(x) =$ $= a_0 p_1(x) p_2(x) \ldots p_s(x)$, где все $p_i(x)$, $i = 1, 2, \ldots, s$, являются неприводимыми многочленами со старшими коэффициентами, равными единице. Для всякого многочлена такое разложение уже вполне однозначно с точностью до нумерации множителей. Если неприводимый многочлен $p(x)$ встречается в указанном разложении многочлена $f(x)$ k раз, то $p(x)$ называется *k-кратным множителем* для $f(x)$.

Если даны разложения многочленов $f(x)$ и $g(x)$ на неприводимые множители, то наибольший общий делитель $d(x)$ этих многочленов равен произведению множителей, входящих одновременно в оба разложения, причем каждый множитель берется в степени, равной меньшей из его кратностей в обоих данных многочленах.

Пример 3. $f(x) = 2(x+1)(x^2+1)(x^2-x+1)^3$, $g(x) =$ $= 5(x^2+1)^4(x^2-x+1)^2(x^2+3x+4)^3$ — разложения многочленов на неприводимые множители над полем действительных чисел. $d(x) = (x^2+1)(x^2-x+1)^2$.

8. Разложение многочленов на неприводимые множители над полем рациональных чисел. Многочлен $f(x)$ с рациональными коэффициентами неприводим над полем рациональных чисел (п. 7) тогда и только тогда, когда над этим полем неприводим многочлен с целыми коэффициентами, полученный умножением $f(x)$ на общее наименьшее кратное знаменателей всех его коэффициентов. Многочлен с целыми коэффициентами неприводим над полем рациональных чисел тогда и только тогда, когда он не разлагается на множители тоже с целыми коэффициентами, каждый из которых имеет более низкую степень, чем сам многочлен.

Пусть

$$g(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n \tag{3.23}$$

— многочлен с целыми коэффициентами.

Критерий Эйзенштейна. Если можно найти такое простое число p, что в многочлене (3.23) a_0 не делится на p, все остальные коэффициенты делятся на p, но a_n, делясь на p, не делится на p^2, то этот многочлен неприводим над полем рациональных чисел.

Пример 1. $g(x)=3x^7-4x^6+2x^5-6x^3-8x-2$.

Здесь подходит число $p=2$, многочлен $g(x)$ неприводим над полем рациональных чисел.

Пример 2. $g(x)=x^4+x^3+x^2+x+1$.

К этому многочлену критерий Эйзенштейна непосредственно применить нельзя, но можно сделать замену $x=y+1$, в результате которой получится многочлен $h(y)=y^4+5y^3+10y^2+10y+5$, неприводимый в силу критерия Эйзенштейна ($p=5$). Следовательно, и многочлен $g(x)$ неприводим над полем рациональных чисел.

Критерий Полиа ([31], стр. 111, 114). Пусть для многочлена (3.23) A — наибольшее из чисел $|a_1|, \ldots, |a_n|$ и k — целое число со свойством $k \geqslant \frac{A}{|a_0|}+\frac{3}{2}$. Тогда если $g(k-1)\neq 0$ и $g(k)$ — простое число, то многочлен $g(x)$ неприводим над полем рациональных чисел.

Пример 3. $g(x)=x^3-x^2+x+1$.

Здесь $A=1$; $k=3$ — наименьшее целое число со свойством $k \geqslant \frac{A}{|a_0|}+\frac{3}{2}$. $g(3)=22$ — составное число, но $g(4)=53$ — число простое. Многочлен $g(x)$ неприводим над полем рациональных чисел.

Критерий Кона ([31], стр. 118). Если коэффициенты многочлена (3.23) удовлетворяют условию $0\leqslant a_i\leqslant 9$, $i=0, 1, \ldots, n$, и $N=a_0 10^n+a_1 10^{n-1}+\ldots+a_{n-1}10+a_n$ — простое число, то этот многочлен неприводим над полем рациональных чисел.

Пример 4. $g(x)=2x^4+x^3+5x^2+6x+3$; $N=21\,563$ — простое число; многочлен $g(x)$ неприводим.

Приведем еще один критерий неприводимости ([24], задача 681). Если многочлен (3.23) принимает значения ± 1 более чем при $2l$ целых значениях x, причем $n=2l$ или $2l+1$, то он неприводим над полем рациональных чисел.

Пример 5. $g(x)=x^3-3x^2+3$.

Здесь $g(-1)=-1$, $g(1)=1$, $g(2)=-1$. Многочлен $g(x)$ неприводим.

Другие достаточные условия неприводимости многочлена с рациональными коэффициентами см. в [5], гл. IV, § 24 и [28], гл. I, § 4.

Для многочленов 2-й, 3-й и 4-й степени с рациональными коэффициентами можно указать не только достаточные, но и необходимые условия неприводимости над полем рациональных чисел ([17], стр. 174—182).

Многочлен 2-й или 3-й степени с рациональными коэффициентами приводим над полем рациональных чисел тогда и только тогда, когда он имеет хотя бы один рациональный корень.

Многочлен 4-й степени

$$f(x)=x^4+b_1x^3+b_2x^2+b_3x+b_4$$

с рациональными коэффициентами приводим над полем рациональных чисел тогда и только тогда, когда-либо он имеет рациональный корень, либо связанный с ним многочлен 3-й степени

$$k(t)=8t^3-4b_2t^2+(2b_1b_3-8b_4)t-(b_1^2b_4-4b_2b_4+b_3^2)\ ^{*)}$$

имеет такой рациональный корень t_0, что

$$\lambda=\sqrt{2t_0+\frac{b_1^2}{4}-b_2} \quad \text{и} \quad \mu=\sqrt{t_0^2-b_4}$$

— рациональные числа. В последнем случае

$$f(x)=\left[x^2+\left(\frac{b_1}{2}+\lambda\right)x+(t_0+\mu)\right]\times$$
$$\times\left[x^2+\left(\frac{b_1}{2}-\lambda\right)x+(t_0-\mu)\right],$$

---

*) Если составить многочлен $k(t)$ для левой части уравнения (3.16) и сделать в нем замену $t=\frac{z+p}{2}$, то получится левая часть кубической резольвенты (3.17) уравнения (3.16).

где значения корней λ и μ нужно брать одного знака, если $b_1t_0 - b_3 > 0$, и противоположных знаков, если $b_1t_0 - b_3 < 0$.

Пример 6. $f(x) = x^4 - 4x^3 + 5x^2 - 2$.

Здесь $x = 1$ — корень $f(x)$, $f(x) = (x-1)(x^3 - 3x^2 + 2x + 2)$. Многочлен $x^3 - 3x^2 + 2x + 2$ рациональных корней не имеет (п. 5) и над полем рациональных чисел неприводим.

Пример 7. $f(x) = x^4 - 2x^3 - 7x^2 - 10x - 2$.

Рациональных корней нет. Многочлен $k(t) = 8t^3 + 28t^2 + 56t - 36$ имеет рациональный корень $t_0 = \frac{1}{2}$ и $\lambda = \sqrt{1+1+7} = \pm 3$, $\mu = \sqrt{\frac{1}{4} + 2} = \pm\frac{3}{2}$ — числа рациональные. $b_1t_0 - b_3 = -1 + 10 > 0$, $f(x) = (x^2 + 2x + 2)(x^2 - 4x - 1)$.

Пример 8. $f(x) = 2x^4 - 3x^3 - x^2 - 3x + 1$.

По способу, изложенному в п. 5, определяем, что $f(x)$ рациональных корней не имеет. Делим все коэффициенты $f(x)$ на старший коэффициент $a_0 = 2$ и составляем многочлен

$$k(t) = 8t^3 - 4\left(-\frac{1}{2}\right)t^2 + \left[2\left(-\frac{3}{2}\right)\left(-\frac{3}{2}\right) - 8\cdot\frac{1}{2}\right]t - \\ - \left[\frac{9}{4}\,\frac{1}{2} - 4\left(-\frac{1}{2}\right)\frac{1}{2} + \frac{9}{4}\right] = 8t^3 + 2t^2 + \frac{1}{2}t - \frac{35}{8}.$$

Умножая $k(t)$ на 8 и делая замену $v = 4t$, получаем многочлен $v^3 + v^2 + v - 35$, не имеющий рациональных корней. Следовательно, $k(t)$ также не имеет рациональных корней. Многочлен $f(x)$ неприводим над полем рациональных чисел.

Если к многочлену (3.23) ни один из указанных выше критериев неприводимости применить нельзя, то приводим этот многочлен над полем рациональных чисел или нет, можно установить следующим образом ([27], стр. 66—68 или [13], стр. 173—176).

Переменному x придают m целых значений $x_1, x_2, \ldots, x_m$, где $m = \frac{n}{2} + 1$ при четном n и $m = \frac{n+1}{2}$ при нечетном n. Затем составляют всевозможные наборы целых чисел вида $c_1, c_2, \ldots, c_m$, где c_i $(i = 1, 2, \ldots, m)$ есть делитель числа $g(x_i)$ (всего получается $S = 2^m s_1 \ldots s_m$ наборов, где s_i — число всех положительных делителей $g(x_i)$). Для каждого такого набора строят многочлен $h_j(x)$ $(j = 1, 2, \ldots, S)$ степени $\leqslant m - 1$ со свойствами $h_j(x_i) = c_i$, где $i = 1, 2, \ldots, m$, а числа c_i взяты из j-го набора (§ 2). После этого для тех из многочленов $h_j(x)$, все коэффициенты которых получились целыми и степень которых >0, непосредственным делением проверяют, делится ли $g(x)$ на $h_j(x)$. Если $g(x)$ не делится ни на один из этих многочленов, он неприводим над полем рациональных чисел. В противном случае получается разложение многочлена $g(x)$ на множители более низкой степени с рациональными коэффициентами, из которого легко получить разложение на множители с целыми коэффициентами, с которыми дальше можно действовать так же, как раньше с многочленом $g(x)$.

При применении к многочлену $g(x)$ указанного выше процесса возможны следующие упрощения:

если два набора чисел $c_1, c_2, \ldots, c_m$ и $c'_1, c'_2, \ldots, c'_m$ отличаются друг от друга лишь знаком, то многочлен $h_j(x)$ строится только для одного из этих наборов;

можно взять еще несколько целых значений переменного $x_{m+1}, x_{m+2}, \ldots, x_{m+l}$; если хотя бы для одного из них $g(x_{m+k})$ не делится на $h_j(x_{m+k})$, то делится ли $g(x)$ на $h_j(x)$, проверять не нужно.

Пример 9. $g(x)=x^5-2x^4-3x^3+6x^2+x-1$.

Здесь $g(0)=-1$, $g(1)=2$, $g(-1)=4$, $g(2)=1$. Выбираем значения $g(0)$, $g(1)$, $g(2)$, как имеющие меньше делителей, и составляем наборы делителей выбранных чисел:

$$\begin{array}{l|r|r|r|r||r|r|r|r||}
c_1=1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & 1\\
c_2=1 & 1 & -1 & 1 & 2 & 2 & -2 & 2 & \\
c_3=1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 &
\end{array}{}^{*)}$$

По формуле Ньютона (§ 2) строим для этих наборов многочлены $h_1(x)\equiv 1$, $h_2(x)=-x^2+x+1$, $h_3(x)=2x^2-4x+1$, $h_4(x)=-x^2+3x-1$, $h_5(x)=-x^2+2x+1$, $h_6(x)=-2x^2+3x+1$, $h_7(x)=3x^2-6x+1$, $h_8(x)=-2x^2+5x-1$, где $h_j(x)$ ($j=1, 2, \ldots, 8$) — многочлен, принимающий при $x=0, 1, 2$ значения, стоящие в j-м столбце выписанной выше таблицы. Числа $h_3(-1)=7$, $h_4(-1)=-5$, $h_7(-1)=10$ и $h_8(-1)=-8$ не являются делителями числа $g(-1)=4$, поэтому многочлены $h_3(x)$, $h_4(x)$, $h_7(x)$ и $h_8(x)$ не рассматриваем. Возьмем еще значение $g(-2)=-19$. Оно не делится на $h_2(-2)=-5$, $h_5(-2)=-7$ и $h_6(-2)=-13$. Многочлен $g(x)$ неприводим над полем рациональных чисел.

---

*) Наборы, отличающиеся от выписанных здесь лишь знаком, не берем.

Пример 10. $g(x)=x^4+x^3-6x^2-3x+9$.

$$g(-2)=-1,\ g(1)=2,\ g(2)=3.$$

Составляем наборы делителей этих чисел:

$$\begin{array}{l|r|r|r|r||r|r|r|r||r|r|r|r||}
c_1=1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & 1 \\
c_2=1 & 1 & -1 & 1 & 2 & 2 & -2 & 2 & 1 & 1 & -1 & 1 \\
c_3=1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & 3 & 3 & 3 & -3
\end{array}$$

$$\begin{array}{||r|r|r|r||}
1 & -1 & 1 & 1 \\
2 & 2 & -2 & 2 \\
3 & 3 & 3 & -3
\end{array}\ {}^{*)}$$

Строим соответствующие им многочлены (§ 2, формула Ньютона):

$h_1(x)\equiv 1$; $h_2(x)$, $h_3(x)$, $h_4(x)$, $h_5(x)$, $h_6(x)$ — многочлены с дробными коэффициентами; $h_7(x)=x^2-3$ — делитель $g(x)$, $g(x)=(x^2-3)(x^2+x-3)$, причем многочлены x^2-3 и x^2+x-3 над полем рациональных чисел неприводимы (так как не имеют рациональных корней).

Другой способ разложения многочлена с рациональными коэффициентами на множители более низкой степени с рациональными коэффициентами см. в [27], стр. 68.

9. Дробно-рациональные функции. *Дробно-рациональной функцией* (или просто *рациональной дробью*) называется выражение вида $\frac{f(x)}{g(x)}$, где $f(x)$ и $g(x)$ — многочлены над некоторым полем K (§ 1), причем $g(x)\neq 0$. Над дробно-рациональными функциями можно производить алгебраические операции по таким же законам, как над рациональными числами. Рациональная дробь $\frac{f_1(x)}{g_1(x)}$ называется *несократимой*, если многочлены $f_1(x)$ и $g_1(x)$ взаимно просты (§ 1, п. 6). Всякая рациональная дробь равна несократимой дроби, получающейся из нее путем деления ее числителя $f(x)$ и знаменателя $g(x)$ на наибольший общий делитель $f(x)$ и $g(x)$.

Пример 1. $\dfrac{x^2-x}{x^3+2x^2-x}=\dfrac{x-1}{x^2+2x-1}$.

Рациональная дробь называется *правильной*, если в ней степень числителя меньше степени знаменателя. Всякая рациональная дробь, отличная от многочлена (т. е. такая, что в ее

---

*) См. сноску на стр. 209.

несократимой записи знаменатель имеет ненулевую степень), может быть единственным образом представлена в виде суммы некоторого многочлена и некоторой правильной дроби. Чтобы получить такое представление дроби $\frac{f(x)}{g(x)}$, нужно разделить с остатком многочлен $f(x)$ на $g(x)$ (§ 1, п. 2); если при этом получится частное $q(x)$ и остаток $r(x)$, то будем иметь

$$\frac{f(x)}{g(x)}=q(x)+\frac{r(x)}{g(x)},$$

где $\frac{r(x)}{g(x)}$ — правильная рациональная дробь.

Пример 2. $\frac{f(x)}{g(x)}=\frac{3x^4+2x^2-4x+6}{x^2-2x+3}$.
Здесь
$3x^4+2x^2-4x+6=(x^2-2x+3)(3x^2+6x+5)-12x-9$,

$$\frac{f(x)}{g(x)}=3x^2+6x+5+\frac{-12x-9}{x^2-2x+3}.$$

Правильная рациональная дробь $\frac{f(x)}{g(x)}$ называется *простейшей*, если $g(x)$ является степенью некоторого многочлена $p(x)$, неприводимого (п. 7) над заданным полем K ($g(x)=p^k(x)$, где $k\geqslant 1$), а степень многочлена $f(x)$ меньше степени $p(x)$. Если K — поле всех комплексных чисел, то простейшие дроби — это дроби вида $\frac{a}{(x-b)^l}$, где a и b — комплексные числа и $l\geqslant 1$; если K — поле всех действительных чисел, то простейшими являются рациональные дроби вида $\frac{a}{(x-b)^l}$ (a, b — действительные числа, $l\geqslant 1$) и $\frac{a_1x+a_2}{(x^2+b_1x+b_2)^m}$, где $x^2+b_1x+b_2$ — многочлен с действительными коэффициентами, не имеющий действительных корней, a_1, a_2 — действительные числа, $m\geqslant 1$.

Основная теорема. Всякая правильная рациональная дробь разлагается в сумму простейших дробей, причем это разложение единственно.

Чтобы получить разложение дроби $\frac{f(x)}{g(x)}$ на простейшие над данным числовым полем K, нужно разложить многочлен $g(x)$ на неприводимые над K множители и, если получится $g(x)=p_1^{s_1}(x)\,\ldots\,p_t^{s_t}(x)$ ($p_1(x), \ldots, p_t(x)$ — неприводимые многочлены), написать

$$\frac{f(x)}{g(x)}=\frac{u_1^{(1)}(x)}{p_1^{s_1}(x)}+\frac{u_2^{(1)}(x)}{p_1^{s_1-1}(x)}+\ldots+\frac{u_{s_1}^{(1)}(x)}{p_1(x)}+\ldots$$
$$\ldots+\frac{u_1^{(t)}(x)}{p_t^{s_t}(x)}+\frac{u_2^{(t)}(x)}{p_t^{s_t-1}(x)}+\ldots+\frac{u_{s_t}^{(t)}(x)}{p_t(x)}, \tag{3.24}$$

где $u_j^{(l)}(x)$ ($l=1, \ldots, t$; $j=1, \ldots, s_l$ при фиксированном l) — многочлен степени меньшей, чем степень $p_l(x)$, взятый с неопределенными коэффициентами. Затем нужно все дроби правой части равенства (3.24) привести к общему знаменателю $g(x)$ и приравнять сумму получившихся при этом числителей многочлену $f(x)$. Сравнивая коэффициенты при одинаковых степенях x в левой и правой частях получившегося равенства (или придавая переменному x в левой и правой части равенства числовые значения), получим систему линейных уравнений, решая которую (гл. I, § 2, п. 1—3) найдем коэффициенты многочленов $u_j^{(l)}(x)$.

Пример 3. Рациональную дробь

$$\frac{f(x)}{g(x)}=\frac{2x^4-2x^3+6x^2+2x+7}{x^5+2x^4+2x^3+4x^2+x+2}$$

разложить на простейшие над полем действительных чисел.

Здесь $g(x)=(x+2)(x^2+1)^2$, причем многочлены $x+2$ и x^2+1 неприводимы. Пишем

$$\frac{f(x)}{g(x)}=\frac{a}{x+2}+\frac{a_1x+a_2}{(x^2+1)^2}+\frac{c_1x+c_2}{x^2+1}.$$

Приводим все дроби к общему знаменателю $g(x)$ и получаем равенство

$$f(x)=a(x^2+1)^2+(a_1x+a_2)(x+2)+(c_1x+c_2)(x+2)(x^2+1),$$

или

$$2x^4-2x^3+6x^2+2x+7=(a+c_1)x^4+(2c_1+c_2)x^3+$$
$$+(2a+a_1+c_1+2c_2)x^2+(2a_1+a_2+2c_1+c_2)x+(a+2a_2+2c_2).$$

Приравнивая коэффициенты при одинаковых степенях x, получаем систему уравнений

$$\begin{cases} a + c_1 = 2, \\ 2c_1 + c_2 = -2, \\ 2a + a_1 + c_1 + 2c_2 = 6, \\ 2a_1 + a_2 + 2c_1 + c_2 = 2, \\ a + 2a_2 + 2c_2 = 7. \end{cases}$$

Решая эту систему, находим: $a=3$, $a_1=1$, $a_2=2$, $c_1=-1$, $c_2=0$. Таким образом, $\frac{f(x)}{g(x)}=\frac{3}{x+2}+\frac{x+2}{(x^2+1)^2}-\frac{x}{x^2+1}$.

Пример 4. Ту же рациональную дробь разложить на простейшие над полем комплексных чисел.

Здесь $g(x)=(x+2)(x+i)^2(x-i)^2$. В соответствии с этим пишем

$$\frac{f(x)}{g(x)}=\frac{a}{x+2}+\frac{b}{(x+i)^2}+\frac{c}{x+i}+\frac{d}{(x-i)^2}+\frac{e}{x-i}.$$

Приводим все дроби к общему знаменателю $g(x)$ и получаем равенство

$$2x^4-2x^3+6x^2+2x+7=a(x+i)^2(x-i)^2+b(x+2)(x-i)^2+$$
$$+c(x+2)(x+i)(x-i)^2+d(x+2)(x+i)^2+$$
$$+e(x+2)(x+i)^2(x-i).$$

Беря значение $x=-2$, получаем: $75=a\cdot 25$, т. е. $a=3$. Берем $x=-i$. Это дает: $3-4i=b(2-i)(-4)$, откуда $b=-\frac{1}{2}+\frac{1}{4}i$. При $x=i$ получаем: $3+4i=d(2+i)(-4)$, т. е. $d=-\frac{1}{2}-\frac{1}{4}i$. Полагаем теперь последовательно $x=0$ и $x=-1$ и, используя уже найденные значения коэффициентов a, b, d, получаем

$$\left\{\begin{aligned} 7&=3+\left(1-\frac{1}{2}i\right)-2ci+\left(1+\frac{1}{2}i\right)+2ei,\\ 15&=12+\left(-\frac{1}{2}-i\right)+c(-2-2i)+\left(-\frac{1}{2}+i\right)+e(-2+2i),\end{aligned}\right.$$

или

$$\left\{\begin{aligned} &-2ci+2ei=2,\\ &c(-2-2i)+e(-2+2i)=4,\end{aligned}\right.$$

откуда $c=\frac{-1+i}{2}$, $e=\frac{-1-i}{2}$. Таким образом,

$$\frac{f(x)}{g(x)}=\frac{3}{x+2}-\frac{2-i}{4(x+i)^2}-\frac{1-i}{2(x+i)}-\frac{2+i}{4(x-i)^2}-\frac{1+i}{2(x-i)}.$$

Пример 5. Разложить на простейшие дробь $\frac{f(x)}{g(x)}=\frac{x^2+2}{(x-1)^3}$.

Пишем: $\frac{f(x)}{g(x)}=\frac{a}{(x-1)^3}+\frac{b}{(x-1)^2}+\frac{c}{x-1}$. Приводим все дроби к общему знаменателю и получаем: $f(x)=a+b(x-1)+c(x-1)^2$. Полагая последовательно $x=1$, $x=0$, $x=2$, находим: $a=3$, $b=2$, $c=1$. Таким образом,

$$\frac{x^2+2}{(x-1)^3}=\frac{3}{(x-1)^3}+\frac{2}{(x-1)^2}+\frac{1}{x-1}.$$

Еще о разложении дробей на простейшие см. § 1, п. 4 и § 2 (формула Лагранжа).

§ 4. Границы корней многочлена. Теоремы о числе корней

1. Границы действительных корней многочлена с действительными коэффициентами. Если

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n \quad (a_0\neq 0) \quad (3.25)$$

— многочлен с действительными коэффициентами, то все его действительные корни (если они вообще существуют) заключены в интервале $(-N, N)$, где $N=1+\frac{A}{|a_0|}$, A — наибольшее из чисел $|a_1|, \ldots, |a_n|$ (ср. § 4, п. 3).

Пример 1. $f(x)=5x^6-11x^5+38x^4+x^3-100x^2-3$.

Здесь $A=100$, $N=1+\frac{100}{5}=21$.

Можно указать и более точные границы корней многочлена (3.25). Число K называется *верхней границей действительных корней* многочлена (3.25), если этот многочлен не имеет действительных корней, больших или равных K; аналогично определяются *верхняя* и *нижняя границы положительных* или *отрицательных корней* многочлена. Если K_1 — верхняя граница положительных корней $f(x)$, K_2 — верхняя граница положительных корней $f(-x)$, $K_3>0$ — верхняя граница положительных корней многочлена $x^nf\left(\frac{1}{x}\right)$*) и $K_4>0$ — верхняя граница положительных корней многочлена $x^nf\left(-\frac{1}{x}\right)$, то все отличные от нуля действительные корни многочлена $f(x)$, если они вообще существуют, лежат внутри интервалов $\left(-K_2, -\frac{1}{K_4}\right)$ и $\left(\frac{1}{K_3}, K_1\right)$.

Для определения верхней границы положительных корней многочлена можно применить один из следующих методов.

1) Если в многочлене (3.25) $a_0>0$, $a_1, a_2, \ldots, a_{l-1}\geqslant 0$ и $a_l<0$, то верхняя граница положительных корней этого многочлена находится по формуле

$$K=1+\sqrt[l]{\frac{B}{a_0}},$$

*) Этот многочлен можно получить, написав все коэффициенты многочлена $f(x)$ в обратном порядке:

$$x^nf\left(\frac{1}{x}\right)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0.$$

где B — наибольшая из абсолютных величин отрицательных коэффициентов $f(x)$ (способ Маклорена).

Пример 2. $f(x)=9x^4-9x^2-36x+1$.

Здесь $l=2$, $B=36$, $K=K_1=1+\sqrt{\frac{36}{9}}=3$. Рассмотрим многочлен $f(-x)=9x^4-9x^2+36x+1$. Для него верхней границей положительных корней служит число $K_2=1+\sqrt{\frac{9}{9}}=2$. Берем $x^4 f\left(\frac{1}{x}\right)=x^4-36x^3-9x^2+9$. Здесь $l=1$, $K_3=1+36=37$. Наконец, для $x^4 f\left(-\frac{1}{x}\right)=x^4+36x^3-9x^2+9$ имеем $K_4=1+$ $+\sqrt{9}=4$. Следовательно, если многочлен $f(x)$ имеет действительные корни, то они обязательно должны лежать в интервалах $\left(-2,-\frac{1}{4}\right)$ и $\left(\frac{1}{37}, 3\right)$ (по способу, указанному в начале этого пункта, получился бы интервал $(-5, 5)$).

2) Если при $x=c$ многочлен (3.25) и его производные $f'(x)$, $f''(x), \ldots, f^{(n)}(x)$ принимают положительные значения, то c является верхней границей положительных корней $f(x)$ (способ Ньютона). Применяя способ Ньютона, значения $f(c)$, $f'(c), \ldots, f^{(n)}(c)$ проще всего находить по схеме Горнера (§ 1, пп. 3—4).

Пример 3. $f(x)=3x^4-18x^3+14x^2+36x+25$.

По способу Маклорена находим: $K=1+\frac{18}{3}=7$. Примеияем способ Ньютона. Испытываем число $c=3$:

	3	−18	14	36	25
3	3	−9	−13	−3	16
	3	0	−13	−42	

Получили: $f(3)=16>0$, но $f'(3)=-42<0$; $c=3$ не подходит. Испытываем $c=4$:

	3	−18	14	36	25
4	3	−6	−10	−4	9
	3	6	14	52	
	. . .	. . .	. . .		

Так как получилась строка, состоящая целиком из положительных чисел, можно дальше не считать (везде дальше будут получаться положительные числа). $f(4)$, $f'(4)$, $f''(4)$, $f'''(4)$, $f^{\mathrm{IV}}(4)>0$, $c=4$ — верхняя граница положительных корней $f(x)$.

Для нахождения числа c, для которого $f(c)>0$, $f'(c)>0, \ldots, f^{(n)}(c)>0$, можно поступить и так: взять число c_1, для которого $f^{(n-1)}(c_1)>0$, затем число $c_2 \geqslant c_1$, для которого $f^{n-2}(c_2)>0$, затем число $c_3 \geqslant c_2$, для которого $f^{(n-3)}(c_3)>0$, и т. д. до $c_n=c$.

Пример 4. $f(x)=x^3-6x^2+7x+5$.
Здесь $f'(x)=3x^2-12x+7$, $f''(x)=6x-12$ и можно взять $c_1=3$, $c_2=4$, $c_3=4$. $c=c_3=4$ — верхняя граница положительных корней $f(x)$.

3) Предполагая $a_0>0$, представляем многочлен (3.25) (не переставляя его членов) в виде $f(x)=f_1(x)+\ldots+f_s(x)$, где у каждого многочлена $f_i(x)$ $(i=1, \ldots, s)$ старший коэффициент положителен и в ряду коэффициентов имеется только одна перемена знака*) ($f_s(x)$ может вообще не иметь перемен знака в ряду коэффициентов). Если при $c>0$ $f_1(c)>0, \ldots, f_s(c)>0$, то c — верхняя граница положительных корней $f(x)$ ([31], стр. 170).

Пример 5. $f(x)=x^7+2x^6-3x^5-3x^4+x^3-2x^2+x+2$. Полагаем

$$f_1(x)=x^7+2x^6-3x^5-3x^4, \quad f_2(x)=x^3-2x^2, \quad f_3(x)=x+2.$$

Тогда $f_1(3)>0$, $f_2(3)>0$, $f_3(3)>0$ и $c=3$ — верхняя граница положительных корней $f(x)$.

Если все коэффициенты многочлена $f(x)$ положительны, он вообще не имеет положительных корней (и верхней границей его положительных корней можно считать 0 или любое положительное число).

Если в многочлене (3.25) $a_0<0$, то нужно рассмотреть многочлен $-f(x)$, имеющий те же корни, что и $f(x)$.

Пример 6. $f(x)=-2x^4-3x^3+5x^2+6$.
Здесь $-f(x)=2x^4+3x^3-5x^2-6$. По способу Маклорена для многочлена $-f(x)$ получаем: $K=1+\sqrt{\frac{6}{2}}=1+\sqrt{3}\approx 2{,}74$ (приближенное значение $\sqrt{3}$ берем с избытком). Переходя к способу Ньютона, испытываем $c=2$:

	2	3	−5	0	−6
2	2	7	9	18	30

*) См. сноску на стр. 217.

Получилась строка из положительных чисел. $c=2$ — верхняя граница положительных корней многочлена $-f(x)$, т. е. и многочлена $f(x)$.

2. Теоремы о числе действительных корней. Если каждый корень многочлена считать столько раз, какова его кратность (§ 1, п. 10), то число всех действительных корней многочлена $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ с действительными коэффициентами всегда имеет ту же четность, что и степень n многочлена. В частности, если n нечетно, то $f(x)$ необходимо имеет хотя бы один действительный корень. Число положительных корней четно или нечетно в зависимости от того, одинаковые или противоположные знаки имеют a_0 и последний отличный от нуля коэффициент $f(x)$.

Если при $x=b$ и $x=c$ $f(x)$ принимает значения противоположных знаков, то на отрезке $[b, c]$ лежит хотя бы один действительный корень многочлена $f(x)$, а общее число корней $f(x)$ на отрезке $[b, c]$ нечетно (если опять каждый корень считать столько раз, какова его кратность). Если значения $f(b)$ и $f(c)$ имеют одинаковые знаки, то на отрезке $[b, c]$ либо вообще нет корней $f(x)$, либо их там четное число.

Число всех отрицательных корней многочлена $f(x)$ равно числу всех положительных корней многочлена $f(-x)$.

Теорема Декарта. Число положительных корней многочлена $f(x)$ с действительными коэффициентами (подсчитываемых каждый столько раз, какова его кратность) либо равно числу перемен знака*) в ряду коэффициентов $f(x)$, либо на четное число меньше.

Пример 1. $f(x)=x^3-2x+1$.

Знаки коэффициентов: $+\ -\ +$, т. е. имеем две перемены знака. Число положительных корней многочлена $f(x)$ — либо 2, либо 0. $f(-x)=-x^3+2x+1$. Здесь в ряду коэффициентов одна перемена знака. Число отрицательных корней $f(x)$ равно 1.

Теорема Декарта позволяет определить точное число положительных корней многочлена, если в ряду его коэффициентов нет перемен знака или имеется одна перемена знака. Если перемен знака больше одной, то иногда точное число

---

*) Если в ряду чисел, расположенных в определенном порядке, имеется k переходов от одного знака к другому (числа, равные нулю, пропускаем), то говорят, что в этом ряду содержится k перемен знака.

положительных корней многочлена удается определить, применяя теорему Декарта и рассматривая, кроме того, знаки значений многочлена в отдельных точках или же поведение отдельных членов многочлена при изменении x.

Пример 2. $f(x)=x^4-2x^3-2x+1$.

В ряду коэффициентов $f(x)$ две перемены знака, т. е. $f(x)$ имеет 2 или 0 положительных корней. $f(0)=1>0$, $f(1)=-2<0$, т. е. отрезок $[0, 1]$ содержит по крайней мере один корень многочлена $f(x)$. Следовательно, число всех положительных корней $f(x)$ равно двум. $f(-x)=x^4+2x^3+2x+1$. Отрицательных корней многочлен $f(x)$ не имеет.

Пример 3. $f(x)=x^7+x^2-x+1$.

В ряду коэффициентов — две перемены знака. Но при $0<x\leqslant 1$ единственный отрицательный член многочлена, $-x$, «покрывается» свободным членом 1, а при $x>1$ член x^2 больше чем x. Следовательно, при любом положительном значении x многочлен $f(x)$ принимает положительное значение и число положительных корней многочлена $f(x)$ равно 0.

Иногда точное число всех действительных корней многочлена с действительными коэффициентами можно определить, применяя теорему Декарта и более детально рассматривая график этого многочлена на участках, заключенных между границами его действительных корней.

Пример 4. $f(x)=x^5-2x^4-8x^3-x^2-9x+1$.

По теореме Декарта, имеем 2 или 0 положительных корней. $f(0)>0$, $f(1)<0$, поэтому положительных корней два. Берем $f(-x)=-x^5-2x^4+8x^3-x^2+9x+1$. Здесь три перемены знака, т. е. $f(x)$ имеет один или три отрицательных корня. По методам п. 1 определяем, что все отрицательные корни многочлена $f(x)$ лежат на отрезке $[-3, -2]$. Тем же способом определяем, что -2 нижняя граница действительных корней производной $f'(x)$, т. е. на отрезке $[-3, -2]$ $f(x)$ — монотонная функция ([34], стр. 53). Следовательно, $f(x)$ имеет один отрицательный корень. Общее число действительных корней многочлена $f(x)$ равно 3.

В некоторых случаях точное число всех действительных корней многочлена можно определить, применяя теорему Декарта и вычисляя дискриминант $D(f)$ этого многочлена (§ 1, п. 12).

Пример 5. $f(x)=x^4+x^2-4x+3$.

С помощью теоремы Декарта определяем, что число положительных корней $f(x)$ равно 2 или 0, а число отрицательных корней равно нулю, т. е. либо все корни многочлена $f(x)$ комплексные, либо у него два действительных и два комплексных корня. $D(f)=5744>0$, т. е. число пар комплексно сопряженных корней $f(x)$ четно (§ 1, п. 12). Следовательно, все корни многочлена $f(x)$ комплексные.

Усиление теоремы Декарта. Если все корни многочлена $f(x)$ действительны, то число положительных корней $f(x)$ равно числу перемен знака в ряду его коэффициентов, а число отрицательных корней равно числу перемен знака в ряду коэффициентов многочлена $f(-x)$.

Пример 6. $f(x)=\begin{vmatrix} -2-x & 1 & 1 \\ 1 & 1-x & 3 \\ 1 & 3 & 1-x \end{vmatrix}=-x^3+14x+20.$

Здесь $f(x)$ — характеристический многочлен симметрической матрицы (гл. II, § 1, п. 17), поэтому все его корни действительны. $f(-x)=x^3-14x+20$. Число положительных корней многочлена $f(x)$ равно 1, число отрицательных корней равно 2.

Если два рядом стоящих коэффициента многочлена $f(x)$ равны нулю, но $a_n \neq 0$, то заведомо не все корни $f(x)$ являются действительными ([23], стр. 151); в частности, усиление теоремы Декарта применять нельзя.

Число действительных корней многочлена $f(x)$ степени n с действительными коэффициентами, бо́льших фиксированного действительного числа c, или равно числу перемен знака в ряду $f(c)$, $f'(c)$, ..., $f^{(n)}(c)$, или на четное число меньше. Если все корни многочлена $f(x)$ действительны, то число корней, больших c, в точности равно числу перемен знака в ряду $f(c)$, $f'(c)$, ..., $f^{(n)}(c)$.

Пример 7. $f(x)=x^4-3x^3+x^2-2x+1$.

Здесь $f(1)=-2$, $f'(1)=-5$, $f''(1)=-4$, $f'''(1)=6$, $f^{IV}(1)=24$. Получили одну перемену знака, многочлен $f(x)$ имеет один действительный корень, больший 1.

Теорема Бюдана — Фурье. Если действительные числа a и b $(a<b)$ не являются корнями многочлена $f(x)$ степени n с действительными коэффициентами, то число действительных корней этого многочлена, лежащих между a и b и подсчитываемых каждый столько раз, какова его кратность, либо равно, либо на четное число меньше разности между числом S_1 перемен знака в ряду чисел $f(a)$, $f'(a)$, ..., $f^{(n)}(a)$ и числом S_2 перемен знака в ряду, полученном из ряда чисел $f(b)$, $f'(b)$, ..., $f^{(n)}(b)$ заменой встречающихся в нем нулей такими отличными от нуля числами, что если $f^{(k-1)}(b)\neq 0$, $f^{(k)}(b)=\ldots=f^{(k+l-1)}(b)=0$, $f^{(k+l)}(b)\neq 0$, то число c_{k+i}, заменяющее $f^{(k+i)}(b)$ $(i=0, 1, \ldots, l-1)$, имеет

тот же знак, что и $f^{(k+l)}(b)$, если разность $l-i$ четная, и противоположный знак, если эта разность нечетная.

Пример 8. $f(x)=3x^8-2x^5+x^4+4x^2-x-1$. Определить число корней $f(x)$ на отрезке $[-2, 0]$.

Имеем: $f(-2)>0$, $f'(-2)<0$, $f''(-2)>0$, $f'''(-2)<0$, $f^{(4)}(-2)>0$, $f^{(5)}(-2)<0$, $f^{(6)}(-2)>0$, $f^{(7)}(-2)<0$, $f^{(8)}(-2)>0$, откуда $S_1=8$; $f(0)<0$, $f'(0)<0$, $f''(0)>0$, $f'''(0)=0$, $f^4(0)>0$, $f^{(5)}(0)<0$, $f^{(6)}(0)=0$, $f_1^{(7)}(0)=0$, $f^{(8)}(0)>0$; составляем ряд знаков $--+-+-+-+$; получаем $S_2=7$. $S_1-S_2=1$, на отрезке $[-2, 0]$ лежит один корень многочлена $f(x)$.

Метод Штурма определения числа действительных корней. Последовательность отличных от нуля многочленов

$$f(x)=f_0(x),\ f_1(x), \ldots, f_m(x) \tag{3.26}$$

с действительными коэффициентами называется *рядом Штурма* для многочлена $f(x)$, если:

1) когда действительное переменное x, возрастая, проходит через корень многочлена $f(x)$, произведение $f(x)\cdot f_1(x)$ меняет знак с $-$ на $+$;

2) никакие два соседних многочлена $f_k(x)$, $f_{k+1}(x)$ $(k=0, 1, \ldots, m-1)$ не имеют общих действительных корней;

3) если для некоторого действительного числа a и некоторого номера $k\,(1\leqslant k\leqslant m-1)$ $f_k(a)=0$, то числа $f_{k-1}(a)$ и $f_{k+1}(a)$ имеют разные знаки;

4) $f_m(x)$ не имеет действительных корней.

Последовательность многочленов (3.26) называется *рядом Штурма* для многочлена $f(x)$ *на отрезке* $[a, b]$, если свойства 1), 2), 3), 4) выполнены на этом отрезке ([23], стр. 141).

Теорема Штурма. Пусть $f(x)$ — многочлен с действительными коэффициентами, a и b $(a<b)$ — действительные числа, не являющиеся его корнями, (3.26) — ряд Штурма для $f(x)$ на отрезке $[a, b]$. Тогда число различных действительных корней многочлена $f(x)$, лежащих на отрезке $[a, b]$, равно разности между числом перемен знака*) $W(a)$ в ряду значений $f(a), f_1(a), \ldots, f_m(a)$ и числом перемен знака $W(b)$ в ряду значений $f(b), f_1(b), \ldots, f_m(b)$.

Построение ряда Штурма. Берется многочлен $f(x)=f_0(x)$, его производная $f'(x)=f_1(x)$, затем остаток $r_1(x)$, получающийся при делении $f(x)$ на $f'(x)$, с обратным

*) См. сноску на стр. 217.

знаком ($-r_1(x)=f_2(x)$), затем остаток $r_2(x)$, получающийся при делении $f'(x)$ на $f_2(x)$, с обратным знаком ($-r_2(x)=f_3(x)$), остаток от деления $f_2(x)$ на $f_3(x)$ с обратным знаком ($-r_3(x)=f_4(x)$) и т. д. до последнего не равного нулю остатка $r_{m-1}(x)$ ($-r_{m-1}(x)=f_m(x)$). Если $f_m(x)$ не имеет действительных корней, то построенный ряд многочленов есть ряд Штурма для многочлена $f(x)$; если $f_m(x)$ имеет действительные корни, но все они лежат вне отрезка $[a, b]$, то это — ряд Штурма для многочлена $f(x)$ на отрезке $[a, b]$. Если многочлен $f_m(x)$ имеет действительные корни на отрезке $[a, b]$, то нужно разделить многочлены $f_0(x)$, $f_1(x), \ldots, f_m(x)$ на $f_m(x)$ и взять получившиеся частные. Это будет ряд Штурма для многочлена $q_0(x)=\frac{f_0(x)}{f_m(x)}$ ([23], стр. 145), имеющего точно те же корни, что и $f(x)$, только кратности 1 (§ 1, п. 10); разность $W(a)-W(b)$ (a, b — действительные числа, не являющиеся корнями $f(x)$), вычисленная для этого ряда, дает число различных действительных корней многочлена $f(x)$ на отрезке $[a, b]$; при этом фактически деление на $f_m(x)$ можно не производить, так как при любом действительном значении $x=c$, не являющемся корнем многочлена $f(x)$, число перемен знака в получающемся ряду значений частных совпадает с числом перемен знака в ряду значений $f_0(c), f_1(c), \ldots, f_m(c)$.

Пример 9. Найти число различных действительных корней многочлена $f(x)=x^4-3x^3-x^2+8x-4$, расположенных на отрезке $[-1, 1]$.

Берем многочлены $f_0(x)=f(x)$, $f_1(x)=f'(x)=4x^3-9x^2-2x+8$, $f_2(x)=-r_1(x)=\frac{35}{16}x^2-\frac{45}{8}x+\frac{5}{2}$, $f_3(x)=-r_2(x)=\frac{160}{49}x-\frac{320}{49}$ ($f_2(x)$ на $f_3(x)$ делится уже без остатка). Так как на отрезке $[-1, 1]$ многочлен $f_3(x)$ корней не имеет, то получился ряд Штурма для многочлена $f(x)$ на отрезке $[-1, 1]$. Составляем таблицу, в первом столбце которой выписываем значения, которые придаются переменному x, в следующих столбцах выписываем знаки получающихся при этом значений многочленов ряда Штурма, в последнем столбце — число перемен знака

x	$f_0(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	W
-1	$-$	$-$	$+$	$-$	2
1	$+$	$+$	$-$	$-$	1

$W(-1)-W(1)=2-1=1$. На отрезке $[-1, 1]$ $f(x)$ имеет один корень.

Пример 10. Для того же многочлена $f(x)$ определить число различных действительных корней на отрезке $[-5, 5]$. Построенный выше ряд многочленов не есть ряд Штурма на этом отрезке, так как $f_3(x)$ имеет на этом отрезке корень $x=2$. Тем не менее выписываем таблицу:

x	$f_0(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	W
-5	$+$	$-$	$+$	$-$	3
5	$+$	$+$	$+$	$+$	0

$W(-5)-W(5)=3-0=3$. На отрезке $[-5, 5]$ многочлен $f(x)$ имеет три различных корня.

При построении ряда Штурма возможны упрощения:

1) если $f'(x)$ или один из остатков $r_i(x)$ $(1 \leqslant i < m-1)$ не имеет действительных корней (не имеет корней на отрезке $[a, b]$), то на нем построение ряда Штурма (построение ряда Штурма на отрезке $[a, b]$) можно оборвать ([23], стр. 143);

2) производя деление с остатком, можно умножать делимое и делитель на произвольные положительные числа (там же);

3) если $f'(x)$ (или один из многочленов $f_i(x)$, $1 < i \leqslant m-1$) делится без остатка на многочлен $g(x)$, принимающий при всех действительных значениях x неотрицательные значения и обращающийся в 0 только при таких значениях $x=\alpha$, при которых $f(\alpha)\neq 0$, то $f'(x)$ (или соответственно $f_i(x)$) можно на $g(x)$ разделить и дальше при построении ряда Штурма брать вместо $f'(x)$ (вместо $f_i(x)$) получившееся частное.

Пример 11. $f(x)=x^6+4x^3-3x^2+12x-24$. Здесь $f'(x)=6(x^5+2x^2-x+2)$ имеет корнем число i и поэтому делится на x^2+1 (§ 3, п. 6). После сокращения на этот многочлен и на число 6 получаем $f_1(x)=x^3-x+2$. Делим $f(x)$ на $f_1(x)$ и берем остаток от этого деления (сокращенный на 2) с обратным знаком: $f_2(x)=x^2-6x+14$. Этот многочлен не имеет действительных корней (§ 3, п. 1), и на нем можно остановиться. Получили ряд Штурма для многочлена $f(x)$.

Если нужно определить общее число всех различных действительных корней многочлена $f(x)$ с действительными коэффициентами, то можно применить к $f(x)$ теорему Штурма на отрезке $[a, b]$, где a — нижняя, b — верхняя граница действительных корней $f(x)$ (п. 1). Более простой способ — применить теорему Штурма на отрезке $[a, b]$, где a — такое большое по абсолютной величине отрицательное число, а b — такое большое положительное число, что при $x=a$ и при $x=b$ знаки всех членов ряда Штурма, построенного для многочлена $f(x)$, совпадают со знаками старших членов этих многочленов; такие числа a и b всегда существуют; символическое обозначение для них: $-\infty$ и $+\infty$; фактически находить числа a и b не нужно.

Если нужно узнать отдельно, сколько различных положительных и сколько различных отрицательных корней имеет многочлен $f(x)$, то теорему Штурма применяют на отрезках $[-\infty, 0]$ и $[0, +\infty]$. Если нужно узнать точнее, в каких интервалах лежат корни, то рассматривают знаки значений многочленов ряда Штурма в различных точках этих отрезков.

Пример 12. $f(x)=4x^6+4x^5-15x^4-16x^3+12x^2+16x+4$. Берем $f_1(x)=\frac{1}{4}f'(x)=6x^5+5x^4-15x^3-12x^2+6x+4$,

$$f_2(x)=50x^4+57x^3-84x^2-114x-32$$

(это взятый с обратным знаком остаток от деления многочлена $9f(x)$ на $f_1(x)$), $f_3(x)=2x^3+x^2-4x-2$ (остаток с обратным знаком от деления $f_1(x)$ на $f_2(x)$, умноженный на $\frac{625}{882}$). $f_2(x)$ делится на $f_3(x)$ уже без остатка. Составляем таблицу:

x	$f(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	W
$-\infty$	+	−	+	−	3
$+\infty$	+	+	+	+	0
0	+	+	−	−	1
−1	+	0	−	+	2
1	+	−	−	−	1
2	+	+	+	+	0

$W(-\infty)-W(+\infty)=3$, т. е. $f(x)$ имеет три различных действительных корня. $W(-\infty)-W(0)=2$, $W(0)-W(+\infty)=1$, т. е. два из них отрицательны и один положителен.

$$W(-\infty)-W(-1)=1, \quad W(-1)-W(0)=1,$$

т. е. один из отрицательных корней меньше -1, другой заключен между -1 и 0. $W(1)-W(2)=1$, т. е. положительный корень лежит на отрезке $[1, 2]$.

Если нужно определить, сколько корней каждой заданной кратности (§ 1, п. 10) имеет многочлен $f(x)$ на отрезке $[a, b]$, то нужно составить многочлены $d_1(x) = (f(x), f'(x))$, $d_2(x) = (d_1(x), d_1'(x))$ и т. д. до $d_s(x) = (d_{s-1}(x), d'_{s-1}(x))$ (§ 1, п. 5) и для каждого из них определить, сколько различных корней имеет этот многочлен на отрезке $[a, b]$. Если $f(x)$ имеет на отрезке $[a, b]$ k_0 различных корней, а $d_i(x)$ — k_i различных корней ($i = 1, 2, \ldots, s$), то на отрезке $[a, b]$ лежат $k_0 - k_1$ простых корней $f(x)$, $k_1 - k_2$ двукратных корней, $k_2 - k_3$ трехкратных корней и т. д.

Пример 13. Пусть $f(x)$ — многочлен, рассмотренный в предыдущем примере. Для него $d_1(x) = f_2(x) = 2x^3 + x^2 - 4x - 2$, $d_2(x) = 1$. По методу Штурма определяем, что на каждом из отрезков $[-\infty, -1]$, $[-1, 0]$, $[1, 2]$ $d_1(x)$ имеет по одному корню. Многочлен $d_2(x)$ вообще корней не имеет. Следовательно, на каждом из рассматриваемых отрезков $f(x)$ имеет по одному двукратному корню.

Метод квадратичных форм. Имеем многочлен

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$$

с действительными коэффициентами. Составляем квадратичную форму (гл. II, § 2, п. 1)

$$\varphi_0(x_0, \ldots, x_{n-1}) = \sum_{i,\,k=0}^{n-1} s_{i+k}x_ix_k,$$

где s_{i+k} — сумма $(i+k)$-х степеней корней этого многочлена (каждый корень берется столько раз, какова его кратность) *). Ранг этой квадратичной формы (гл. II, § 2, п. 1) равен числу всех различных корней многочлена $f(x)$ (действительных и комплексных вместе), а сигнатура (гл. II, § 2, п. 2) равна числу различных действительных корней $f(x)$ ([6], стр. 442).

Пример 14. $f(x) = x^3 - 2x^2 + 5x - 8$.
Здесь

$$s_0 = 3, \quad s_1 = 2, \quad s_2 = -6, \quad s_3 = 2, \quad s_4 = 50;$$

$$\varphi_0(x_0, x_1, x_2) = 3x_0^2 + 4x_0x_1 - 12x_0x_2 - 6x_1^2 + 4x_1x_2 + 50x_2^2.$$

Ранг квадратичной формы φ_0 равен 3, сигнатура равна 1 (способы определения ранга и сигнатуры см. в гл. II, § 2, пп. 1, 2). Следовательно, все три корня многочлена $f(x)$ различны, причем только один из них действительный.

---

*) Коэффициенты s_{i+k} можно найти, пользуясь их выражениями через коэффициенты многочлена $f(x)$ (§ 5, п. 3).

Если нужно определить отдельно число всех положительных или отрицательных корней $f(x)$, то указанный выше результат можно объединить с теоремой: число различных пар комплексно сопряженных корней многочлена $f(x)$ плюс число различных отрицательных корней этого многочлена равно числу отрицательных квадратов в каноническом виде (стр. 141) квадратичной формы

$$\varphi_1(x_0, \ldots, x_{n-1}) = \sum_{i,\, k=0}^{n-1} s_{1+i+k} x_i x_k$$

([15], стр. 161).

Пример 15. $f(x) = 2x^4 - 2x^3 - 2x + 1$. Здесь

$$s_0 = 4,\quad s_1 = 0,\quad s_2 = 2,\quad s_3 = 3,\quad s_4 = 0,\quad s_5 = 5,\quad s_6 = 2,\quad s_7 = \frac{7}{2};$$

$$\varphi_0(x_0, x_1, x_2, x_3) = 4x_0^2 + 4x_0x_2 + 6x_0x_3 + 2x_1^2 + \\ + 6x_1x_2 + 10x_2x_3 + 2x_3^2;$$

$$\varphi_1(x_0, x_1, x_2, x_3) = 4x_0x_1 + 6x_0x_2 + 3x_1^2 + 10x_1x_3 + \\ + 5x_2^2 + 4x_2x_3 + \frac{7}{2}x_3^2 = z_0^2 + z_1^2 + z_2^2 - z_3^2$$

(канонический вид φ_1). Ранг квадратичной формы φ_0 равен 4, сигнатура равна 2, поэтому $f(x)$ имеет два различных действительных и два различных комплексных корня (сопряженных между собой (§ 3, п. 6)). В каноническом виде квадратичной формы φ_1 имеется один отрицательный квадрат. Так как уже известно, что $f(x)$ имеет одну пару комплексно сопряженных корней, то оба действительных корня $f(x)$ положительны.

Замечание. Если

$$f(x) = (a_0 + b_0 i)x^n + (a_1 + b_1 i)x^{n-1} + \ldots \\ \ldots + (a_{n-1} + b_{n-1} i)x + (a_n + b_n i)$$

— многочлен с комплексными коэффициентами (a_k, b_k, $k = 0, 1, \ldots, n$, — действительные числа), то его можно представить в виде $f(x) = g(x) + h(x)i$, где $g(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ и $h(x) = b_0x^n + b_1x^{n-1} + \ldots \ldots + b_{n-1}x + b_n$ — многочлены с действительными коэффициентами. Действительными корнями многочлена $f(x)$ являются все действительные корни наибольшего общего делителя многочленов $g(x)$ и $h(x)$ и только они.

3. Границы комплексных корней многочлена. Если

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \tag{3.27}$$

— многочлен с действительными или комплексными коэффициентами, то все его корни изображаются точками комплексной плоскости ([11], стр. 115), лежащими внутри круга с центром в начале координат радиуса $N=1+\frac{A}{|a_0|}$, где A — наибольшее из чисел $|a_1|, \ldots, |a_n|$.

Пример 1. $f(x)=(1+i)x^3-3x^2+ix+2$.
Здесь $A=3$, $N=1+\frac{3}{\sqrt{2}}\approx 3{,}122$.

Число N можно заменить числом $N_1=\rho+\frac{A_1}{|a_0|}$, где ρ — любое положительное число, A_1 — наибольшее из чисел

$$|a_1|, \quad \frac{|a_2|}{\rho}, \quad \frac{|a_3|}{\rho^2}, \ \ldots, \ \frac{|a_n|}{\rho^{n-1}}$$

([24], задача 686).

Пример 2. $f(x)=x^4-ix^3+(3-4i)x+16$.
Здесь $N=17$. Возьмем $\rho=2$. Тогда получим $A_1=2$ (как наибольшее из чисел $1, \frac{0}{2}, \frac{5}{4}, \frac{16}{8}$), $N_1=2+2=4$. Все корни $f(x)$ лежат внутри круга с центром в начале координат радиуса 4.

Другие формулы для чисел, которыми можно заменить N, см. в [24], гл. V, § 7.

Если все коэффициенты многочлена (3.27) действительны и положительны, то при условии $a_0>a_1>\ldots>a_n$ все его корни лежат внутри круга радиуса 1 с центром в начале координат, а при условии $a_0<a_1<\ldots<a_n$ все корни лежат вне этого круга. Более общий результат: если все коэффициенты многочлена (3.27) строго положительны и если все отношения $\frac{a_k}{a_{k-1}}$ последующего коэффициента к предыдущему заключены между положительными числами m и M $\left(m<\frac{a_k}{a_{k-1}}<M,\ k=1, 2, \ldots, n\right)$, то все корни этого многочлена лежат в кольце, заключенном между окружностями с центром в начале координат соответственно радиусов m и M ([11], изд. 4, стр. 295—296).

Пример 3. $f(x)=11x^5+8x^4+7x^3+5x^2+2x+1$.
Коэффициенты положительны и убывают, все корни $f(x)$ лежат внутри круга радиуса 1 с центром в начале координат.
Пример 4. $f(x)=3x^4+3x^3+5x^2+2x+4$.

Отношения коэффициентов: $1, \frac{5}{3}, \frac{2}{5}, 2$. Все корни $f(x)$ изображаются точками комплексной плоскости, лежащими между окружностями радиусов $\frac{2}{5}$ и 2 с центром в начале координат (или на самих этих окружностях).

Об определении числа корней многочлена с действительными коэффициентами, лежащих внутри или вне произвольно заданной окружности, см. [15], § 10 или [29], гл. I, § 8.

Если коэффициенты a_0, a_1 многочлена (3.27) действительны, причем $a_0 \geqslant 1$, $a_1 \geqslant 0$, то для действительной части $\operatorname{Re}(\alpha)$ всякого корня α многочлена (3.27) справедливо неравенство $\operatorname{Re}(\alpha) < \frac{1+\sqrt{4c+1}}{2}$, где c — любое такое число, что $|a_i| \leqslant c$, $i=2, 3, \ldots, n$ ([31], стр. 114—117). Если коэффициенты a_0, a_1 многочлена (3.27) указанным условиям не удовлетворяют, то нужно сделать замену $y = x + \frac{a_1}{na_0}$, дающую многочлен, в котором отсутствует член с $n-1$-й степенью неизвестного, а затем умножить полученный многочлен на такое число, что в результате получится многочлен $h(y) = b_0 y^n + b_2 y^{n-2} + \ldots$ $\ldots + b_{n-1} y + b_n$ с действительным старшим коэффициентом $b_0 \geqslant 1$. Действительная часть $\operatorname{Re}(\beta)$ всякого корня β многочлена $h(y)$ удовлетворяет неравенствам

$$-\frac{1+\sqrt{4c_1+1}}{2} < \operatorname{Re}(\beta) < \frac{1+\sqrt{4c_1+1}}{2},$$

где c_1 — любое такое число, что $|b_i| \leqslant c_1$, $i=2, 3, \ldots, n$, а для действительных частей $\operatorname{Re}(\alpha)$ корней α многочлена (3.27) справедливы неравенства

$$-\frac{1+\sqrt{4c_1+1}}{2} - \operatorname{Re}\left(\frac{a_1}{na_0}\right) < \operatorname{Re}(\alpha) < \frac{1+\sqrt{4c_1+1}}{2} - \operatorname{Re}\left(\frac{a_1}{na_0}\right) \quad (3.28)$$

($\operatorname{Re}\left(\frac{a_1}{na_0}\right)$ — действительная часть числа $\frac{a_1}{na_0}$). Геометрически это означает, что все корни многочлена (3.27) изображаются точками комплексной плоскости, лежащими в полосе, ограниченной прямыми, параллельными мнимой оси и проходящими через точки, изображающие числа, стоящие в левой и правой частях неравенств (3.28).

Пример 5. $f(x)=(1+i)x^3-3x^2+ix+2.$

Замена $y=x-\frac{1}{1+i}$ дает многочлен

$$g(y)=(1+i)y^3+\frac{-4+i}{1+i}y+\frac{-3+5i}{2i}.$$

Умножаем этот многочлен на $\frac{1}{1+i}$ и получаем

$$h(y)=y^3+\frac{1+4i}{2}y+\frac{4-i}{2}.$$

Берем

$$c_1=2,\ 1\left(>\left|\frac{1+4i}{2}\right|=\left|\frac{4-i}{2}\right|=\frac{\sqrt{17}}{2}\right).$$

Тогда

$$\frac{1+\sqrt{4c_1+1}}{2}=\frac{1+\sqrt{8,4+1}}{2}<\frac{1+3,1}{2}=2,05.$$

Так как

$$\mathrm{Re}\left(\frac{a_1}{3a_0}\right)=\mathrm{Re}\left(\frac{-1}{1+i}\right)=-\frac{1}{2},$$

то для действительных частей корней α многочлена $f(x)$ получаем: $-1,55<\mathrm{Re}(\alpha)<2,55$.

Если нужно установить границы мнимых частей корней многочлена $f(x)$, то можно сделать в этом многочлене замену $x=iy$. Если действительные части корней многочлена $g(y)=f(iy)$ заключены между числами c и d, то между теми же числами заключены коэффициенты при мнимых частях у всех корней многочлена $f(x)$.

Пример 6. $f(x)=ix^5-(1+i)x^3-ix^2+(2-i)x+6.$

Замена $x=iy$ дает

$$g(y)=-y^5+(-1+i)y^3+iy^2+(1+2i)y+6.$$

Для действительных частей $\mathrm{Re}(\beta)$ корней многочлена $g(y)$ справедливы неравенства

$$-\frac{1+\sqrt{4\cdot 6+1}}{2}=-3<\mathrm{Re}(\beta)<3=\frac{1+\sqrt{4\cdot 6+1}}{2}.$$

Следовательно, все корни многочлена $f(x)$ изображаются точками комплексной плоскости, лежащими в полосе, ограниченной прямыми, параллельными действительной оси и проходящими через точки, изображающие числа $-3i$ и $3i$.

4. Число корней в левой и правой полуплоскости.

а) Индексы Коши ([6], гл. XV, §§ 2 и 3). Имеем рациональную функцию

$$\Phi(x)=\frac{k(x)}{h(x)},$$

где $k(x)$ и $h(x)$ — многочлены с действительными коэффициентами. *Индексом Коши* функции $\Phi(x)$ в пределах от a до b (a, b — действительные числа) называется разность между числом разрывов $\Phi(x)$ ([34], стр. 48) с переходом от $-\infty$ к $+\infty$ и числом разрывов с переходом от $+\infty$ к $-\infty$, получающихся при изменении x от a к b. Индекс обозначается через $I_a^b\Phi(x)$. Если $\Phi(x)$ имеет разрыв при $x=a$ или при $x=b$, то такой разрыв не учитывается.

Пример 1. $\Phi(x)=\dfrac{2x+1}{(x+2)^3(x-1)(x-4)(x-5)^2}$.

Если здесь x, возрастая, переходит через точку $x=-2$ или $x=1$, то знак $\Phi(x)$ меняется с $+$ на $-$; если x переходит через точку $x=4$, то знак меняется с $-$ на $+$; при переходе через точку $x=5$ знак $\Phi(x)$ не меняется. Поэтому, например, $I_{-3}^{6}\Phi(x)=1-2=-1$. Если заставить x изменяться от 6 к -3, то получится $I_6^{-3}\Phi(x)=2-1=1$.

В определении индекса Коши вместо действительных чисел a и b можно брать $\pm\infty$. В предыдущем примере

$$I_{-\infty}^{+\infty}\Phi(x)=-1, \quad I_{+\infty}^{-\infty}\Phi(x)=1.$$

Если известны все действительные корни многочлена $h(x)$, причем $h(x)=(x-\alpha_1)^{l_1}\ldots(x-\alpha_m)^{l_m}h_1(x)$ ($\alpha_1, \ldots, \alpha_m$ — все различные между собой действительные корни $h(x)$, $h_1(x)$ — многочлен без действительных корней), то индекс функции $\Phi(x)=\dfrac{k(x)}{h(x)}$ проще всего вычислить, представив ее в виде

$$\Phi(x)=\sum_{i=1}^{m}\left(\frac{A_1^{(i)}}{x-\alpha_i}+\ldots+\frac{A_{l_i}^{(i)}}{(x-\alpha_i)^{l_i}}\right)+\frac{q(x)}{h_1(x)}, \tag{3.29}$$

где коэффициенты $A_j^{(i)}$ и коэффициенты многочлена $q(x)$ — действительные числа, причем $A_{l_i}^{(i)}\neq 0$, $i=1, 2, \ldots, m$ (такое представление функции $\Phi(x)$ всегда возможно); если обозначить через $\overline{A}_{l_i}^{(i)}$ те из коэффициентов $A_{l_i}^{(i)}$, $l=1, 2, \ldots m$, которым соответствуют нечетные показатели l_i, то будем иметь

$$I_{-\infty}^{+\infty} \Phi(x)=\sum \operatorname{sign} \overline{A}_{l_i}^{(i)}, \tag{3.30}$$

где $\operatorname{sign} \overline{A}_{l_i}^{(i)}$ есть 1 или -1 в зависимости от того, $\overline{A}_{l_i}^{(i)}>0$ или $\overline{A}_{l_i}^{(i)}<0$. При практическом нахождении $\overline{A}_{l_i}^{(i)}$ сначала пишут равенство (3.29) с неопределенными коэффициентами, затем приводят все дроби, стоящие в правой части этого равенства, к общему знаменателю и находят $\overline{A}_{l_i}^{(i)}$, исходя из равенства получившегося при этом числителя правой части (3.29) и числителя $\Phi(x)$. Если нужно определить $I_a^b \Phi(x)$, то в формуле (3.30) берут только те $\overline{A}_{l_i}^{(i)}$, которым соответствуют корни α_i, лежащие между a и b $(a<\alpha_i<b)$.

Пример 2. $\Phi(x)=\dfrac{x^8+1}{x^3(x-2)^2(x+3)(x^4+3x^2+5)}$.

Представляем $\Phi(x)$ в виде (3.29):

$$\Phi(x)=\frac{A_1^{(1)}}{x}+\frac{A_2^{(1)}}{x^2}+\frac{A_3^{(1)}}{x^3}+\frac{A_1^{(2)}}{x-2}+\frac{A_2^{(2)}}{(x-2)^2}+$$
$$+\frac{A_1^{(3)}}{x+3}+\frac{q(x)}{x^4+3x^2+5}$$

(многочлен x^4+3x^2+5 действительных корней не имеет). Приводим все дроби к общему знаменателю, приравниваем числители в левой и правой частях равенства и полагаем сначала $x=0$, а затем $x=-3$. Это дает $A_3^{(1)}=\dfrac{1}{4\cdot3\cdot5}>0$, $A_1^{(3)}=\dfrac{-27+1}{(-27)\cdot25\cdot113}>$ >0. Следовательно, $I_{-\infty}^{+\infty}\Phi(x)=2$.

Другие способы определения индекса Коши рациональной дроби $\Phi(x)$ см. в [6], гл. XV, §§ 2 и 11.

Пусть теперь $f(x)$ — многочлен степени n с действительными коэффициентами. Запишем его в виде

$$f(x)=a_0x^n+b_0x^{n-1}+a_1x^{n-2}+b_1x^{n-3}+\ldots \tag{3.31}$$

и положим

$$\begin{aligned} f_1(x)&=a_0x^n-a_1x^{n-2}+a_2x^{n-4}-\ldots, \\ f_2(x)&=b_0x^{n-1}-b_1x^{n-3}+b_2x^{n-5}-\ldots \end{aligned} \tag{3.32}$$

Тогда если t — число корней многочлена $f(x)$, лежащих в правой полуплоскости (т. е. имеющих положительную действительную часть — см. [11], стр. 115), а s — число его чисто мнимых корней, то

$$I_{-\infty}^{+\infty}\frac{f_2(x)}{f_1(x)}=n-2t-s$$

(каждый корень считается столько раз, какова его кратность, см. § 1, п. 10).

Число чисто мнимых корней многочлена $f(x)$ равно числу действительных корней наибольшего общего делителя многочленов (3.32).

Пример 3. $f(x)=x^4+x^3-2x-1$.

Здесь многочлены $f_1(x)=x^4-1$ и $f_2(x)=x^3+2x$ взаимно просты (§ 1, п. 6), так что $s=0$. $I_{-\infty}^{+\infty}\frac{x^3+2x}{x^4-1}=1+1=2=4-2t$, $t=1$.

б) Алгоритм Рауса ([6], гл. XV, §§ 3 и 4). Для многочлена (3.31) составляется схема Рауса

$$\begin{array}{l} a_0,\ a_1,\ a_2,\ \ldots; \\ b_0,\ b_1,\ b_2,\ \ldots, \\ c_0,\ c_1,\ c_2,\ \ldots, \\ d_0,\ d_1,\ d_2,\ \ldots, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot, \end{array}$$

в первых двух строчках которой выписываются коэффициенты многочлена (3.31), а каждая следующая строчка получается из двух предыдущих так: из чисел первой из этих строк вычитаются стоящие под ними числа второй строки, умноженные на такое число, чтобы первая разность была равна нулю; эта первая разность отбрасывается, и строка начинается со второй разности. Такое построение схемы возможно, если в ряду b_0, c_0, d_0, ... не встречается числа, равного нулю (так называемый *регулярный случай*). В этом случае число корней многочлена (3.31), лежащих в правой полуплоскости, равно числу перемен знака*) в первом столбце

*) См. сноску на стр. 217.

построенной схемы (каждый корень считается столько раз, какова его кратность). Чисто мнимых корней многочлен $f(x)$ в регулярном случае не имеет.

Пример 4. $f(x) = x^6 - 5x^5 + 2x^4 - x^3 + x + 3$.
Составляем схему Рауса:

$$\begin{array}{rrrr}
1 & 2 & -1 & 3 \\
-5 & 0 & 1 & \\
2 & -\frac{4}{5} & 3 & \\
-2 & \frac{17}{2} & & \\
\frac{77}{10} & 3 & & \\
\frac{1429}{154} & & & \\
3 & & &
\end{array}$$

В первом столбце этой схемы четыре перемены знака, т. е. $f(x)$ имеет четыре корня с положительной действительной частью. Остальные два корня многочлена $f(x)$ имеют отрицательную действительную часть.

Если при построении схемы Рауса для многочлена (3.31) в ряду $b_0, c_0, d_0, \ldots$ встречается число $h_0 = 0$, причем и вся строка, содержащая h_0, оказывается состоящей из нулей, то нужно эту строку заменить строкой чисел $(n-m+1)e_0$, $(n-m-1)e_1$, $(n-m-3)e_2, \ldots$, где $e_0, e_1, e_2, \ldots$ — элементы предыдущей строки, а m — число уже построенных строк схемы, и продолжать построение схемы Рауса дальше. Если это построение дойдет до конца (быть может, с еще несколькими заменами получающихся нулевых строк соответствующими числами), то опять число перемен знака в первом столбце получившейся схемы будет равно числу корней многочлена (3.31), лежащих в правой полуплоскости. Число чисто мнимых корней этого многочлена (подсчитываемых каждый столько раз, какова его кратность) в данном случае равно $n - m + 1 - 2l$, где m — число строк, получившихся до появления первой нулевой строки, а l — число перемен знака в первом столбце построенной схемы, начиная с m-го места.

Пример 5. $f(x)=x^8-x^7+2x^6-x^5+x^3-2x^2+x-1$.
Составляем схему Рауса:

1	2	0	−2	−1	
−1	−1	1	1		
1	1	−1	−1		
6	4	−2			(заменили строку из нулей соответствующими числами)
$\frac{1}{3}$	$-\frac{2}{3}$	−1			
16	16				
−1	−1				
−2					(опять замена нулевой строки)
−1					

В первом столбце схемы три перемены знака, т. е. три корня многочлена $f(x)$ лежат в правой полуплоскости. Число чисто мнимых корней $f(x)$: $s=8-3+1-2\cdot 1=4$.

Пример 6. $f(x)=3x^5+x^4+6x^3+2x^2+3x+1$.
Составляем схему Рауса:

3	6	3	
1	2	1	
4	4		(замена нулевой строки)
1	1		
2			(опять замена нулевой строки)
1			

В первом столбце схемы перемен знака нет, т. е. $f(x)$ не имеет корней в правой полуплоскости. Число чисто мнимых корней $s=5-2+1-2\cdot 0=4$.

Если получилось $h_0=0$, но в строке, где стоит h_0, имеются и числа, отличные от нуля, то нужно разделить многочлен (3.31) на наибольший общий делитель $D(x)$ многочленов $F_1(x)=a_0x^n+a_1x^{n-2}+a_2x^{n-4}+\dots$ и $F_2(x)=b_0x^{n-1}+b_1x^{n-3}+b_2x^{n-5}+\dots$ (§ 1, п. 5). Если

$$f(x)=D(x)f^*(x), \tag{3.33}$$

то число корней $f(x)$, лежащих в правой полуплоскости, равно t_1+t_2, где t_1 (соответственно t_2) — число корней многочлена $D(x)$ (соответственно $f^*(x)$), лежащих в правой полуплоскости. При этом $t_1=\frac{q-s}{2}$, где q — степень $D(x)$, s — число действительных корней многочлена $D(iy)$ (способ нахождения s см. в п. 2), а t_2 определяется с помощью алгоритма Рауса, примененного к многочлену $f^*(x)$. Если при построении схемы Рауса для $f^*(x)$ в первом столбце получится число $h_0=0$, причем в строке, где оно получилось, будут стоять и числа, отличные от нуля, то нужно будет заменить h_0 малой величиной ε определенного (но произвольного) знака и продолжать построение схемы. Знаки получающихся дальше элементов будут определяться, исходя из малости и знака ε. Если какой-то из элементов первого столбца схемы окажется тождественным нулем относительно ε, то его нужно будет заменить другой малой величиной η и продолжать вычисления. Число перемен знака в первом столбце построенной схемы будет равняться t_2. Число чисто мнимых корней многочлена $f(x)$ в рассматриваемом случае равно s.

Пример 7. $f(x)=x^6-x^4+3x^3-x^2+3x+1$.

Здесь $b_0=0$. Наибольшим общим делителем многочленов $F_1(x)=x^6-x^4-x^2+1$ и $F_2(x)=3x^3+3x$ является $D(x)=x^2+1$, $f(x)=D(x)f^*(x)$, где $f^*(x)=x^4-2x^2+3x+1$. $q=2$, $s=2$, $t_1=0$. Применяем алгоритм Рауса к многочлену $f^*(x)$ (причем вводим малый параметр ε):

$$\begin{array}{lcc}
1 & -2 & 1 \\
\varepsilon & 3 & \\
-2-\dfrac{3}{\varepsilon} & 1 & \\
3+\dfrac{\varepsilon^2}{2\varepsilon+3} & & \\
1 & &
\end{array}$$

При любом достаточно малом по абсолютной величине значении ε в первом столбце этой схемы две перемены знака, $t_2=2$. Многочлен $f(x)$ имеет два корня в правой полуплоскости и два чисто мнимых корня.

в) Теорема Рауса—Гурвица ([6], гл. XV, §§ 6, 8 и 12). Для многочлена (3.31) составляется *матрица Гурвица*

$$H=\begin{pmatrix}
b_0 & b_1 & b_2 & \dots & b_{n-1} \\
a_0 & a_1 & a_2 & \dots & a_{n-1} \\
0 & b_0 & b_1 & \dots & b_{n-2} \\
0 & a_0 & a_1 & \dots & a_{n-2} \\
0 & 0 & b_0 & \dots & b_{n-3} \\
\cdot & \cdot & \cdot & \cdot & \cdot
\end{pmatrix}$$

(квадратная матрица n-го порядка (см. Введение), в которой те элементы a_i, b_i, которые в многочлене (3.31) отсутствуют, равны нулю). Миноры этой матрицы порядков $1, 2, \ldots, n$ (стр. 44), стоящие в левом верхнем углу, называются *определителями Гурвица* и обозначаются через

$$\Delta_1, \Delta_2, \ldots, \Delta_n. \tag{3.34}$$

Если все эти миноры отличны от нуля, то число t корней многочлена (3.31), лежащих в правой полуплоскости, равно числу перемен знака в ряду

$$a_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \frac{\Delta_3}{\Delta_2}, \ldots, \frac{\Delta_n}{\Delta_{n-1}}. \tag{3.35}$$

Чисто мнимых корней многочлен (3.31) в этом случае не имеет.

Пример 8. $f(x) = 3x^5 + x^4 + 4x^3 + 2x^2 + 6x + 2$.

Составляем матрицу Гурвица:

$$H = \begin{pmatrix} 1 & 2 & 2 & 0 & 0 \\ 3 & 4 & 6 & 0 & 0 \\ 0 & 1 & 2 & 2 & 0 \\ 0 & 3 & 4 & 6 & 0 \\ 0 & 0 & 1 & 2 & 2 \end{pmatrix}.$$

$a_0 = 3$, $\Delta_1 = 1$, $\Delta_2 = -2$, $\Delta_3 = -4$, $\Delta_4 = -8$, $\Delta_5 = -16$. В ряду 3, 1, $\frac{-2}{1} = -2$, $\frac{-4}{-2} = 2$, $\frac{-8}{-4} = 2$, $\frac{-16}{-8} = 2$ две перемены знака, $f(x)$ имеет 2 корня в правой полуплоскости.

Если некоторые из миноров (3.34) равны нулю, но $\Delta_n \neq 0$, то необходимо $\Delta_{n-1} \neq 0$, и если

$$\Delta_s \neq 0,\ \Delta_{s+1} = \ldots = \Delta_{s+p} = 0,\ \Delta_{s+p+1} \neq 0, \tag{3.36}$$

то $\Delta_{s-1} \neq 0$ (если $s > 1$), $\Delta_{s+p+2} \neq 0$ и p — число нечетное. Число t корней многочлена (3.31), лежащих в правой полуплоскости, равно числу перемен знака в ряду (3.35), причем только нужно считать, что при условии (3.36) между числами $\frac{\Delta_s}{\Delta_{s-1}}$ и $\frac{\Delta_{s+p+2}}{\Delta_{s+p+1}}$ имеется $\frac{p+2}{2} - (-1)^{\frac{p+1}{2}} \frac{\delta}{2}$ перемен знака (δ есть 1 или -1 в зависимости от того, имеют ли $\frac{\Delta_s}{\Delta_{s-1}}$ и $\frac{\Delta_{s+p+2}}{\Delta_{s+p+1}}$

одинаковые или противоположные знаки); при $s=1$ $\frac{\Delta_s}{\Delta_{s-1}}$ заменяется на Δ_1; если $\Delta_1=\ldots=\Delta_p=0$, $\Delta_{p+1}\neq 0$, то по тому же правилу определяется число перемен знака между числами a_0 и $\frac{\Delta_{p+2}}{\Delta_{p+1}}$ (здесь опять необходимо $\Delta_{p+2}\neq 0$ и p — число нечетное). При $\Delta_n\neq 0$ чисто мнимых корней многочлен (3.31) не имеет.

Пример 9. $f(x)=x^3+x+2$.
Здесь

$$\Delta_3=\begin{vmatrix} 0 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{vmatrix}=-4, \quad \Delta_2=-2,$$

$\Delta_1=0$, $a_0=1$; $p=1$. Число перемен знака между a_0 и $\frac{\Delta_3}{\Delta_2}$: $t=\frac{3}{2}-(-1)\frac{1}{2}=2$, $f(x)$ имеет два корня в правой полуплоскости.

Если $\Delta_n=0$, то нужно представить многочлен (3.31) в виде (3.33) и для определения числа t_2 корней многочлена $f^*(x)$, лежащих в правой полуплоскости, составить определители Гурвица для этого многочлена (тот из них, порядок которого равен степени $f^*(x)$, необходимо отличен от нуля).

Теоремой Рауса—Гурвица удобно пользоваться, когда коэффициенты многочлена зависят от параметров и требуется определить, при каких значениях параметров число t корней многочлена, лежащих в правой полуплоскости, будет иметь то или иное значение.

Некоторые дополнения к теореме Рауса—Гурвица и другие способы определения числа корней многочлена, лежащих в левой или правой полуплоскости, см. в [6], гл. XV и [29], гл. I (особенно §§ 8 и 9).

г) Случай комплексных коэффициентов ([6], гл. XV, § 18). Если $f(x)$ — многочлен с комплексными коэффициентами, то для определения числа корней $f(x)$, лежащих в правой полуплоскости, нужно сначала сделать замену $x=iy$. В результате получится многочлен

$$f(iy)=c_0y^n+c_1y^{n-1}+\ldots+c_n+i\,(d_0y^n+d_1y^{n-1}+\ldots+d_n),$$

где c_k, d_k $(k=0, 1, \ldots, n)$ — действительные числа. Если

$d_0 \neq 0$, то нужно затем составить квадратную матрицу порядка $2n$:

$$A = \begin{pmatrix} d_0 & d_1 & \dots & d_n & 0 & \dots 0 \\ c_0 & c_1 & \dots & c_n & 0 & \dots 0 \\ 0 & d_0 & \dots & d_{n-1} & d_n & \dots 0 \\ 0 & c_0 & \dots & c_{n-1} & c_n & \dots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

(если $d_0 = 0$, то вместо $f(iy)$ нужно взять многочлен $if(iy)$ и для него составить соответствующую матрицу). Миноры этой матрицы (стр. 44) четных порядков, стоящие в левом верхнем углу, обозначим через $\nabla_2, \nabla_4, \dots, \nabla_{2n}$. Если $\nabla_{2n} \neq 0$, то число t корней многочлена $f(x)$, лежащих в правой полуплоскости, равно числу перемен знака в ряду $1, \nabla_2, \nabla_4, \dots, \nabla_{2n}$, причем если в этом ряду имеются нули, то для каждой группы подряд идущих нулей

$$(\nabla_{2h} \neq 0), \quad \nabla_{2h+2} = \dots = \nabla_{2h+2p} = 0, \quad (\nabla_{2h+2p+2} \neq 0)$$

нужно считать, что между ∇_{2h} и $\nabla_{2h+2p+2}$ имеется $\frac{p+1}{2}$ при p нечетном и $\frac{p+1-\delta}{2}$ при p четном перемен знака (δ равно 1 или -1 в зависимости от того, одинаковые или противоположные знаки имеют числа $(-1)^{\frac{p}{2}}$ и $\frac{\nabla_{2h+2p+2}}{\nabla_{2h}}$). Чисто мнимых корней многочлен $f(x)$ при $\nabla_{2n} \neq 0$ не имеет.

Пример 10. $f(x) = ix^3 - x^2 + (5 - 11i)x - (4 + 20i)$. Здесь

$$if(iy) = iy^3 + iy^2 + (-5 + 11i)y - (-20 + 4i) = \\ = -5y + 20 + i(y^3 + y^2 + 11y - 4),$$

$$A = \begin{pmatrix} 1 & 1 & 11 & -4 & 0 & 0 \\ 0 & 0 & -5 & 20 & 0 & 0 \\ 0 & 1 & 1 & 11 & -4 & 0 \\ 0 & 0 & 0 & -5 & 20 & 0 \\ 0 & 0 & 1 & 1 & 11 & -4 \\ 0 & 0 & 0 & 0 & -5 & 20 \end{pmatrix}.$$

Выписываем ряд: 1, $\nabla_2 = 0$, $\nabla_4 = -25$, $\nabla_6 = -15\,000$. $p = 1$; считаем, что между 1 и ∇_4 имеем одну перемену знака. Дальше перемен знака нет, $t = 1$.

Если $\nabla_{2n}=0$, то можно найти наибольший общий делитель $\overline{D}(y)$ (§ 1, п. 5) многочленов $\overline{F}_1(y)=c_0y^n+c_1y^{n-1}+\ldots+c_n$ и $\overline{F}_2(y)=d_0y^n+d_1y^{n-1}+\ldots+d_n$ и разделить $f(x)$ на многочлен $\overline{D}(-ix)$. В частном получится многочлен $\overline{f}(x)$ степени k, для которого $\nabla_{2k}\neq 0$. Число t корней $f(x)$, лежащих в правой полуплоскости, равно сумме числа $\overline{t}_1$ корней $\overline{f}(x)$, лежащих в правой полуплоскости (это число находится указанным выше способом), и числа $\overline{t}_2$ корней $\overline{D}(-ix)$, лежащих в правой полуплоскости (оно равно $\frac{\overline{q}-\overline{s}}{2}$, где $\overline{q}$ — степень, $\overline{s}$ — число действительных корней (§ 4, п. 2) многочлена $\overline{D}(y)$. Число чисто мнимых корней $f(x)$ в данном случае равно $\overline{s}$.

Пример 11. $f(x)=x^5+(-1+2i)x^4+(-1-i)x^3+ix^2+(-2-i)x+(1-i)$.

Здесь

$$f(iy)=iy^5+(-1+2i)y^4+(-1+i)y^3-iy^2+(1-2i)y+(1-i)=$$
$$=(-y^4-y^3+y+1)+i(y^5+2y^4+y^3-y^2-2y-1);\quad \nabla_{10}=0$$

Наибольшим общим делителем многочленов

$$\overline{F}_1(y)=-y^4-y^3+y+1 \text{ и } \overline{F}_2(y)=y^5+2y^4+y^3-y^2-2y-1$$

является

$$\overline{D}(y)=y^4+y^3-y-1.\ \overline{D}(-ix)=x^4+ix^3+ix-1;$$

$$\overline{f}(x)=x-1+i;\ \overline{t}_1=1,\ \overline{s}=2,\ \overline{t}_2=\frac{4-2}{2}=1.$$

$f(x)$ имеет два корня в правой полуплоскости и два чисто мнимых корня.

5. Критерии устойчивости. В ряде прикладных областей, где исследуется устойчивость механических и электрических систем, приходится иметь дело с условиями, при которых все корни многочлена лежат в левой полуплоскости, т. е. имеют отрицательную действительную часть ([11], стр. 115). Многочлен с этим свойством называется *многочленом Гурвица.*

Если $f(x)$ — многочлен Гурвица, имеющий действительные коэффициенты, то все его коэффициенты необходимо положительны ([11], изд. 4, стр. 318).

Критерий Рауса ([6], стр. 425). Чтобы все корни многочлена $f(x)$ с действительными коэффициентами лежали в левой полуплоскости, необходимо и достаточно, чтобы все элементы первого столбца схемы Рауса (п. 4, б)) получались отличными от нуля и одного знака.

Пример 1. $f(x)=x^3+x^2+3x+2$.

Составляем схему Рауса:

$$\begin{matrix} 1 & 3 \\ 1 & 2 \\ 1 & \\ 2 & \end{matrix}$$

Все элементы первого столбца схемы положительны, $f(x)$ — многочлен Гурвица.

Пример 2. $f(x)=3x^5+x^4+6x^3+2x^2+3x+1$ (п. 4, б)).

Этот многочлен не есть многочлен Гурвица, что видно уже из того, что при составлении для него схемы Рауса в первом ее столбце появляется нуль (строка из нулей).

Критерий Гурвица. ([6], стр. 436; [11], изд. 4, стр. 316). Чтобы все корни многочлена $f(x)$ с действительными коэффициентами и положительным старшим коэффициентом лежали в левой полуплоскости, необходимо и достаточно, чтобы все определители Гурвица (п. 4, в)), составленные для этого многочлена, были положительны.

Если старший коэффициент $f(x)$ отрицателен, то нужно рассмотреть многочлен $-f(x)$, корни которого совпадают с корнями $f(x)$.

Пример 3. $f(x)=x^4+2x^3+9x^2+x+4$.

Составляем матрицу Гурвица (п. 4, в)):

$$H=\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 9 & 4 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 9 & 4 \end{pmatrix}.$$

$\Delta_1=2>0$, $\Delta_2=17>0$, $\Delta_3=1>0$, $\Delta_4=4>0$; $f(x)$ — многочлен Гурвица.

Критерий Льенара и Шипара ([6], стр. 457). Необходимые и достаточные условия для того, чтобы все корни многочлена $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ с действительными коэффициентами и $a_0>0$ лежали в левой полуплоскости, могут быть записаны в любом из следующих четырех видов:

1) $a_n > 0,\ a_{n-2} > 0, \ldots;\ \Delta_1 > 0,\ \Delta_3 > 0, \ldots;$

2) $a_n > 0,\ a_{n-2} > 0, \ldots;\ \Delta_2 > 0,\ \Delta_4 > 0, \ldots;$

3) $a_n > 0,\ a_{n-1} > 0,\ a_{n-3} > 0, \ldots;\ \Delta_1 > 0,\ \Delta_3 > 0, \ldots;$

4) $a_n > 0,\ a_{n-1} > 0,\ a_{n-3} > 0, \ldots;\ \Delta_2 > 0,\ \Delta_4 > 0, \ldots$

($\Delta_1, \Delta_2, \Delta_3, \ldots$ — определители Гурвица (п. 4, в)).

Пример 4. $f(x) = x^6 + x^5 + 3x^4 + x^3 + 5x^2 + 2x + 1.$

Здесь все коэффициенты положительны. Составляем матрицу Гурвица:

$$H = \begin{pmatrix} 1 & 1 & 2 & 0 & 0 & 0 \\ 1 & 3 & 5 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & 5 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 & 5 & 1 \end{pmatrix}$$

$\Delta_1 = 1 > 0$, но $\Delta_3 = -1 < 0$, т. е. не все корни данного многочлена имеют отрицательную действительную часть.

Другие критерии устойчивости см. в [6], гл. XV, §§ 5, 14 и 15, [15] и [29].

6. Область устойчивости. Каждому многочлену n-й степени с действительными коэффициентами можно поставить в соответствие точку n-мерного пространства, координаты которой равны частным от деления на старший коэффициент всех остальных коэффициентов многочлена. В таком «пространстве коэффициентов» многочлены Гурвица образуют некоторую n-мерную область, которая называется *областью устойчивости*. Если коэффициенты многочлена заданы как функции p параметров, то область устойчивости строится в пространстве этих параметров. Об исследовании области устойчивости (существенном, например, при проектировании новых систем регулирования) см. [6], гл. XV, §§ 15 и 17.

§ 5. Многочлены от нескольких переменных

1. Определения. *Многочленом* $f(x_1, x_2, \ldots, x_n)$ *от* n *переменных* $x_1, x_2, \ldots, x_n$ *над числовым полем* K (см. введение) называется сумма конечного числа членов вида $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$, где $k_1, k_2, \ldots, k_n$ — целые неотрицательные числа, с коэффициентами из поля K; при этом предполагается, что многочлен подобных членов не содержит, и члены с коэффициентами, равными нулю, не рассматриваются. Два многочлена $f_1(x_1, x_2, \ldots, x_n)$ и $f_2(x_1, x_2, \ldots, x_n)$ называются *равными*, если равны их коэффициенты при одинаковых членах.

Сумма $k_1+k_2+\ldots+k_n$ называется *степенью члена* $x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ многочлена $f(x_1, x_2, \ldots, x_n)$; наибольшая из этих сумм называется *степенью многочлена* $f(x_1, x_2, \ldots, x_n)$ *по совокупности переменных*. Многочлены нулевой степени — это отличные от нуля числа из K. Число нуль также считается многочленом; это единственный многочлен от n переменных, степень которого не определена. Если все члены многочлена $f(x_1, x_2, \ldots, x_n)$ имеют по совокупности переменных одну и ту же степень s, то такой многочлен называется *однородным многочленом* или *формой* s-й степени от n переменных (примеры: линейные формы, квадратичные формы — см. гл. II, § 2, п. 1).

Степенью многочлена $f(x_1, x_2, \ldots, x_n)$ *по отношению к одному из переменных* x_i $(i=1, 2, \ldots, n)$ называется наивысший показатель, с которым x_i входит в члены этого многочлена (эта степень может быть и нулевой).

Из двух членов $x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ и $x_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$ многочлена $f(x_1, x_2, \ldots, x_n)$ тот считается *выше*, у которого показатель при x_1 больше, а если эти показатели равны, то тот, у которого показатель при x_2 больше, и т. д. (т. е. первый член выше второго, если первая из разностей k_i-l_i $(i=1, 2, \ldots, n)$, отличная от нуля, положительна). Если все члены многочлена $f(x_1, x_2, \ldots, x_n)$ расположены в таком порядке, что каждый следующий член ниже предыдущего, то говорят, что члены этого многочлена расположены *лексикографически*. Тот член, который при этом стоит на первом месте, называется *высшим членом* многочлена.

Пример. $f(x_1, x_2, x_3)=x_1^3x_2-2x_1^3x_3-8x_1x_2^5x_3^7+x_2^6x_3+x_2^3$ — лексикографическая запись многочлена, $x_1^3x_2$ — высший член.

Суммой многочленов $f_1(x_1, x_2, \ldots, x_n)$ и $f_2(x_1, x_2, \ldots, x_n)$ называется многочлен, коэффициенты которого получаются сложением соответствующих коэффициентов f_1 и f_2; если при этом какой-то член входит лишь в один из многочленов f_1, f_2, то он считается входящим и в другой многочлен с коэффициентом, равным нулю. *Произведением* многочленов $f_1(x_1, x_2, \ldots, x_n)$ и $f_2(x_1, x_2, \ldots, x_n)$ называется многочлен, полученный почленным умножением f_1 на f_2 с последующим приведением подобных членов. Степень (по совокупности переменных) произведения двух отличных от нуля многочленов от n переменных равна сумме степеней этих многочленов, высший член произведения равен произведению высших членов сомножителей.

2. Симметрические многочлены. Многочлен $f(x_1, x_2, \ldots, x_n)$ от n переменных называется *симметрическим*, если он не меняется ни при какой перестановке этих переменных. Чтобы проверить, является ли данный многочлен симметрическим, достаточно проверить, меняется ли он при перестановках любых двух из входящих в него переменных между собой.

Пример 1. Многочлен $f(x_1, x_2, x_3) = x_1^2x_2x_3 + x_1x_2^2x_3 + x_1^2$ не является симметрическим, так как, если заменить в нем везде x_1 на x_3, а x_3 на x_1, получится многочлен

$$x_3^2x_2x_1 + x_3x_2^2x_1 + x_3^2 \neq f(x_1, x_2, x_3).$$

Пример 2. $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 + 10$ — симметрический многочлен от переменных x_1, x_2, x_3.

Симметрические многочлены от n переменных

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + \ldots + x_n, \\
\sigma_2 &= x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n, \\
\sigma_3 &= x_1x_2x_3 + x_1x_2x_4 + \ldots + x_{n-2}x_{n-1}x_n, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\sigma_{n-1} &= x_1x_2\ldots x_{n-1} + x_1x_2\ldots x_{n-2}x_n + \ldots + x_2x_3\ldots x_n, \\
\sigma_n &= x_1x_2\ldots x_n
\end{aligned}$$

называются *элементарными* (или *основными*) симметрическими многочленами. Если $g(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ — многочлен с коэффициентами из поля K от одного переменного (§ 1, п. 1), то значения элементарных симметрических многочленов от n переменных при значениях переменных, равных корням многочлена $g(x)$, равны соответственно $-\frac{a_1}{a_0}, \frac{a_2}{a_0}, \ldots, (-1)^n\frac{a_n}{a_0}$ (ср. § 1, п. 11).

Симметрический многочлен от переменных $x_1, x_2, \ldots, x_n$, все члены которого могут быть получены из одного из них путем перестановок переменных $x_1, x_2, \ldots, x_n$, называется *моногенным* многочленом ([23], стр. 214). Если $ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ — высший член моногенного многочлена, то этот многочлен обозначается через $S(ax_1^{k_1}x_2^{k_2}\ldots x_n^{k_n})$.

Пример 3. Моногенный многочлен от трех переменных

$$S(x_1^2x_2)=x_1^2x_2+x_1^2x_3+x_2^2x_1+x_2^2x_3+x_3^2x_1+x_3^2x_2.$$

Пример 4. $f(x_1, x_2, x_3)=x_1^2+x_2^2+x_3^2-2x_1x_2-2x_1x_3-$ $-2x_2x_3+10=S(x_1^2)+S(-2x_1x_2)+10.$

Основная теорема о симметрических многочленах. Всякий симметрический многочлен от переменных $x_1, x_2, \ldots, x_n$ можно представить в виде многочлена от элементарных симметрических многочленов (с коэффициентами из того же поля, в котором лежали коэффициенты симметрического многочлена); это представление единственно.

Пример 5. $f(x_1, x_2)=x_1^2+x_2^2-x_1x_2+3=\sigma_1^2-3\sigma_2+3.$

Чтобы найти выражение данного симметрического многочлена через элементарные, нужно сначала разбить этот многочлен на однородные части (п. 1), собирая вместе все члены многочлена, имеющие одну и ту же степень по совокупности переменных, и затем выражать через элементарные симметрические многочлены каждую однородную часть отдельно. Чтобы выразить через элементарные симметрические многочлены однородный симметрический многочлен $f(x_1, x_2, \ldots, x_n)$, нужно взять его высший член $a_0x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$ (п. 1), выписать набор показателей в нем $k_1, k_2, \ldots, k_n$ и составить всевозможные наборы чисел вида $l_1, l_2, \ldots, l_n$ со свойствами: 1) сумма чисел $l_1+l_2+\ldots+l_n$ в каждом наборе одна и та же и равна $k_1+k_2+\ldots+k_n$; 2) числа каждого набора идут, не возрастая, т. е. $l_1\geqslant l_2\geqslant\ldots\geqslant l_n$; 3) член $x_1^{l_1}x_2^{l_2}\ldots x_n^{l_n}$ не выше члена $x_1^{k_1}x_2^{k_2}\ldots x_n^{k_n}$. После этого для каждого набора $l_1, l_2, \ldots, l_n$ нужно составить произведение $\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\ldots\sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n}$ и многочлен $f(x_1, x_2, \ldots, x_n)$ приравнять сумме построенных так произведений, взятых с неопределенными коэффициентами *). Если найти эти коэффициенты, придавая различными способами численные значения переменным $x_1, x_2, \ldots, x_n$ в обеих частях равенства, то

*) Коэффициент при произведении $\sigma_1^{k_1-k_2}\ldots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n}$ берется сразу равным a_0.

получится выражение $f(x_1, x_2, \ldots, x_n)$ через элементарные симметрические многочлены.

Пример 6. $f(x_1, x_2, x_3) = -x_1^3x_2 - x_1x_2^3 - x_1^3x_3 - x_1x_3^3 - x_2^3x_3 - x_2x_3^3 + x_1^2x_2x_3 + x_1x_2^2x_3 + x_1x_2x_3^2$.

Здесь $f(x_1, x_2, x_3)$ — уже однородный симметрический многочлен, его высший член — $x_1^3x_2$. Составляем наборы показателей:

$$\begin{matrix} 3 & 1 & 0 \\ 2 & 2 & 0 \\ 2 & 1 & 1 \end{matrix}$$

Полагаем $f(x_1, x_2, x_3) = -\sigma_1^2\sigma_2 + a\sigma_2^2 + b\sigma_1\sigma_3$. При $x_1 = 1, x_2 = -1$, $x_3 = 0$ имеем: $f(x_1, x_2, x_3) = 2$, $\sigma_1 = 0$, $\sigma_2 = -1$, $\sigma_3 = 0$, откуда $2 = a$. При $x_1 = x_2 = x_3 = 1$ имеем: $f(x_1, x_2, x_3) = -3$, $\sigma_1 = 3$, $\sigma_2 = 3$, $\sigma_3 = 1$, откуда $-3 = -27 + 18 + 3b$, $b = 2$.

$$f(x_1, x_2, x_3) = -\sigma_1^2\sigma_2 + 2\sigma_2^2 + 2\sigma_1\sigma_3.$$

Пример 7. $f(x_1, x_2, \ldots, x_n) = S(x_1^2x_2) + 2x_1^2 + 2x_2^2 + \ldots + 2x_n^2$.

Разбиваем $f(x_1, x_2, \ldots, x_n)$ на однородные части $f_1(x_1, x_2, \ldots, x_n) = S(x_1^2x_2)$ и $f_2(x_1, x_2, \ldots, x_n) = 2x_1^2 + 2x_2^2 + \ldots + 2x_n^2 = S(2x_1^2)$.

Чтобы выразить f_1 через элементарные симметрические многочлены, выписываем наборы показателей

$$\begin{matrix} 2 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 1 & 0 & \ldots & 0 \end{matrix}$$

Это дает $f_1(x_1, x_2, \ldots, x_n) = \sigma_1\sigma_2 + a\sigma_3$. Полагая $x_1 = x_2 = x_3 = 1$, $x_4 = \ldots = x_n = 0$, находим: $6 = 3 \cdot 3 + a \cdot 1$, $a = -3$, $f_1(x_1, x_2, \ldots, x_n) = \sigma_1\sigma_2 - 3\sigma_3$. Аналогично находим, что $f_2(x_1, x_2, \ldots, x_n) = 2\sigma_1^2 - 4\sigma_2$. Таким образом, $f(x_1, x_2, \ldots, x_n) = \sigma_1\sigma_2 - 3\sigma_3 + 2\sigma_1^2 - 4\sigma_2$.

Пользуясь основной теоремой о симметрических многочленах, можно найти значение любого симметрического многочлена от n переменных при значениях входящих в него переменных, равных корням многочлена $g(x)$ n-й степени от одного переменного, не зная самих этих корней.

Пример 8. Найти сумму кубов корней $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ многочлена $g(x) = 7x^4 - 14x^3 - 7x + 2$.

Берем симметрический многочлен $f(x_1, x_2, x_3, x_4) = x_1^3 + x_2^3 + x_3^3 + x_4^3$ и выражаем его через элементарные симметрические многочлены: $f(x_1, x_2, x_3, x_4) = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$. При значениях переменных $x_1 = \alpha_1$, $x_2 = \alpha_2$, $x_3 = \alpha_3$, $x_4 = \alpha_4$ имеем: $\sigma_1 = 2$, $\sigma_2 = 0$, $\sigma_3 = 1$ (см. начало этого пункта). Отсюда $f(\alpha_1, \ldots, \alpha_4) = \alpha_1^3 + \alpha_2^3 + \alpha_3^3 + \alpha_4^3 = 8 - 0 + 3 = 11$.

Пример 9. Найти дискриминант (§ 1, п. 12) многочлена $g(x) = x^3 - 7x - 6$.

Берем симметрический многочлен

$$f(x_1, x_2, x_3) = (x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2.$$

Его высший член находим как произведение высших членов сомножителей, произведению которых равен $f(x_1, x_2, x_3)$ (п. 1); получаем $x_1^4x_2^2$.

Составляем наборы показателей:

$$\begin{matrix} 4 & 2 & 0 \\ 4 & 1 & 1 \\ 3 & 3 & 0 \\ 3 & 2 & 1 \\ 2 & 2 & 2 \end{matrix}$$

и получаем: $f(x_1, x_2, x_3) = \sigma_1^2\sigma_2^2 + a\sigma_1^3\sigma_3 + b\sigma_2^3 + c\sigma_1\sigma_2\sigma_3 + d\sigma_3^2$. Придавая переменным значения $x_1 = x_2 = 1$, $x_3 = 0$, затем $x_1 = x_2 = 1$, $x_3 = -2$, затем $x_1 = x_2 = 2$, $x_3 = -1$, затем $x_1 = x_2 = x_3 = 1$, последовательно находим: $b = -4$, $d = -27$, $a = -4$, $c = 18$. При значениях переменных, равных корням α_1, α_2, α_3 многочлена $g(x)$, получаем $f(\alpha_1, \alpha_2, \alpha_3) = D(g) = 0 + 0 + (-4)(-343) + 0 + (-27)\cdot 36 = 400$.

3. Выражение степенных сумм через элементарные симметрические многочлены. *Степенными суммами* называются симметрические многочлены $s_k = x_1^k + x_2^k + \ldots + x_n^k$ $(k = 1, 2, \ldots)$. С элементарными симметрическими многочленами (см. предыдущий пункт) они связаны *формулами Ньютона* $s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \ldots + (-1)^{k-1} s_1 \sigma_{k-1} + (-1)^k k\sigma_k = 0$ $(k \leqslant n)$, $s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \ldots + (-1)^n s_{k-n}\sigma_n = 0$ (при $k > n$). Из этих формул можно последовательно находить выражения $s_1, s_2, s_3, \ldots$ через $\sigma_1, \sigma_2, \ldots, \sigma_n$ (или наоборот).

Пример 1. Пусть $n \geqslant 3$. Тогда $s_1 = \sigma_1$; $s_2 - s_1\sigma_1 + 2\sigma_2 = 0$, откуда $s_2 = \sigma_1^2 - 2\sigma_2$; $s_3 - s_2\sigma_1 + s_1\sigma_2 - 3\sigma_3 = 0$, откуда $s_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$.

Общие формулы, выражающие $s_k (k = 1, 2, 3, \ldots)$ через $\sigma_1, \sigma_2, \ldots, \sigma_n$ ([23], стр. 211):

$$s_k = k\sum(-1)^{2\lambda_1+3\lambda_2+\ldots+(n+1)\lambda_n} \frac{(\lambda_1 + \ldots + \lambda_n - 1)!\, \sigma_1^{\lambda_1}\ldots\sigma_n^{\lambda_n}}{\lambda_1!\ldots\lambda_n!},$$

где суммирование распространяется на все наборы неотрицательных целых чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ со свойством $\lambda_1 + 2\lambda_2 + \ldots + n\lambda_n = k$ (*первая формула Варинга*).

Пример 2. Найти s_4.

При $k=4$ нужно взять наборы: 1) $\lambda_1=4$, $\lambda_2=\ldots=\lambda_n=0$; 2) $\lambda_1=2$, $\lambda_2=1$, $\lambda_3=\ldots=\lambda_n=0$; 3) $\lambda_1=\lambda_3=1$, $\lambda_2=\lambda_4=\ldots=\lambda_n=0$; 4) $\lambda_2=2$, $\lambda_1=\lambda_3=\ldots=\lambda_n=0$; 5) $\lambda_1=\lambda_2=\lambda_3=0$, $\lambda_4=1$, $\lambda_5=\ldots=\lambda_n=0$. Это дает $s_4=4\left[(-1)^8\frac{3!\,\sigma_1^4}{4!}+(-1)^7\frac{2!\,\sigma_1^2\sigma_2}{2!\,1!}+(-1)^6\frac{1!\,\sigma_1\sigma_3}{1!\,1!}+(-1)^6\frac{1!\,\sigma_2^2}{2!}+(-1)^5\frac{\sigma_4}{1!}\right]=\sigma_1^4-4\sigma_1^2\sigma_2+4\sigma_1\sigma_3+2\sigma_2^2-4\sigma_4$.

Обратно, элементарные симметрические многочлены от n переменных выражаются через степенные суммы по формулам ([23], стр. 213)

$$\sigma_k=\sum\frac{(-1)^{\lambda_1+\ldots+\lambda_k+k}}{1^{\lambda_1}2^{\lambda_2}\ldots k^{\lambda_k}\lambda_1!\ldots\lambda_k!}s_1^{\lambda_1}\ldots s_k^{\lambda_k},$$

где суммирование распространяется на те же наборы чисел λ, что и в первой формуле Варинга (*вторая формула Варинга*).

Пример 3. $\sigma_3=\frac{(-1)^{3+3}}{1^3\cdot 3!}s_1^3+\frac{(-1)^{1+1+3}}{1\cdot 2\cdot 1!\,1!}s_1s_2+\frac{(-1)^{1+3}}{3\cdot 1!}s_3=\frac{1}{6}s_1^3-\frac{1}{2}s_1s_2+\frac{1}{3}s_3$.

4. Системы алгебраических уравнений с несколькими неизвестными. Имеем систему уравнений

$$\left.\begin{array}{l} f_1(x_1, x_2, \ldots, x_n)=0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ f_s(x_1, x_2, \ldots, x_n)=0, \end{array}\right\} \qquad (3.37)$$

где левые части всех уравнений — многочлены от n переменных $x_1, x_2, \ldots, x_n$ с некоторыми числовыми коэффициентами. Говорят, что эта система уравнений имеет решение, если существуют значения переменных $x_1=\alpha_1$, $x_2=\alpha_2, \ldots, x_n=\alpha_n$, при подстановке которых во все уравнения системы эти уравнения обращаются в тождества. Частным случаем системы уравнений (3.37) является система линейных уравнений от n неизвестных; способы решения такой системы см. в гл. I, § 2.

Другой частный случай — система двух уравнений произвольной степени от двух неизвестных

$$\left.\begin{array}{l} f(x, y)=0, \\ g(x, y)=0. \end{array}\right\} \qquad (3.38)$$

Для решения системы (3.38) нужно в каждом из многочленов $f(x, y)$, $g(x, y)$ отнести одно из переменных, например y, к коэффициентам, т. е. записать эти многочлены в виде

$$f(x, y) = a_0(y)\,x^n + a_1(y)\,x^{n-1} + \ldots + a_{n-1}(y)\,x + a_n(y),$$
$$g(x, y) = b_0(y)\,x^s + b_1(y)\,x^{s-1} + \ldots + b_{s-1}(y)\,x + b_s(y).$$

Затем, рассматривая f, g как многочлены только от переменного x, нужно составить результант $R(f, g)$ этих многочленов (§ 1, п. 12) и найти его корни ($R(f, g)$ — многочлен от переменного y). После этого для каждого корня β_0 многочлена $R(f, g)$ нужно найти все общие корни многочленов

$$\left.\begin{aligned} f(x, \beta_0) &= a_0(\beta_0)\,x^n + a_1(\beta_0)\,x^{n-1} + \ldots \\ &\quad \ldots + a_{n-1}(\beta_0)\,x + a_n(\beta_0), \\ g(x, \beta_0) &= b_0(\beta_0)\,x^s + b_1(\beta_0)\,x^{s-1} + \ldots \\ &\quad \ldots + b_{s-1}(\beta_0)\,x + b_s(\beta_0). \end{aligned}\right\} \tag{3.39}$$

Если такими общими корнями являются числа $\gamma_0^{(1)}, \gamma_0^{(2)}, \ldots, \gamma_0^{(k)}$, то пары чисел $(\gamma_0^{(1)}, \beta_0), (\gamma_0^{(2)}, \beta_0), \ldots, (\gamma_0^{(k)}, \beta_0)$ служат решениями системы (3.38). Никаких решений, кроме получающихся описанным способом, система (3.38) не имеет.

Если для корня β_0 результанта $R(f, g)$ $a_0(\beta_0) \neq 0$ или $b_0(\beta_0) \neq 0$, то многочлены (3.39) заведомо имеют хотя бы один общий корень, если же $a_0(\beta_0) = 0$ и $b_0(\beta_0) = 0$, то многочлены (3.39) общих корней могут и не иметь; в последнем случае корню β_0 результанта никакое решение системы (3.38) не соответствует.

Пример 1. Решить систему

$$\begin{cases} x^2y + x^2 + 2xy + y^3 = 0, \\ x^2 - 6x - 3y^2 = 0. \end{cases}$$

Относим переменное y к коэффициентам и составляем результант многочленов $f(x, y) = (y+1)\,x^2 + 2yx + y^3$ и $g(x, y) = x^2 - 6x - 3y^2$:

$$R(f, g) = \begin{vmatrix} y+1 & 2y & y^3 & 0 \\ 0 & y+1 & 2y & y^3 \\ 1 & -6 & -3y^2 & 0 \\ 0 & 1 & -6 & -3y^2 \end{vmatrix} = y^4(4y+3)^2.$$

Корни результанта: $\beta_1=0$, $\beta_2=-\frac{3}{4}$. Рассматриваем многочлены $f(x,\,0)=x^2$ и $g(x,\,0)=x^2-6x$; их общий корень: $\gamma_1=0$. Рассматриваем теперь многочлены $f\left(x,-\frac{3}{4}\right)=\frac{1}{4}x^2-\frac{3}{2}x-\frac{27}{64}$ и $g\left(x,-\frac{3}{4}\right)=x^2-6x-\frac{27}{16}$; их общие корни: $\gamma_2^{(1)}=3+\frac{3}{4}\sqrt{19}$, $\gamma_2^{(2)}=3-\frac{3}{4}\sqrt{19}$. Решения данной системы уравнений: 1) $x=0$, $y=0$; 2) $x=3+\frac{3}{4}\sqrt{19}$, $y=-\frac{3}{4}$; 3) $x=3-\frac{3}{4}\sqrt{19}$, $y=-\frac{3}{4}$.

Пример 2. Решить систему

$$\begin{cases} x^2y^2+xy^2+x+y=0,\\ x^2y\;\;+xy\;\;+1\;\;\;\;=0.\end{cases}$$

Здесь удобнее отнести к коэффициентам переменное x. Получаем многочлены $f(x,y)=(x^2+x)\,y^2+y+x$, $g(x,y)=(x^2+x)\,y+1$ и составляем их результант:

$$R(f,g)=\begin{vmatrix} x^2+x & 1 & x\\ x^2+x & 1 & 0\\ 0 & x^2+x & 1\end{vmatrix}=(x+1)^2x^2.$$

Корни результанта: $\gamma_1=0$, $\gamma_2=-1$. При этих значениях переменного x старшие коэффициенты многочленов $f(x,y)$ и $g(x,y)$ обращаются в нуль, причем получающиеся многочлены общих корней не имеют ($g(0,y)=g(-1,y)\equiv 1$). Следовательно, заданная система уравнений не имеет решений.

Еще о решении систем алгебраических уравнений с несколькими неизвестными см. [13], гл. III, § 5.

ГЛАВА IV
ОБЩАЯ АЛГЕБРА

§ 1. Группы

1. Определение группы, примеры. Пусть G — какое-то множество ([34], стр. 23—24) элементов (например, множество чисел, или функций, или каких-нибудь объектов геометрической природы и т. д.). Говорят, что в множестве G определена *алгебраическая операция*, если каждым двум элементам a, b из G, взятым в определенном порядке, однозначным образом поставлен в соответствие некоторый третий элемент c из G.

Примеры алгебраических операций: сложение целых чисел (каждым двум целым числам ставится в соответствие их сумма), их вычитание; сложение векторов на плоскости; сложение или умножение квадратных матриц n-го порядка (гл. II, § 1, п. 1); векторное умножение векторов трехмерного пространства.

Алгебраическая операция, определенная в множестве G, обычно называется или *умножением*, или *сложением*. В первом случае, если элементам a, $b \in G$ поставлен в соответствие элемент $c \in G$, пишут $c = ab$, во втором случае пишут $c = a + b$.

Множество G с определенной в нем алгебраической операцией (пусть она сейчас называется умножением) является *группой*, если выполнены следующие условия:

1) для любых трех элементов a, b, $c \in G$

$$a(bc) = (ab)c$$

(*ассоциативность* умножения);

2) в множестве G существует такой элемент e, что для любого $a \in G$

$$ae = ea = a;$$

3) для любого $a \in G$ существует такое $a' \in G$, что

$$aa' = a'a = e.$$

Элемент e называется *единицей* группы, элемент a' называется элементом, *обратным* a, и обозначается обычно через a^{-1}. Если операция, определенная в группе, называется сложением, то вместо единицы группы говорят о *нуле* группы (это элемент, обозначаемый символом 0 и обладающий свойством $a+0=0+a=a$ для любого $a \in G$), а вместо элемента a^{-1}, обратного элементу a, говорят об элементе $-a$, *противоположном* a $(a+(-a)=0)$.

Примерами групп являются:

1) множество всех целых чисел относительно операции сложения (так называемая *аддитивная группа* всех целых чисел);

2) множество всех четных чисел относительно операции сложения (аддитивная группа всех четных чисел);

3) множество всех отличных от нуля рациональных чисел относительно операции умножения (*мультипликативная группа* отличных от нуля рациональных чисел);

4) множество всех векторов на плоскости относительно обычного сложения векторов;

5) множество всех невырожденных квадратных матриц n-го порядка с действительными элементами относительно операции умножения матриц (гл. II, § 1, п. 1).

Целые числа относительно операции умножения группы не образуют, так как для целого числа, отличного от ± 1, не существует обратного ему целого числа. По той же причине не образуют группы относительно умножения все квадратные матрицы n-го порядка. Не является группой и множество всех векторов трехмерного пространства относительно векторного умножения векторов, так как эта операция не ассоциативна.

Второе определение группы (равносильное приведенному выше). Множество G с определенной в нем алгебраической операцией называется группой, если:

1) эта операция ассоциативна,

2) для любых двух элементов $a, b \in G$ существует ровно один такой элемент $x \in G$, что $ax = b$, и ровно один такой элемент $y \in G$, что $ya = b$.

Если операция, определенная в группе, *коммутативна* (т. е. для любых элементов a, b группы $ab = ba$), то сама группа называется *коммутативной* или *абелевой*. Группы приведенных выше примеров 1, 2, 3, 4 абелевы, группа примера 5 не абелева. Еще пример не абелевой группы — множество всех вращений трехмерного пространства вокруг начала координат, где произведением двух вращений называется их последовательное выполнение.

Если группа G состоит из конечного числа элементов, то она называется *конечной группой*, а число элементов в ней называется *порядком группы*.

Пример 1. Произведением двух подстановок n-й степени (гл. I, § 1, п. 1) называется результат их последовательного выполнения (например, если $n = 4$ и $a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$, то $ab = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$, так как, например, число 1 при подстановке a переходит в 2, а число 2 при подстановке b переходит в 4, т. е. в итоге 1 переходит в 4, и т. д.). Это — алгебраическая операция, относительно которой множество всех подстановок n-й степени является группой. Единицей группы служит тождественная подстановка $\begin{pmatrix} 1 & 2 \dots & n \\ 1 & 2 \dots & n \end{pmatrix}$, элементом, обратным к подстановке $\begin{pmatrix} 1 & 2 \dots & n \\ i_1 & i_2 \dots & i_n \end{pmatrix}$, является подстановка $\begin{pmatrix} i_1 & i_2 \dots & i_n \\ 1 & 2 \dots & n \end{pmatrix}$. Эта группа называется *симметрической группой n-й степени* и обычно обозначается через S_n. Это — конечная группа порядка $n!$ При $n \geqslant 3$ она некоммутативна (например, при $n = 4$ для указанных выше подстановок a, b имеем: $ba = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \neq ab$).

Если группа конечна, то умножение в ней часто задается с помощью *таблицы Кэли*, в первой строке и первом столбце которой выписываются все элементы группы, начиная с единицы (пусть эти элементы: $a_1 = e, a_2, \dots, a_n$), а затем на пересечении строки, где стоит элемент a_i, и столбца, где стоит элемент a_j, пишется элемент, равный произведению $a_i a_j$.

Пример 2. В симметрической группе 3-й степени пусть

$$a_1 = e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\quad a_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},\quad a_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$
$$a_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\quad a_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},\quad a_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Тогда таблица Кэли имеет вид

	e	a_2	a_3	a_4	a_5	a_6
e	e	a_2	a_3	a_4	a_5	a_6
a_2	a_2	e	a_4	a_3	a_6	a_5
a_3	a_3	a_5	e	a_6	a_2	a_4
a_4	a_4	a_6	a_2	a_5	e	a_3
a_5	a_5	a_3	a_6	e	a_4	a_2
a_6	a_6	a_4	a_5	a_2	a_3	e

2. Изоморфизм групп. Говорят, что между элементами множеств (см. введение) M и N установлено *взаимно однозначное соответствие*, если каждому элементу множества M поставлен в соответствие некоторый вполне определенный элемент множества N, причем различным элементам из M соответствуют различные элементы в N и всякий элемент из N соответствует некоторому элементу из M.

Две группы G и H называются *изоморфными*, если между их элементами можно установить взаимно однозначное соответствие, при котором для любых элементов $a, b \in G$ и соответствующих им элементов $a', b' \in H$ элементу $ab = c$ будет соответствовать элемент $c' = a'b'$.

Пример 1. Аддитивная группа G всех целых чисел (п. 1) изоморфна аддитивной группе H всех четных чисел (для установления изоморфизма между ними можно каждому числу $k \in G$ поставить в соответствие число $2k \in H$).

Пример 2. Мультипликативная группа (п. 1) всех положительных действительных чисел изоморфна аддитивной группе всех действительных чисел (изоморфизм: $a \to \lg a$).

Изоморфное отображение группы на себя называется *автоморфизмом* этой группы.

Пример 3. Одним из автоморфизмов аддитивной группы всех целых чисел является отображение, при котором каждому целому числу a ставится в соответствие число $-a$.

Если две группы изоморфны, то изоморфное соответствие между ними можно установить, вообще говоря, многими способами. Изоморфные группы могут отличаться друг от друга только природой своих элементов и, быть может, названием операции, определенной в группе. Но все свойства изоморфных между собой групп, вытекающие из свойств определенных в них операций и не зависящие от природы элементов группы, одинаковы. Например, если группа абелева (п. 1), то и все изоморфные ей группы абелевы. В алгебре изоморфные друг другу группы не считаются существенно различными.

3. Гомоморфизм. Пусть каждому элементу группы G соответствует однозначно определенный элемент группы H, причем если элементам $a, b \in G$ соответствуют элементы $a', b' \in H$, то элементу $ab = c$ соответствует элемент $c' = a'b'$. Такое отображение $G \to H$ называется *гомоморфизмом*; говорят, что *группа G гомоморфно отображена в группу H*. Если при этом на каждый элемент группы H отображается хотя бы один элемент группы G, то говорят о *гомоморфном отображении* группы G *на* группу H; гомоморфизм в этом случае называется *эпиморфизмом*.

Вообще говоря, при гомоморфизме в данный элемент группы H могут переходить различные элементы группы G, а также может не переходить ни одного.

Пример 1. Если каждому четному числу поставить в соответствие число 1, а каждому нечетному поставить в соответствие число -1, то получится гомоморфное отображение аддитивной группы всех целых чисел в мультипликативную группу всех отличных от нуля рациональных чисел (п. 1, примеры 1 и 3). Это отображение будет также гомоморфным отображением аддитивной группы всех целых чисел на мультипликативную группу, состоящую из чисел -1, 1.

Пример 2. Если каждой невырожденной квадратной матрице n-го порядка с действительными элементами поставить в соответствие определитель этой матрицы (гл. I, § 1, п. 2), то получится гомоморфное отображение группы (по умножению) всех действительных невырожденных квадратных матриц n-го порядка на мультипликативную группу всех отличных от нуля действительных чисел.

При всяком гомоморфном отображении группы G в группу H единица группы G (или нуль, если групповая операция — сложение) переходит в единицу (в нуль) группы H. Совокупность всех элементов группы G, переходящих в единицу (в нуль) группы H, называется *ядром* данного гомоморфизма. В приведенном выше примере 1 ядро гомоморфизма составляют все четные числа, в примере 2 — все матрицы с определителем, равным 1.

Гомоморфное отображение группы в себя называется *эндоморфизмом* этой группы. Примером эндоморфизма может служить отображение, при котором каждому элементу группы ставится в соответствие единица этой группы.

Если группа G гомоморфно отображена на какое-то множество M, в котором определена алгебраическая операция (п. 1), то относительно этой алгебраической операции множество M само необходимо является группой.

4. Подгруппы. Циклические группы. Подмножество A (см. введение) группы G называется *подгруппой* этой группы, если вместе с каждым элементом a оно содержит также обратный ему элемент a^{-1} и вместе с каждыми двумя элементами a, b оно содержит и их произведение ab. Эти два требования можно заменить одним: для любых элементов $a, b \in A$ элемент ab^{-1} должен лежать в A.

Всякая подгруппа данной группы G сама является группой относительно той операции, которая определена в G.

Пример 1. Аддитивная группа всех четных чисел (п. 1) является подгруппой аддитивной группы всех целых чисел, которая сама является подгруппой аддитивной группы всех комплексных чисел.

Пример 2. Множество всех четных подстановок n элементов (гл. I, § 1, п. 1) является подгруппой в группе всех подстановок n элементов (п. 1). Эта подгруппа называется *знакопеременной группой n-й степени* и обычно обозначается через A_n.

В любой группе подмножество, состоящее только из единичного элемента группы, является подгруппой. Эта подгруппа называется *единичной подгруппой* данной группы и обычно обозначается символом E. Сама группа также всегда является своей подгруппой. Всякая подгруппа, отличная от всей группы, называется *истинной подгруппой* этой группы.

Подгруппами симметрических групп (п. 1) исчерпываются, по существу, все конечные группы: всякая конечная группа порядка n (п. 1) изоморфна (п. 2) некоторой подгруппе группы всех подстановок n-й степени (*теорема Кэли*).

Если в группе G взять какой-нибудь элемент g и все степени этого элемента*) (или все его кратные, если опера-

*) g^n при $n > 0$ — произведение n элементов, равных g; а g^{-n} — произведение n элементов, равных g^{-1}, или, что то же, элемент, обратный g^n: $g^{-n} = (g^{-1})^n = (g^n)^{-1}$; $g^0 = e$ (по определению).

ция в группе — сложение), то также получится подгруппа группы G. Эта подгруппа называется *циклической подгруппой, порожденной элементом* g, и обозначается через $\{g\}$. Если подгруппа $\{g\}$ совпадает со всей группой G, то сама группа G называется *циклической группой.*

Если для элемента g группы G существует такое натуральное число k, что $g^k = e$ (e — единица группы G), то g называется *элементом конечного порядка.* Наименьшее натуральное число n со свойством $g^n = e$ в этом случае называется *порядком элемента* g. Если ни для какого натурального числа k равенство $g^k = e$ не выполнено, то g называется *элементом бесконечного порядка*; в этом случае все степени элемента g (включая и отрицательные) различны между собой; в частности, ни одна из этих степеней не равна e, кроме $g^0 = e$.

Если элемент g группы G имеет бесконечный порядок, то циклическая подгруппа $\{g\}$ бесконечна. Если g — элемент порядка n, то подгруппа $\{g\}$ также имеет порядок n (п. 1); она состоит из (различных между собой) элементов e, g, g^2, ..., g^{n-1}, которым равны все остальные степени элемента g, положительные и отрицательные.

Пример 3. Аддитивная группа всех целых чисел (п. 1) есть бесконечная циклическая группа (она состоит из всех кратных числа 1).

Пример 4. Все значения корня n-й степени из 1 образуют (относительно умножения) циклическую группу порядка n, порожденную любым из первообразных корней n-й степени из 1 ([11], стр. 128).

Все бесконечные циклические группы изоморфны между собой (п. 2). Изоморфны между собой также все конечные циклические группы одного и того же порядка n. Всякая подгруппа циклической группы сама является циклической группой.

5. Системы образующих. Возрастающие последовательности подгрупп. Пусть M — какое-то непустое подмножество группы G.

Совокупность всех элементов группы G, равных произведениям (положительных и отрицательных) степеней элементов из M (в каждое произведение входит конечное число множителей), является подгруппой группы G. Эта подгруппа называется *подгруппой, порожденной* множеством M, и обозначается через $\{M\}$. Общий вид элемента подгруппы $\{M\}$: $a_{\alpha_1}^{k_1} \ldots a_{\alpha_n}^{k_n}$, где каждое k_i $(i=1, \ldots, n)$ — какое-то целое число, а $a_{\alpha_1}, \ldots, a_{\alpha_n}$ — элементы множества M, среди которых могут быть и одинаковые (например, $a_1^2 a_2^{-1} a_1^{-3} a_3 a_2^2 \in \{M\}$, если $a_1, a_2, a_3 \in M$). Подгруппа $\{M\}$ содержится в любой подгруппе группы G, содержащей целиком множество M, и является пересечением всех подгрупп группы G, содержащих M.

Множество M называется *системой образующих* подгруппы $\{M\}$. Если подгруппа $\{M\}$ совпадает со всей группой G, то M — система образующих группы G. Если при этом множество M конечно, то G называется группой *с конечным числом образующих*.

Пример 1. Множество всех элементов группы G всегда является ее системой образующих.

Пример 2. Одной из систем образующих мультипликативной группы (п. 1) всех положительных рациональных чисел является множество всех простых чисел.

Пример 3. Число 1 является системой образующих аддитивной группы всех целых чисел (п. 1). Другая система образующих той же группы — числа 2, 3.

Если множество M состоит из одного элемента, то порожденная им подгруппа есть циклическая подгруппа, порожденная этим элементом (п. 4).

Если множество M состоит из всех элементов, принадлежащих некоторым подгруппам $A, B, \ldots$ группы G, то подгруппа $\{M\}$ называется *подгруппой, порожденной заданными подгруппами*, и обозначается через $\{A, B, \ldots\}$. Подгруппа, порожденная некоторым множеством подгрупп группы G, состоит из всех элементов группы, равных произведениям элементов, взятых из этих подгрупп. Это — минимальная подгруппа группы G, содержащая все заданные подгруппы.

Пусть подгруппы $A_1, A_2, \ldots, A_n, \ldots$ группы G таковы, что

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots \tag{4:1}$$

(*возрастающая последовательность подгрупп*). Тогда подгруппа, порожденная этими подгруппами, совпадает просто с множеством всех элементов группы G, каждый из которых принадлежит хоть одной из этих подгрупп, и называется *объединением возрастающей последовательности под-*

групп (4.1). Может случиться, что объединение возрастающей последовательности подгрупп группы G совпадает с самой этой группой.

Пример 4. Аддитивная группа (п. 1) всех рациональных чисел является объединением возрастающей последовательности своих циклических подгрупп

$$\{1\} \subset \left\{\frac{1}{2}\right\} \subset \left\{\frac{1}{6}\right\} \subset \ldots \subset \left\{\frac{1}{n!}\right\} \subset \ldots$$

Пример 5. Множество всех комплексных чисел, являющихся корнями из 1 ([11], стр. 127) степеней, равных степеням некоторого фиксированного простого числа p, представляет собой группу относительно умножения. Эта группа есть объединение возрастающей последовательности (циклических) подгрупп $A_1 \subset A_2 \subset \ldots$, где A_l ($l = 1, 2, \ldots$) состоит из всех корней p^l-й степени из 1. Эта группа (как и всякая группа, ей изоморфная) называется *группой типа* p^∞ или *квазициклической* группой. Группы типа p^∞ играют большую роль в теории абелевых групп.

6. Смежные классы. Разложение группы по подгруппе. Если A — подгруппа, g — произвольный элемент группы G, то через gA обозначается множество всех элементов группы G, получающихся при умножении элемента g на всевозможные элементы из подгруппы A (т. е. множество всех элементов вида ga, где $a \in A$). Это множество называется *левым смежным классом группы G по подгруппе A, определяемым элементом g*. Аналогично *правым смежным классом* Ag группы G по подгруппе A, определяемым элементом g, называется множество всех элементов вида ag, где $a \in A$. Если групповая операция — сложение, то левые и правые смежные классы обозначаются соответственно через $g+A$ и $A+g$.

Пример 1. В группе всех подстановок третьей степени (п. 1) возьмем подгруппу A, состоящую из подстановок $e = \begin{pmatrix}1 & 2 & 3\\ 1 & 2 & 3\end{pmatrix}$ и $a = \begin{pmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{pmatrix}$, и возьмем элемент $g = \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}$. Тогда левый смежный класс gA будет состоять из подстановок $ge = g = \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}$ и $ga = \begin{pmatrix}1 & 2 & 3\\ 3 & 2 & 1\end{pmatrix}$, а правый смежный класс Ag — из подстановок $eg = g = \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix}$ и $ag = \begin{pmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{pmatrix}$.

Каждый левый смежный класс определяется любым из входящих в него элементов, т. е. если $g_1 \in gA$, то $g_1A = gA$.

Два любых левых смежных класса группы G по подгруппе A или совпадают, или не имеют ни одного общего элемента. Одним из левых смежных классов группы G по подгруппе A является сама подгруппа A ($A = eA$, где e — единица группы G); все остальные левые смежные классы по подгруппе A подгруппами группы G не являются.

Эти же утверждения верны для правых смежных классов.

Число всех различных левых смежных классов группы G по подгруппе A всегда равно числу всех различных правых смежных классов группы G по этой же подгруппе (в бесконечном случае это означает, что мощность множества всех различных левых смежных классов группы G по подгруппе A совпадает с мощностью множества правых смежных классов). Это число (в бесконечном случае — мощность) называется *индексом* подгруппы A в группе G. Если группа G конечна, то ее порядок (п. 1) равен произведению порядка любой ее подгруппы A на индекс этой подгруппы в группе G (*теорема Лагранжа*). Отсюда следует, что порядок всякой подгруппы конечной группы является делителем порядка группы. Также порядок любого элемента (п. 4) конечной группы является делителем порядка этой группы. Обратно, если порядок конечной группы G делится на простое число p, то G обладает элементами порядка p (*теорема Коши*).

Если в произвольной группе G выбрана какая-то подгруппа A, то все элементы группы G можно разбить на непересекающиеся классы, объединяя вместе те элементы, которые лежат в одном и том же левом смежном классе группы G по подгруппе A. Такое разбиение называется *левосторонним разложением группы G по подгруппе A*. Если вместо левых смежных классов взять правые смежные классы по подгруппе A, то получится *правостороннее разложение группы G по подгруппе A*.

Пример 2. В симметрической группе 3-й степени S_3 (п. 1), элементы которой $a_1 = e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$, $a_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $a_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, $a_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $a_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, $a_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, возьмем подгруппу A, состоящую из элементов e и a_6. Тогда при левостороннем разложении группы S_3 по подгруппе A множество ее элементов разобьется на классы: 1) e, a_6 (смежный класс $eA = A$), 2) a_2, a_5 (смежный класс a_2A), 3) a_3, a_4 (смежный класс a_3A), а при правостороннем разложении получатся классы: 1) e, a_6 (смежный класс $Ae = A$), 2) a_2, a_4

(смежный класс Aa_2), 3) a_3, a_5 (смежный класс Aa_3). Индекс подгруппы A в группе S_3 равен 3.

Пример 3. Как при левостороннем, так и при правостороннем разложении аддитивной группы всех целых чисел (п. 1) по подгруппе, состоящей из всех целых чисел, делящихся на 3, в один класс будут объединяться те целые числа, которые при делении на 3 дают один и тот же остаток. Всего получится три класса: 1) класс B_0 чисел, делящихся на 3 без остатка, 2) класс B_1 чисел, дающих при делении на 3 в остатке 1, и 3) класс B_2 чисел, дающих при делении на 3 в остатке 2.

Если в произвольной группе G в качестве подгруппы A взять саму группу G, то (как правостороннее, так и левостороннее) разложение группы G по подгруппе A будет состоять только из одного класса (в который войдут все элементы группы G). Если же в качестве подгруппы A взять единичную подгруппу E (п. 4), то каждый элемент группы G будет составлять отдельный смежный класс.

7. Нормальный делитель группы. Если при левостороннем и при правостороннем разложении группы G по некоторой ее подгруппе H (п. 6) классы, на которые распадаются элементы группы G, получаются одинаковыми, то подгруппа H называется *нормальным делителем* группы G (или *инвариантной подгруппой.*)

Пример 1. В примере 2 п. 6 подгруппа A группы S_3 не является нормальным делителем этой группы.

Пример 2. Если индекс подгруппы H в группе G равен 2, то H — нормальный делитель группы G: и при левостороннем, и при правостороннем разложении группы G по подгруппе H одним из смежных классов является сама подгруппа H, а второй смежный класс состоит из всех элементов группы G, не лежащих в H. Например, знакопеременная группа n-й степени (п. 4), индекс которой в симметрической группе n-й степени S_n (п. 1) равен 2, является нормальным делителем группы S_n.

Подгруппа H тогда и только тогда является нормальным делителем группы G, когда для любого элемента g группы G

$$gH = Hg.$$

Это равенство означает, что для всякого элемента h из H можно найти в H такие элементы h' и h'', что

$$gh = h'g, \qquad hg = gh''. \tag{4.2}$$

Пример 3. Если группа G абелева (п. 1), то всякая ее подгруппа H является нормальным делителем (для получения равенств (4.2) достаточно взять $h' = h'' = h$).

Элементы g_1 и g_2 группы G называются *сопряженными* в этой группе, если в G существует такой элемент g, что

$$g_2 = g^{-1} g_1 g$$

(в этом случае говорят также, что элемент g_2 получен из g_1 *трансформированием* элементом g).

Если A — подгруппа, g — фиксированный элемент некоторой группы G, то совокупность всех элементов вида $g^{-1}ag$, где a пробегает всю подгруппу A, есть снова подгруппа группы G. Эта подгруппа называется подгруппой, *сопряженной* с A в группе G.

Пример 4. Подгруппа A_1 группы S_3 примера 2 п. 6, состоящая из элементов e и a_2, сопряжена с рассматривавшейся там подгруппой A: $A_1 = a_3^{-1} A a_3$.

Подгруппа H группы G тогда и только тогда является нормальным делителем этой группы, когда вместе с каждым элементом h она содержит и все элементы, сопряженные с h в группе G.

Пример 5. Пусть G — группа всех невырожденных квадратных матриц n-го порядка с действительными элементами (п. 1, пример 5), H — ее подгруппа, состоящая из всех матриц, определитель которых (гл. I, § 1, п. 2) равен 1. Тогда если матрица A лежит в H, то в H лежит и всякая матрица вида $B^{-1}AB$ (B — любая невырожденная матрица), так как $|B^{-1}AB| = |B^{-1}||A||B| = |A| = 1$ (стр. 34, 71). Следовательно, H — нормальный делитель группы G.

Нормальные делители группы G и только они совпадают со всеми подгруппами, сопряженными с ними в G.

Всякий нормальный делитель группы G является ядром некоторого гомоморфизма (п. 3) этой группы; обратно, ядро всякого гомоморфизма группы G является нормальным делителем в G.

В любой группе единичная подгруппа (п. 4) и сама группа являются нормальными делителями. Если других нормальных делителей в группе нет, она называется *простой* группой. Примером простой группы служит знакопеременная группа n-й степени (п. 4) при $n \geqslant 5$. Существуют и бесконечные простые группы.

Если в группе G все подгруппы являются нормальными делителями, причем группа G не абелева (п. 1), то она называется *гамильтоновой* группой. Пример такой группы см. в [12], стр. 58.

8. Фактор-группа. Если в множестве всех смежных классов группы G по нормальному делителю H (пп. 6 и 7) ввести операцию по правилу $(g_1H)(g_2H)=(g_1g_2)H$ (при аддитивной записи $(g_1+H)+(g_2+H)=(g_1+g_2)+H$), то это будет алгебраическая операция *) (п. 1), относительно которой множество всех смежных классов группы G по нормальному делителю H само окажется группой. Эта группа называется *фактор-группой группы G по нормальному делителю H* и обозначается через G/H. Единицей фактор-группы G/H является смежный класс H. Элементом, обратным элементу фактор-группы gH, является смежный класс $g^{-1}H$. Порядок фактор-группы G/H равен индексу H в G (пп. 1 и 6).

Пример 1. Фактор-группа аддитивной группы всех целых чисел (п. 1) по подгруппе всех чисел, делящихся на 3, состоит из трех элементов — классов B_0, B_1 и B_2 (п. 6, пример 3). Это — циклическая группа 3-го порядка, порожденная элементом B_1 (п. 4) (групповая операция — сложение).

Пример 2. В мультипликативной группе G всех отличных от нуля рациональных чисел (п. 1) числа $1, -1$ составляют нормальный делитель H. Смежные классы по этому нормальному делителю являются парами чисел $c, -c$, где c — некоторое положительное рациональное число. Произведением смежных классов $c_1, -c_1$ и $c_2, -c_2$ является смежный класс группы G по нормальному делителю H, состоящий из чисел $c_1c_2, -c_1c_2$ $(c_1Hc_2H=(c_1c_2)H)$. Фактор-группа G/H изоморфна (п. 2) мультипликативной группе всех положительных рациональных чисел (этот изоморфизм можно получить, ставя в соответствие смежному классу $c, -c$ число c).

Пример 3. Если в мультипликативной группе G всех невырожденных квадратных матриц n-го порядка с действительными элементами взять нормальный делитель H, состоящий из всех матриц, определитель которых равен 1 (п. 7, пример 5), то смежные классы по H будут состоять из всех матриц с одинаковым определителем. Если каждому смежному классу поставить в соответствие действительное число, равное определителю всех

---

*) Если $g_1'H=g_1H$, $g_2'H=g_2H$, то $(g_1'g_2')H=(g_1g_2)H$, т. е. смежный класс $(g_1g_2)H$ не зависит от выбора элементов g_1, g_2 в данных смежных классах g_1H, g_2H.

матриц, входящих в этот смежный класс, то получится изоморфизм между фактор-группой G/H и мультипликативной группой всех отличных от нуля действительных чисел (пп. 1 и 2).

Пример 4. Пусть G — аддитивная группа всех векторов на плоскости, выходящих из начала координат, H — ее подгруппа, состоящая из всех векторов, лежащих на оси ординат. Тогда H — нормальный делитель группы G, а каждый смежный класс группы G по H представляет собой совокупность всех векторов плоскости, выходящих из начала координат и имеющих конец на некоторой (фиксированной) прямой, параллельной оси ординат. Фактор-группа G/H изоморфна (п. 2) аддитивной группе всех действительных чисел (для установления изоморфизма между этими группами можно каждому смежному классу группы G по нормальному делителю H поставить в соответствие то действительное число, которое изображается точкой пересечения прямой, на которой лежат концы всех векторов данного смежного класса, с осью абсцисс).

Если каждому элементу g группы G поставить в соответствие смежный класс gH группы G по некоторому ее нормальному делителю H, то получится гомоморфное отображение группы G на фактор-группу G/H. Оно называется *естественным гомоморфным отображением* G на G/H. Ядром этого гомоморфизма (п. 3) является нормальный делитель H. Совокупность всех элементов группы G, переходящих при естественном гомоморфизме в один и тот же фиксированный элемент группы G/H, является смежным классом группы G по нормальному делителю H.

Если группа G гомоморфно отображена на какую-то группу $\bar{G}$ и H — ядро этого гомоморфизма, то необходимо H является нормальным делителем группы G, а группа $\bar{G}$ изоморфна (п. 2) фактор-группе G/H (*теорема о гомоморфизмах групп*).

Пример 5. Пусть $\{a\}$ — бесконечная циклическая группа, $\{b\}$ — циклическая группа порядка n (п. 4). Если каждому элементу a^k группы $\{a\}$ (k — любое целое число) поставить в соответствие элемент b^k группы $\{b\}$, то получится гомоморфное отображение группы $\{a\}$ на группу $\{b\}$. Ядром этого гомоморфизма является циклическая подгруппа группы $\{a\}$, порожденная элементом a^n. Фактор-группа $\{a\}/\{a^n\}$ изоморфна группе $\{b\}$.

9. Группы подстановок ([27]; [23]; [5], ч. 1). *Группы подстановок* — это подгруппы симметрических групп (п. 1). При их изучении обычно для удобства пользуются записью подстановок в виде произведений циклов. *Циклом* или *циклической подстановкой* называется такая подстановка чисел $1, 2, \ldots, n$ *), которая одно из этих чисел i_1 переводит в число i_2, i_2 переводит в i_3 и т. д., i_{k-1} переводит в i_k ($k \leqslant n$), i_k переводит в i_1, а все остальные числа оставляет на месте. Эта подстановка обозначается через $(i_1, i_2, \ldots, i_k)$. Циклы $(i_1, i_2, \ldots, i_k)$ и, например, $(i_2, \ldots, i_k, i_1)$ равны между собой. Число k называется *длиной* цикла. Цикл длины 1 — это тождественная подстановка. Чтобы представить произвольную подстановку чисел $1, 2, \ldots, n$ в виде произведения циклов, нужно взять любое из этих чисел (например, 1), затем число, в которое оно переводится данной подстановкой (пусть это число — j_2), затем число, в которое j_2 переходит при этой подстановке, и т. д. до числа j_l, переводящегося данной подстановкой в первое из взятых чисел (в нашем случае в 1), и выписать цикл $(1, j_2, \ldots, j_l)$; после этого нужно взять какое-нибудь из чисел $1, 2, \ldots, n$, которое пока еще не встречалось (если такое число существует), и, начиная с него, сделать то же самое и т. д., пока не будут использованы все числа $1, 2, \ldots, n$. Тогда мы получим несколько циклов, не имеющих общих действительно перемещаемых ими чисел, произведению которых (в любом порядке) и будет равна данная подстановка. Циклы, не содержащие общих действительно перемещаемых ими чисел, называются *независимыми*. Те циклы, входящие в произведение, длина которых равна 1, обычно не пишут. При этом условии всякая подстановка разлагается в произведение независимых циклов единственным образом.

Пример 1. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 2 & 1 & 5 & 7 & 4 & 6 & 3 \end{pmatrix} = (183)(4576)$.

Пример 2. Подстановки, составляющие симметрическую группу 3-й степени, записываются в виде произведений циклов так: $a_1 = (1)$, $a_2 = (2,3)$, $a_3 = (1,2)$, $a_4 = (1,2,3)$, $a_5 = (1,3,2)$, $a_6 = (1,3)$.

Цикл длины 2 называется *транспозицией*. Всякую подстановку можно представить также в виде произведения транспозиций (например, $(i_1, i_2, \ldots, i_k) = (i_1, i_2)(i_1, i_3) \ldots (i_1, i_k)$). Это представление не единственно, но всякая четная подстановка (гл. I, § 1, п. 1) разлагается всегда в произведение четного числа транспозиций, а всякая нечетная подстановка — в про-

*) Вместо подстановок чисел $1, 2, \ldots, n$ можно говорить также о подстановках n каких-то элементов $a_1, \ldots, a_n$; если вместо этих элементов писать только индексы, с помощью которых они занумерованы, то получится в точности то же самое, как если бы речь шла о подстановках чисел.

изведение нечетного числа транспозиций. Всякую четную подстановку можно представить также в виде произведения циклов длины 3 (например, $(1,2)(1,3)=(1,2,3)$, $(1,2)(3,4)=$ $=(1,3,2)(2,4,3)$).

Порядок (п. 4) цикла как элемента группы подстановок, в которой он лежит, всегда равен его длине. Порядок подстановки, представленной в виде произведения независимых циклов, равен общему наименьшему кратному длин этих циклов.

Если подстановка a представлена в виде произведения циклов, то, чтобы получить сопряженную с нею (п. 7) подстановку $b^{-1}ab$, нужно числа, составляющие эти циклы, подвергнуть подстановке b.

Пример 3. $a=\begin{pmatrix}1&2&3&4&5&6\\1&6&4&5&3&2\end{pmatrix}=(2,6)(3,4,5)$;

$$b=\begin{pmatrix}1&2&3&4&5&6\\4&3&1&6&5&2\end{pmatrix};\quad b^{-1}ab=(3,2)(1,6,5).$$

Группа подстановок n элементов $\alpha_1,\ldots,\alpha_n$ называется *транзитивной*, если для любых двух из этих элементов α_i, α_j в группе имеется подстановка, переводящая α_i в α_j. Если это условие не выполнено, группа называется *интранзитивной*.

Пример 4. Симметрическая группа n-й степени (п. 1) транзитивна. Знакопеременная группа n-й степени (п. 4) при $n\geqslant 3$ также транзитивна.

Пример 5. Подгруппа симметрической группы n-й степени ($n>2$), состоящая из подстановок (1), (1, 2), интранзитивна.

Транзитивная группа подстановок элементов $\alpha_1,\ldots,\alpha_n$ называется *импримитивной*, если множество элементов $\alpha_1,\ldots,\alpha_n$ можно разбить на непересекающиеся части (*области импримитивности*) так, что каждая из этих частей будет содержать больше одного элемента и каждая подстановка группы будет переводить все элементы любой из этих частей либо снова в элементы этой же части, либо в элементы другой части, одной и той же для всех элементов первой части. Все области импримитивности состоят из одинакового числа элементов. Если множество $\alpha_1,\ldots,\alpha_n$ на области импримитивности разбить нельзя, то группа называется *примитивной*.

Пример 6. Симметрическая группа любой степени и знакопеременная группа n-й степени при $n\geqslant 3$ являются примитивными.

Пример 7. Если число n простое, то всякая транзитивная группа подстановок n элементов примитивна.

Пример 8. Подгруппа группы S_4, состоящая из подстановок (1), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3) (так называемая *четверная группа Клейна*), импримитивна; областями импримитивности для нее являются, например, подмножества 1,2 и 3,4.

В любой группе P подстановок элементов $\alpha_1, \ldots, \alpha_n$ совокупность P_{α_i} всех подстановок, оставляющих какой-нибудь фиксированный элемент α_i на месте, является подгруппой. Если группа P транзитивна, то подгруппы $P_{\alpha_1}, \ldots, P_{\alpha_n}$ сопряжены в ней между собой (п. 7), а индекс каждой из них в группе P (п. 6) равен n. Отсюда следует, что порядок (п. 1) транзитивной группы подстановок n элементов всегда делится на число n (п. 6).

Транзитивная группа P подстановок элементов $\alpha_1, \ldots, \alpha_n$ примитивна тогда и только тогда, когда подгруппы P_{α_i} $(i=1, \ldots, n)$ являются ее максимальными истинными подгруппами. Если это условие не выполнено и существует такая подгруппа Q группы P, что $P_{\alpha_i} \subset Q \subset P$ при некотором i, то областями импримитивности для группы P служат подмножества множества $\alpha_1, \ldots, \alpha_n$, каждое из которых состоит из всех элементов, в которые α_i (i фиксировано) переводится подстановками, принадлежащими одному и тому же правому смежному классу группы P по подгруппе Q (п. 6).

10. Прямые произведения групп. Группа G называется *прямым произведением* (или *прямой суммой*, если групповая операция — сложение) своих подгрупп $H_1, \ldots, H_n$, если:

1) каждая подгруппа H_i $(i=1, \ldots, n)$ является нормальным делителем (п. 7) группы G;

2) группа G порождается (п. 5) подгруппами $H_1, \ldots, H_n$;

3) пересечение (см. введение) всякой подгруппы H_i $(i=1, \ldots, n)$ с подгруппой, порожденной всеми подгруппами H_j, $j \neq i$, содержит только единицу группы.

Требование 3) можно заменить более слабым: пересечение всякой подгруппы H_i $(i=2, \ldots, n)$ с подгруппой, порожденной подгруппами $H_1, \ldots, H_{i-1}$, должно состоять лишь из единицы группы.

Второе определение прямого произведения (равносильное первому). Группа G называется прямым произведением своих подгрупп $H_1, \ldots, H_n$, если:

1) каждый элемент g группы G однозначно записывается в виде произведения

$$g = h_1 \ldots h_n,$$

где $h_i \in H_i$ $(i = 1, \ldots, n)$;

2) элементы любых двух подгрупп H_i и H_j $(i \neq j)$ перестановочны между собой (т. е. $h_i h_j = h_j h_i$, если $h_i \in H_i$, $h_j \in H_j$).

Если группа G является прямым произведением (прямой суммой) своих подгрупп $H_1, H_2, \ldots, H_n$, то пишут: $G = H_1 \times H_2 \times \ldots \times H_n$ (соответственно $G = H_1 \dotplus H_2 \dotplus \ldots \dotplus H_n$). Подгруппы $H_1, H_2, \ldots; H_n$ называются *прямыми множителями* (соответственно *прямыми слагаемыми*), входящими в данное прямое разложение группы G.

Если $G = H_1 \times \ldots \times H_n$ и $G \ni g = h_1 \ldots h_n$ $(h_i \in H_i, i = 1, \ldots, n)$, то элементы $h_1, \ldots, h_n$ называются *компонентами* элемента g соответственно в прямых множителях $H_1, \ldots, H_n$. Для любого i $(i = 1, \ldots, n)$ компонента произведения двух элементов группы G в H_i равна произведению компонент этих элементов в H_i, т. е. если $g = h_1 \ldots h_n$, $g' = h'_1 \ldots h'_n$, то $gg' = (h_1 h'_1) \ldots (h_n h'_n)$.

Пример 1. Четверная группа Клейна V (п. 9) является прямым произведением своей подгруппы H_1, состоящей из подстановок (1) и (1,2) (3,4), и подгруппы H_2, состоящей из подстановок (1) и (1,3) (2,4). Она же является прямым произведением подгруппы H_1 и подгруппы H_3, состоящей из подстановок (1) и (1,4) (2,3), а также прямым произведением подгрупп H_2 и H_3: $V = H_1 \times H_2 = H_1 \times H_3 = H_2 \times H_3$.

Пример 2. Все действительные квадратные матрицы n-го порядка (см. Введение) с положительным определителем (гл. I, § 1, п. 2) образуют относительно умножения матриц (гл. II, § 1, п. 1) группу G, в которой множество H_1 всех матриц с определителем, равным 1, является нормальным делителем. Множество H_2 всех скалярных матриц (гл. II, § 1, п. 3) с положительным числом на главной диагонали также является нормальным делителем группы G. В пересечении H_1 с H_2 ([34], стр. 23—24) лежит только единичная матрица, являющаяся единицей группы G; $G = H_1 \times H_2$.

Если $G = H_1 \times H_2 \times \ldots \times H_n$, причем каждое H_i $(i = 1, \ldots, n)$ само является прямым произведением своих подгрупп, $H_i = H_{i1} \times \ldots \times H_{ik_i}$, то для группы G можно получить новое прямое разложение: $G = H_{11} \times \ldots \times H_{1k_1} \times \ldots \times H_{n1} \times \ldots \times H_{nk_n}$ (*продолжение* прежнего прямого разложения группы G). Еще прямое разложение группы G можно получить, разбивая произвольным образом множество подгрупп $H_1, \ldots, H_n$, входящих в ее исходное прямое разложение, на непересекающиеся части и заменяя подгруппы, входящие в каждую из этих частей, их произведением (т. е. подгруппой, ими порожденной (п. 5)).

Пример 3. Пусть G — группа (по сложению) всех векторов трехмерного пространства, выходящих из начала координат, H_1, H_2, H_3 — ее подгруппы, состоящие соответственно из всех векторов оси x, оси y и оси z, выходящих из начала координат, A и B — подгруппы, состоящие из векторов плоскости xOy и векторов плоскости yOz. Тогда $A = H_1 \dot{+} H_2$, $B = H_2 \dot{+} H_3$, $G = A \dot{+} H_3 = H_1 \dot{+} H_2 \dot{+} H_3 = H_1 \dot{+} B$.

Если $G = H_1 \times H_2$, то фактор-группа G/H_1 (п. 8) изоморфна (п. 2) подгруппе H_2; аналогично фактор-группа G/H_2 изоморфна H_1. Если C — подгруппа группы G, содержащая внутри себя прямой множитель H_1, то $C = H_1 \times D$, где D — пересечение ([34], стр. 23—24) подгрупп C и H_2. Если $G = H_1 \times \ldots \times H_n$ и H_i' — подгруппа группы H_i ($i = 1, \ldots, n$), то подгруппа, порожденная в группе G всеми подгруппами H_i' (п. 5), является прямым произведением этих подгрупп: $\{H_1', \ldots, H_n'\} = H_1' \times \ldots \times H_n'$.

Определение прямого произведения (прямой суммы) можно дать и для случая *бесконечного числа прямых множителей* (прямых слагаемых): группа G называется *прямым произведением (прямой суммой) некоторого множества своих подгрупп* H_α (обозначение: $G = \dot{\prod\limits_\alpha} H_\alpha$ или соответственно $G = \dot{\sum\limits_\alpha} H_\alpha$), если G порождается этими подгруппами и если подгруппа, порожденная в группе G любым конечным числом подгрупп H_α, является их прямым произведением (прямой суммой). Указанные выше свойства прямых произведений верны и для случая бесконечного числа множителей.

Пример 4. Мультипликативная группа всех отличных от нуля рациональных чисел (п. 1) является прямым произведением своей подгруппы, состоящей из чисел $-1, 1$, и бесконечного числа бесконечных циклических подгрупп, каждая из которых порождается некоторым простым числом.

Пример 5. Группа (по сложению) всех многочленов (всех степеней) от одного переменного x с действительными коэффициентами является прямой суммой своих подгрупп, каждая из которых состоит из всех многочленов вида ax^k, где k фиксировано ($k = 0, 1, 2, \ldots$), а коэффициент a пробегает все действительные числа.

Если $H_1, \ldots, H_n$ — какие-то группы, то всегда можно построить группу, являющуюся прямым произведением подгрупп, изоморфных (п. 2) $H_1, \ldots, H_n$. Такая группа получится, если взять множество всех последовательностей вида $(h_1, \ldots, h_n)$, где h_i — элемент группы H_i $(i = 1, \ldots, n)$, и перемножать последовательности почленно: $(h_1, \ldots, h_n) \times \times (h'_1, \ldots, h'_n) = (h_1 h'_1, \ldots, h_n h'_n)$ $(h_i, h'_i \in H_i,\ i = 1, \ldots, n)$.

Если дано бесконечное множество групп H_α, то так же можно построить группу, являющуюся прямым произведением подгрупп, изоморфных H_α. Ее элементами будут системы элементов h_α, взятых по одному из каждой группы H_α, причем только в каждой такой системе все элементы, кроме конечного числа, должны быть единицами соответствующих групп.

11. Абелевы группы. Когда рассматриваются абелевы группы (п. 1), групповая операция обычно называется сложением. Абелева группа G тогда и только тогда является прямой суммой (п. 10) своих подгрупп H_α (α пробегает некоторое множество индексов), когда каждый ее элемент g однозначно записывается в виде суммы конечного числа элементов, взятых из различных групп H_α.

Пример 1. Аддитивная группа (п. 1) всех комплексных чисел является прямой суммой аддитивной группы всех действительных чисел и аддитивной группы всех чисто мнимых чисел ([11], стр. 115).

Всякая абелева группа с конечным числом образующих (п. 5) является прямой суммой конечного числа бесконечных циклических (п. 4) подгрупп и конечного числа циклических подгрупп, порядки которых (п. 4) являются степенями простых чисел (*основная теорема об абелевых группах с конечным числом образующих*).

Пример 2. Циклическая группа 12-го порядка, порожденная элементом a, является прямой суммой циклической подгруппы 4-го порядка, порожденной элементом $3a$, и циклической подгруппы 3-го порядка, порожденной элементом $4a$: $\{a\} = \{3a\} \dotplus \{4a\}$.

Абелева группа, которую нельзя разложить в прямую сумму ее истинных подгрупп, называется *неразложимой*. Бесконечные циклические группы и циклические группы, порядки которых — степени простых чисел, являются неразложимыми.

Если абелева группа с конечным числом образующих двумя способами разложена в прямую сумму бесконечных циклических групп и циклических групп, порядки которых —

степени простых чисел, то в обоих разложениях число бесконечных циклических слагаемых, число конечных циклических слагаемых и совокупность их порядков будут одинаковыми.

Всякая подгруппа A абелевой группы G с конечным числом образующих сама является группой с конечным числом образующих. Если при этом в разложении группы G в прямую сумму бесконечных циклических групп и циклических групп, порядки которых — степени простых чисел, имеется r бесконечных циклических слагаемых ($r \geqslant 0$) и простому числу p соответствуют слагаемые порядков $p^{\alpha_1}, \ldots, p^{\alpha_{k_p}}$, где

$$\alpha_1 \geqslant \ldots \geqslant \alpha_{k_p},$$

(всего k_p слагаемых, $k_p \geqslant 0$), а в разложении группы A в прямую сумму бесконечных циклических групп и циклических групп, порядки которых — степени простых чисел, имеется s бесконечных циклических слагаемых и простому числу p соответствуют слагаемые порядков $p^{\beta_1}, \ldots, p^{\beta_{l_p}}$ ($\beta_1 \geqslant \ldots \geqslant \beta_{l_p}$), то необходимо $s \leqslant r$ и для любого простого числа p $l_p \leqslant k_p$ и $\beta_1 \leqslant \alpha_1, \ldots, \beta_{l_p} \leqslant \alpha_{l_p}$. Об абелевых группах, уже не обязательно обладающих конечными системами образующих, см. [12], гл. VI—VIII.

§ 2. Кольца

1. Определения, примеры (см. [5]). Множество R с двумя определенными в нем алгебраическими операциями (§ 1, п. 1), сложением и умножением, называется *кольцом*, если относительно операции сложения оно является абелевой группой (§ 1, п. 1), а операция умножения связана с операцией сложения законами *дистрибутивности*, т. е. для любых трех элементов $a, b, c \in R$

$$a(b+c) = ab + ac \text{ и } (b+c)a = ba + ca. \tag{4.3}$$

Умножение, определенное в кольце, не обязано быть ни ассоциативным, ни коммутативным. Если умножение, определенное в кольце R, ассоциативно (т. е. для любых трех элементов $a, b, c \in R$ $a(bc) = (ab)c$), то кольцо R называется *ассоциативным кольцом*. Если, кроме того, умножение, определенное в R, коммутативно, то R называется *коммутативным кольцом*. В коммутативном кольце второе из равенств (4.3) является следствием первого.

Если в кольце R для любого элемента a выполнено условие $a^2=0$ и для любых трех элементов a, b, c

$$a(bc)+b(ca)+c(ab)=0$$

(*тождество Якоби*), то R называется *кольцом Ли*.

Пример 1. Все целые числа относительно обычных операций сложения и умножения чисел образуют коммутативное кольцо.

Пример 2. Все рациональные числа, все действительные числа, все комплексные числа относительно обычных операций сложения и умножения образуют коммутативные кольца.

Пример 3. Все многочлены от одного переменного с произвольными числовыми коэффициентами относительно обычных операций сложения и умножения многочленов образуют коммутативное кольцо.

Пример 4. Ассоциативное, но не коммутативное, кольцо образуют все квадратные матрицы n-го порядка с произвольными числовыми элементами (если опять складывать и умножать матрицы обычным образом, см. гл. II, § 1, п. 1).

Пример 5. Множество всех векторов трехмерного пространства, где векторы складываются обычным образом, а произведением двух векторов называется их векторное произведение, является неассоциативным кольцом. Это кольцо есть кольцо Ли.

Абелева группа, которая получится, если в кольце рассматривать только одну операцию сложения, называется *аддитивной группой* кольца. Нулевой элемент этой группы называется *нулем кольца*. Произведение любого элемента a кольца на нуль равно нулю кольца: $a\cdot 0=0$.

Если для элементов a, b некоторого кольца $ab=0$, но $a\neq 0$ и $b\neq 0$, то a и b называются *делителями нуля* (a — левый делитель нуля, b — правый). Если в кольце R делителей нуля нет, то R называется *кольцом без делителей нуля*. Коммутативное кольцо без делителей нуля называется *областью целостности*.

Пример 6. Всякое кольцо, в котором элементы — числа, а операции — обычное сложение и умножение чисел, является областью целостности.

Пример 7. Все функции, определенные и непрерывные ([34], стр. 48) на отрезке $[-1,1]$, относительно обычных операций сложения и умножения функций образуют кольцо с делителями нуля (например, произведение функций

$$f_1(x)=\begin{cases}0 \text{ при } -1\leqslant x\leqslant 0,\\ x \text{ при } 0<x\leqslant 1;\end{cases}\qquad f_2(x)=\begin{cases}x \text{ при } -1\leqslant x\leqslant 0,\\ 0 \text{ при } 0<x\leqslant 1,\end{cases}$$

ни одна из которых не равна нулю кольца, является нулем).

Если для элементов a, b_1, b_2 кольца выполнено равенство $ab_1 = ab_2$ (или $b_1a = b_2a$), причем $a \neq 0$ не есть левый (соответственно — правый) делитель нуля, то $b_1 = b_2$, т. е. на отличные от нуля элементы, не являющиеся делителями нуля, равенства можно сокращать. Производить сокращение на элемент, являющийся делителем нуля, нельзя.

Пример 8. В кольце всех квадратных матриц 2-го порядка (пример 4) для матриц

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B_1 = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}, B_2 = \begin{pmatrix} 1 & 2 \\ 4 & 0 \end{pmatrix}$$

справедливо равенство $AB_1 = AB_2$, хотя $B_1 \neq B_2$. Здесь A — левый делитель нуля: например, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Элемент e кольца R называется *единицей* этого кольца, если для любого элемента $a \in R$ $ae = ea = a$. Единицы в кольце может и не быть. Если в кольце R единица есть, то R называется *кольцом с единицей*.

Пример 9. Кольцо всех целых чисел есть кольцо с единицей.

Пример 10. Все четные числа образуют кольцо без единицы.

В кольце с единицей e для элемента $a \neq 0$ может существовать *обратный* ему элемент a^{-1} со свойством $aa^{-1} = a^{-1}a = e$ (но может такого элемента и не быть). Элементы кольца с единицей, для которых в этом кольце обратный элемент существует, называются *делителями единицы*.

Пример 11. В кольце всех квадратных матриц n-го порядка единицей является единичная матрица (гл. II, § 1, п. 2). Обратный элемент существует для всякой невырожденной матрицы (гл. II, § 1, п. 2); для вырожденных матриц обратных им элементов не существует.

2. Изоморфизм. Гомоморфизм. Кольца R и Q называются *изоморфными*, если между их элементами можно установить такое взаимно однозначное соответствие (§ 1, п. 2), что если элементам $a_1, a_2 \in R$ соответствуют элементы $b_1, b_2 \in Q$, то элементу $a_1 + a_2$ соответствует элемент $b_1 + b_2$ и элементу a_1a_2 — элемент b_1b_2.

Пример 1. Кольцо всех квадратных матриц n-го порядка с действительными элементами (п. 1, пример 4) изоморфно кольцу всех линейных преобразований, действующих в действительном n-мерном линейном векторном пространстве (сложение и умножение преобразований определяется обычно: см. [8], [11], [14], [16], [32]).

Говорят, что кольцо R *гомоморфно отображено* в кольцо Q, если каждому элементу кольца R поставлен в соответствие однозначно определенный элемент кольца Q, причем если элементам a_1, a_2 кольца R соответствуют элементы $b_1, b_2 \in Q$, то элементу $a_1 + a_2$ соответствует $b_1 + b_2$, элементу $a_1 a_2$ соответствует $b_1 b_2$. Если при этом на каждый элемент кольца Q отображается по крайней мере один элемент кольца R, то говорят о гомоморфном отображении кольца R на кольцо Q; гомоморфизм в этом случае называется *эпиморфизмом*.

Пример 2. Если каждому многочлену от одного переменного с любыми комплексными коэффициентами поставить в соответствие свободный член этого многочлена, то получится гомоморфное отображение кольца всех многочленов с комплексными коэффициентами на кольцо (на самом деле даже поле) всех комплексных чисел.

Если кольцо R гомоморфно (в частности, изоморфно) отображено на кольцо Q, причем кольцо R ассоциативно (п. 1), то и кольцо Q ассоциативно; если кольцо R коммутативно, то и Q коммутативно; если R — кольцо с единицей (п. 1), то и Q — кольцо с единицей, причем единицей кольца Q является образ единицы кольца R. Если кольцо R может быть гомоморфно (или изоморфно) отображено на некоторое множество M с двумя алгебраическими операциями (§ 1, п. 1), то M само является кольцом относительно этих двух операций.

Если кольцо R гомоморфно отображено в кольцо Q, то множество всех элементов кольца R, переходящих при этом в нуль кольца Q, называется *ядром* этого гомоморфизма.

3. Подкольца. Идеалы. Подгруппа A (§ 1, п. 4) аддитивной группы кольца R (п. 1) называется *подкольцом* этого кольца, если она вместе с каждыми двумя элементами a_1, a_2 содержит также их произведение $a_1 a_2$. Подкольцо A кольца R называется *левым* (соответственно *правым*) *идеалом* этого кольца, если оно вместе с каждым элементом a содержит также все элементы вида ra (вида ar), где r — произвольный элемент кольца R. В коммутативных кольцах понятия левого и правого идеала совпадают. Если подкольцо A произвольного кольца одновременно является и левым, и правым идеалом, то оно называется *двусторонним идеалом* этого кольца.

Пример 1. В кольце всех целых чисел числа, кратные некоторому фиксированному числу n, составляют двусторонний идеал.

Пример 2. В кольце всех квадратных матриц n-го порядка множество всех матриц, у которых последний столбец состоит из нулей, является левым идеалом, а множество всех матриц, у которых последняя строка состоит из нулей, является правым идеалом.

В любом кольце элемент 0 и само кольцо являются двусторонними идеалами. Если других двусторонних идеалов в кольце нет, оно называется *простым кольцом.* Примером простого кольца служит кольцо всех квадратных матриц n-го порядка с любыми комплексными (или с любыми действительными) элементами (п. 1, пример 4). Также всякое поле (§ 3) является простым кольцом.

Если в кольце R дано некоторое множество (см. введение) элементов N, то наименьший (левый, правый или двусторонний) идеал кольца R, содержащий все элементы из N, называется (соответственно левым, правым или двусторонним) *идеалом, порожденным множеством* N; этот идеал является пересечением ([34], стр. 23—24) всех (соответственно левых, правых или двусторонних) идеалов кольца R, содержащих множество N. Идеал, порожденный одним элементом a, называется *главным идеалом* (обозначение: $(a)_l$, $(a)_r$ или (a), в зависимости от того, идеал левый, правый или двусторонний). Если R — коммутативное кольцо, то главный идеал (a), порожденный элементом a этого кольца, состоит из всех элементов вида $ra+na$, где $r \in R$, n — целое число. Если R — коммутативное кольцо с единицей, то (a) состоит просто из всех элементов вида ra, где $r \in R$.

Если R — кольцо с единицей e, то само оно является главным идеалом, порожденным элементом e. Область целостности (п. 1) с единицей, в которой каждый идеал — главный, называется *кольцом главных идеалов.* Например, кольцами главных идеалов являются: кольцо всех целых чисел, кольцо всех многочленов от одного переменного с коэффициентами из любого поля (гл. III, § 1).

Если R — кольцо главных идеалов и ε — делитель единицы этого кольца (п. 1), то любой элемент $a \in R$ можно представить в виде $a=a\ \varepsilon^{-1}\varepsilon$. Такое разложение элемента a называется *тривиальным.* Если для элемента $p \neq 0$ кольца R нетривиальных разложений не существует (т. е. из равенства $p=cd$ всегда следует, что c или d — делитель единицы), то p называется *неразложимым* или *простым элементом* (например, простое число в кольце целых чисел, многочлен, неприводимый над данным полем (гл. III, § 3, п. 7), в кольце всех многочленов с коэффициентами из этого поля).

В кольце главных идеалов всякий отличный от нуля элемент является произведением простых элементов. Если при этом $a = p_1 p_2 \ldots p_s = q_1 q_2 \ldots q_t$ — два разложения элемента a на простые множители (не являющиеся делителями единицы), то $s = t$ и p_i совпадают с q_j с точностью до порядка нумерации и до множителей — делителей единицы.

4. Прямые суммы колец. Если аддитивная группа произвольного кольца R (п. 1) разложена в прямую сумму своих подгрупп A_α (§ 1, п. 10), где α пробегает некоторое множество индексов, причем каждое A_α является двусторонним идеалом (п. 3) кольца R, то говорят, что кольцо R *разложено в* (двустороннюю) *прямую сумму*. Если дано какое-то множество колец R_α, то всегда можно построить кольцо, являющееся двусторонней прямой суммой идеалов, изоморфных R_α. Для этого нужно построить абелеву группу, являющуюся прямой суммой подгрупп $\overline{R}_\alpha$, изоморфных аддитивным группам колец R_α (§ 1, п. 10), и определить на ней умножение по формуле

$$(a_{\alpha_1} + \ldots + a_{\alpha_n})(a'_{\alpha_1} + \ldots + a'_{\alpha_n}) = a_{\alpha_1} a'_{\alpha_1} + \ldots + a_{\alpha_n} a'_{\alpha_n}$$

$(a_{\alpha_i}, a'_{\alpha_i} \in \overline{R}_{\alpha_i}, i = 1, \ldots, n$, причем некоторые из элементов a_{α_i} или a'_{α_i} могут быть и нулями); произведение $a_{\alpha_i} a'_{\alpha_i}$ находим, пользуясь правилом умножения в кольце R_{α_i}.

5. Фактор-кольца. Пусть R — произвольное кольцо, A — его двусторонний идеал. Тогда можно построить фактор-группу (§ 1, п. 8) аддитивной группы кольца R (п. 1) по подгруппе, состоящей из всех элементов идеала A. Если в множестве всех смежных классов $r + A$ ($r \in R$), являющихся элементами этой фактор-группы, ввести еще вторую операцию, умножение, по правилу $(r + A)(r' + A) = rr' + A$ $(r, r' \in R)$, то получится кольцо, называемое *фактор-кольцом кольца R по двустороннему идеалу A* (обозначение: R/A).

Пример. Фактор-кольцо кольца всех целых чисел по идеалу A, состоящему из всех четных чисел, содержит два элемента: смежный класс $a_0 = 0 + A = A$ и смежный класс $a_1 = 1 + A$. Эти элементы складываются по правилу $a_0 + a_0 = a_0$, $a_0 + a_1 = a_1 + a_0 = a_1$ $a_1 + a_1 = a_0$ и умножаются по правилу $a_0 a_0 = a_0$, $a_0 a_1 = a_1 a_0 = a_0$, $a_1 a_1 = a_1$.

Если каждому элементу r кольца R поставить в соответствие определяемый им смежный класс $r+A$ по двустороннему идеалу A, то это даст гомоморфное отображение (п. 2) кольца R на фактор-кольцо R/A, ядром которого будет являться идеал A (п. 2). Если кольцо R гомоморфно отображено на какое-то кольцо Q и A — ядро этого гомоморфизма, то A необходимо является двусторонним идеалом кольца R, а фактор-кольцо R/A изоморфно Q (*теорема о гомоморфизмах колец*).

§ 3. Поля. Тела

1. Поля (см. [5]). *Полем* называется коммутативное кольцо (§ 2, п. 1), состоящее не только из нуля, в котором для любого элемента $a \neq 0$ и любого элемента b существует ровно один такой элемент x, что $ax = b$. Элемент x называется *частным* от деления элемента b на элемент a (обозначение: $x = \frac{b}{a}$).

Примерами полей служат: кольцо всех рациональных чисел, кольцо всех действительных чисел, кольцо всех комплексных чисел. Все комплексные числа, являющиеся корнями многочленов с рациональными коэффициентами, также образуют поле, называемое *полем алгебраических чисел*. Кольцо всех целых чисел полем не является. Но всякое его фактор-кольцо (§ 2, п. 5) по идеалу, состоящему из всех чисел, кратных некоторому простому числу p, является полем (*поле классов вычетов по модулю* p). Еще пример поля — все дробно-рациональные функции $\frac{f(x)}{g(x)}$, где $f(x)$ и $g(x)$ — многочлены с действительными коэффициентами, причем $g(x) \neq 0$.

Всякое поле обладает единицей (§ 2, п. 1). Для любого отличного от нуля элемента поля существует обратный ему элемент (§ 2, п. 1). Множество всех отличных от нуля элементов поля образует относительно умножения, определенного в поле, абелеву группу (§ 1, п. 1) (*мультипликативную группу поля*). Никакое поле не содержит делителей нуля (§ 2, п. 1). Единственными идеалами поля (§ 2, п. 3) являются нулевой идеал и само поле.

Множество с двумя алгебраическими операциями (§ 1, п. 1), изоморфное полю (§ 2, п. 2), само является полем. Всякое гомоморфное отображение одного поля в другое (§ 2, п. 2) является или изоморфизмом, или отображением, переводящим все элементы поля в нуль.

Если некоторое целое положительное кратное единичного элемента e поля P $ne = e + e + \ldots + e$ (n слагаемых) равно нулю, то наименьшее целое положительное число p со свойством $pe = 0$ называется *характеристикой поля* P; p всегда является простым числом. Если никакое целое положительное кратное единичного элемента поля P нулю не равно, то P называется полем *характеристики нуль*. Пример поля характеристики p — поле классов вычетов по модулю p, пример поля характеристики нуль — любое числовое поле (например, поле всех действительных чисел).

Множество P' элементов поля P называется *подполем* этого поля, если оно само является полем по отношению к тем операциям, которые определены в P; P' тогда и только тогда является подполем поля P, когда оно вместе с любыми двумя элементами a, b содержит также $a + b$, ab, $a - b$ *) и $\frac{a}{b}$ (если $b \neq 0$). Если P' — подполе поля P, то P называется *расширением* поля P'. Поле, не имеющее никаких подполей, кроме него самого, называется *простым полем*. Примеры простых полей — поле классов вычетов по модулю p, поле всех рациональных чисел.

Всякое поле характеристики p содержит подполе, изоморфное (§ 2, п. 2) полю классов вычетов по модулю p, а всякое поле характеристики нуль содержит подполе, изоморфное полю всех рациональных чисел.

Если в поле P даны подполе P' и множество элементов N, то наименьшее подполе P'' поля P, содержащее P' и N, называется полем, полученным *присоединением* к полю P' множества N (обозначение: $P'' = P'(N)$). Поле $P'(N)$ состоит из всех элементов, получающихся из элементов поля P' и множества N путем сложения, вычитания, умножения и деления, и является пересечением (см. введение) всех подполей поля P, содержащих P' и N. Если множество N состоит из одного элемента α, то P'' называется *простым расширением* поля P'. Если при этом α является корнем некоторого многочлена $f(x)$ с коэффициентами из поля P' (гл. III, § 1), то $P'' = P'(\alpha)$

*) $a - b$ — элемент, являющийся решением уравнения $b + x = a$.

называется *простым алгебраическим расширением* поля P', а элемент α называется *алгебраическим относительно* P'; если же не существует многочлена $f(x)$ с коэффициентами из P', корнем которого был бы элемент α, то $P'' = P'(\alpha)$ называется *простым трансцендентным расширением* поля P', а элемент α называется *трансцендентным относительно* P'. Если простое алгебраическое расширение $P'(\alpha)$ поля P' получено присоединением к нему элемента α, являющегося корнем многочлена $f(x)$ степени n с коэффициентами из P', неприводимого над полем P' (гл. III, § 3, п. 7), то каждый элемент поля $P'(\alpha)$ однозначно записывается в виде $a_0 + a_1\alpha + \ldots + a_{n-2}\alpha^{n-2} + a_{n-1}\alpha^{n-1}$, где $a_0, \ldots, a_{n-1}$ — элементы поля P'. Если $P'(\eta)$ — простое трансцендентное расширение поля P', то каждый его элемент имеет вид $\frac{f(\eta)}{g(\eta)}$, где $f(\eta)$ и $g(\eta) \neq 0$ — какие-то многочлены от η с коэффициентами из поля P'.

Пример 1. Простое алгебраическое расширение поля рациональных чисел, полученное присоединением к нему числа $\alpha = \sqrt[3]{2}$, состоит из всех чисел вида $a + b\sqrt[3]{2} + c\sqrt[3]{4}$, где a, b, c — рациональные числа.

Пример 2. Поле всех комплексных чисел является простым алгебраическим расширением поля всех действительных чисел, полученным из него присоединением корня i многочлена $x^2 + 1$.

Пример 3. Поле всех дробно-рациональных функций вида $\frac{f(x)}{g(x)}$, где $f(x)$ и $g(x)$—многочлены с рациональными коэффициентами ($g(x) \neq 0$), является простым трансцендентным расширением поля рациональных чисел.

Если $f(x)$ — многочлен с коэффициентами из поля P (гл. III, § 1), то всегда существует такое расширение $\overline{P}$ поля P, что $f(x)$ является произведением многочленов первой степени с коэффициентами из $\overline{P}$, $f(x) = a_0(x - \alpha_1)\ldots(x - \alpha_n)$, где $a_0 \in P, \alpha_i \in \overline{P}, i = 1, \ldots, n$. Подполе поля $\overline{P}$, полученное присоединением к полю P элементов $\alpha_1, \ldots, \alpha_n$, называется *полем разложения* многочлена $f(x)$ ($\alpha_1, \ldots, \alpha_n$ — корни $f(x)$). Если все корни любого многочлена с коэффициентами из поля P лежат уже в самом поле P, то поле P называется *алгебраически замкнутым.* Примерами алгебраически замкнутых полей являются: поле всех комплексных чисел, поле алгебраических чисел.

Если каждый элемент некоторого расширения $\tilde{P}$ поля P является алгебраическим относительно P, то поле $\tilde{P}$ называется *алгебраическим расширением* поля P. Всякое простое алгебраическое расширение поля P является алгебраическим расширением. Для любого поля существует такое алгебраическое расширение, которое является алгебраически замкнутым полем, т. е. уже дальнейшего алгебраического расширения не допускает (*теорема Штейница*).

2. Тела. *Телом* называется ассоциативное кольцо (§ 2, п. 1), состоящее не только из нуля, в котором для любого элемента $a \neq 0$ и любого элемента b существуют такой элемент x и такой элемент y, что $ax = b$, $ya = b$. Отличие тела от поля состоит в том, что умножение в теле не обязано быть коммутативным. Пример тела, не являющегося полем, — тело кватернионов (§ 4).

Иногда в определение тела требование ассоциативности кольца не включают.

§ 4. Алгебры

Кольцо A (§ 2, п. 1) называется *алгеброй* над полем P (§ 3), если его аддитивная группа (§ 2, п. 1) есть векторное пространство ([34], стр. 58) над полем P и если умножение в A связано с умножением на элементы из P формулой $\alpha(ab) = (\alpha a)b = a(\alpha b)$ $(a, b \in A, \alpha \in P)$ ([34], стр. 23).

Если векторное пространство, которое служит аддитивной группой алгебры A, является n-мерным ([34], стр. 58), то число n называется *рангом* алгебры A. Алгебры конечного ранга называются также *гиперкомплексными системами*. Если алгебра A является кольцом Ли (§ 2, п 1), то она называется *алгеброй Ли* ([19], стр. 379). Если какая-то алгебра является не только кольцом, но даже телом, она называется *алгеброй с делением*.

Пример 1. Множество всех векторов трехмерного пространства, в котором векторы складываются и умножаются на числа обычным образом, а произведением двух векторов является их векторное произведение, есть алгебра Ли над полем действительных чисел.

Пример 2. Множество всех квадратных матриц n-го порядка с комплексными элементами, в котором обычным образом определены операции сложения и умножения матриц и операция умножения матрицы на комплексное число (гл. II, §1, п. 1), является алгеброй ранга n^2 над полем комплексных чисел.

Пример 3. Все комплексные числа относительно обычных операций сложения и умножения комплексных чисел и обычной операции умножения комплексных чисел на действительные числа образуют алгебру с делением ранга 2 над полем действительных чисел.

Пример 4. Всякое поле является алгеброй ранга 1 над самим собой.

Если векторное пространство над полем P, служащее аддитивной группой алгебры A, имеет базис $e_1, \ldots, e_n$ ([34], стр. 56), то достаточно знать, чему равны произведения элементов этого базиса друг на друга, чтобы знать, чему равно произведение любого элемента алгебры A на любой другой ее элемент (если $a, b \in A$, причем $a = \alpha_1 e_1 + \ldots + \alpha_n e_n$, $b = \beta_1 e_1 + \ldots + \beta_n e_n$ $(\alpha_i, \beta_i \in P, i = 1, \ldots, n)$, то $ab = \alpha_1\beta_1 e_1^2 + \alpha_1\beta_2 e_1 e_2 + \ldots + \alpha_1\beta_n e_1 e_n + \alpha_2\beta_1 e_2 e_1 + \ldots + \alpha_n\beta_n e_n^2$). Отсюда — задание алгебры конечного ранга *таблицей умножения*, где для любых двух элементов e_i, e_j базиса векторного пространства, служащего аддитивной группой алгебры, указывается линейная комбинация элементов этого базиса, равная произведению $e_i e_j$.

Пример 5. Если в алгебре всех комплексных чисел над полем действительных чисел в качестве базиса взять элементы $1, i$, то таблица умножения будет выглядеть так:

	1	i
1	1	i
i	i	-1.

Пример 6. Если в алгебре всех квадратных матриц n-го порядка над полем комплексных чисел в качестве базиса взять матрицы E_{ij} $(i, j = 1, \ldots, n)$, где на пересечении i-й строки и j-го столбца стоит единица, а остальные элементы — нули, то произведение в этой алгебре будет задаваться таблицей умножения

$$E_{ij}E_{kl} = \begin{cases} E_{il} & \text{при } j = k, \\ 0 & \text{при } j \neq k. \end{cases}$$

Если умножение элементов базиса ассоциативно или коммутативно, то и вообще умножение, определенное в алгебре, соответственно ассоциативно или коммутативно.

Алгебра, аддитивной группой которой служит 4-мерное векторное пространство над полем действительных чисел с базисом 1, i, j, k и в которой умножение задано таблицей умножения

	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1,

называется *алгеброй кватернионов*. Это — некоммутативная, но ассоциативная алгебра с делением. Ее элементы называются *кватернионами*. Они однозначно записываются в виде $z = \alpha \cdot 1 + \beta i + \gamma j + \delta k$, где α, β, γ, δ — действительные числа. Кватернион $\bar{z} = \alpha \cdot 1 - \beta i - \gamma j - \delta k$ называется *сопряженным* с z; $z\bar{z} = \alpha^2 + \beta^2 + \gamma^2 + \delta^2 = |z|^2$ (обобщение квадрата модуля комплексного числа). Обратным для кватерниона $z \neq 0$ является кватернион $z^{-1} = \frac{\bar{z}}{|z|^2}$.

Поле действительных чисел, поле комплексных чисел и алгебра кватернионов являются единственными ассоциативными алгебрами с делением конечного ранга над полем действительных чисел (*теорема Фробениуса*) ([16], стр. 368).

Две алгебры A и A' над одним и тем же полем P называются *изоморфными*, если между их элементами можно установить взаимно однозначное соответствие, которое дает изоморфизм колец A и A' (§ 2, п. 2), причем если элементу $a \in A$ соответствует элемент $a' \in A'$, то для любого $\alpha \in P$ элементу αa соответствует $\alpha a'$. Говорят, что алгебра A над полем P *гомоморфно* отображена в алгебру C над тем же полем, если кольцо A гомоморфно отображено в кольце C (§ 2, п. 2), причем если элементу $a \in A$ соответствует элемент $c \in C$, то при любом $\alpha \in P$ элементу αa соответствует элемент αc.

Пример 7. Алгебра всех комплексных чисел над полем действительных чисел изоморфна алгебре всех квадратных матриц второго порядка вида $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$ с действительными элементами α, β.

Подкольцо B (§ 2, п. 3) алгебры A над полем P называется *подалгеброй* этой алгебры, если его аддитивная группа (§ 2, п. 1) является подпространством ([34], стр. 58) пространства, служащего аддитивной группой алгебры A. Если подалгебра B является левым, правым или двусторонним идеалом кольца A (§ 2, п. 3), то она называется соответственно *левым, правым* или *двусторонним идеалом алгебры* A. Множество B элементов алгебры A тогда и только тогда является левым идеалом этой алгебры, когда оно вместе с каждыми двумя элементами b_1 и b_2 содержит их разность $b_1 - b_2$ и вместе с каждым элементом b содержит также произведение ab любого элемента $a \in A$ на элемент b и произведение αb любого элемента α поля P на элемент b.

Пример 8. В алгебре кватернионов кватернионы вида $\alpha \cdot 1 + \beta i$ образуют подалгебру (изоморфную алгебре всех комплексных чисел над полем действительных чисел).

Пример 9. В алгебре всех квадратных матриц второго порядка над полем комплексных чисел матрицы вида $\begin{pmatrix} 0 & \alpha \\ 0 & \beta \end{pmatrix}$ образуют левый идеал.

Пример 10. Во всякой алгебре с единицей e (§ 2, п. 1) над полем P (состоящей не только из нуля) совокупность всех элементов вида αe, где $\alpha \in P$, является подалгеброй, изоморфной полю P.

Если B — двусторонний идеал алгебры A над полем P, то можно взять фактор-кольцо A/B (§ 2, п. 5). Если в этом фактор-кольце ввести операцию умножения на элементы из поля P по правилу $\alpha(a + B) = \alpha a + B$ ($\alpha \in P$, $a \in A$), то получится алгебра над полем P, называемая *фактор-алгеброй* алгебры A по идеалу B.

§ 5. Структуры

1. Частично упорядоченные множества. Множество S называется *частично упорядоченным*, если для некоторых его элементов a, b определено соотношение $a \leqslant b$ (словами: «a меньше или равно b», или «a предшествует b», или «a содержится в b»), причем выполнены условия:

1) всегда $a \leqslant a$;

2) из $a \leqslant b$ и $b \leqslant a$ следует $a = b$ (т. е. элементы a и b совпадают);

3) из $a \leqslant b$, $b \leqslant c$ следует $a \leqslant c$ (транзитивность).

Если $a \leqslant b$, но $a \neq b$, то пишут $a < b$. Вместо $a \leqslant b$ пишут также $b \geqslant a$ и говорят: «b больше или равно a», или «b следует за a», или «b содержит a»; вместо $a < b$ пишут также $b > a$. Если $b > a$ и в данном частично упорядоченном множестве нет элемента x со свойством $b > x > a$, то говорят, что b *покрывает* a.

Пример 1. Множество S состоит из всех подмножеств A, $B, \ldots$ некоторого множества M (включая пустое множество), причем считается, что $A \leqslant B$, если каждый элемент подмножества A лежит также и в B (пустое множество считается содержащимся в любом другом).

Пример 2. Множество S состоит из всех подгрупп A, $B, \ldots$, некоторой группы G (§ 1, п. 4), а соотношение $A \leqslant B$ означает, что A содержится в B.

Пример 3. Множество S состоит из всех натуральных чисел, а соотношение $a \leqslant b$ означает, что число a является делителем числа b.

Пример 4. Множество S состоит из всех действительных однозначных функций ([34], стр. 48), определенных на отрезке $[0, 1]$, причем считается, что $f(x) \leqslant g(x)$, если в каждой точке отрезка $[0, 1]$ $f(x)$ принимает значение, меньшее или равное значению $g(x)$ в этой точке.

Частично упорядоченное множество называется *упорядоченным* (или *линейно упорядоченным*), если для всякой пары его элементов a, b имеет место или $a \leqslant b$, или $b \leqslant a$. Например, множество всех действительных чисел, где $a \leqslant b$ означает, что число a меньше или равно числу b.

Если множество S частично упорядочено (упорядочено) при помощи некоторого отношения $\leqslant$, то при помощи этого же отношения $\leqslant$ оказывается частично упорядоченным (упорядоченным) и любое его подмножество.

Если между элементами двух частично упорядоченных множеств S и S' можно установить такое взаимно однозначное соответствие (§ 1, п. 2), что для любых a, $b \in S$ и соответствующих им элементов a', $b' \in S'$ из соотношения $a \leqslant b$ всегда следует соотношение $a' \leqslant b'$ и наоборот, то множества S и S' называются *изоморфными*. Изоморфизм частично упорядоченного множества с самим собой называется *автоморфизмом*.

Пример 5. Если S — множество всех действительных чисел и $a \leqslant b$ означает, что число a меньше или равно числу b, то автоморфизмом множества S будет, например, соответствие $a \to a+2$.

Если в частично упорядоченном множестве S заменить заданную упорядоченность на противоположную, т. е. считать, что $b \leqslant a$, тогда и только тогда, когда при прежней упорядоченности было $a \leqslant b$, то получится опять частично упорядоченное множество, называемое *двойственным* множеству S.

Пример 6. Множество S^*, состоящее из всех подмножеств $A, B, \ldots$ некоторого множества M, где считается, что $A \leqslant B$, если каждый элемент подмножества B лежит также в A, двойственно множеству S примера 1.

2. Основные определения ([2]). Частично упорядоченное множество S (п. 1) называется *структурой*, если для любых двух его элементов a, b существуют элементы $c, d \in S$ со свойствами: 1) $c \leqslant a$, $c \leqslant b$; 2) если $c' \leqslant a$ и $c' \leqslant b$ ($c' \in S$), то необходимо $c' \leqslant c$; 3) $d \geqslant a$, $d \geqslant b$; 4) если $d' \geqslant a$ и $d' \geqslant b$ ($d' \in S$), то необходимо $d' \geqslant d$.

Элемент c называется *пересечением* элементов a и b (обозначение: $c = a \cap b$), элемент d называется их *объединением* (обозначение: $d = a \cup b$).

Пример 1. Частично упорядоченные множества примеров 1—4 предыдущего пункта являются структурами, причем в первом из них пересечениями и объединениями подмножеств множества M являются их пересечения и объединения в теоретико-множественном смысле ([34], стр. 23—24), во втором примере пересечением двух подгрупп является их пересечение в теоретико-множественном смысле, а объединением является подгруппа, порожденная (§ 1, п. 5) двумя данными подгруппами, в третьем примере пересечением является общий наибольший делитель, а объединением — общее наименьшее кратное двух заданных чисел, в четвертом примере пересечением (объединением) функций $f(x)$ и $g(x)$ является функция, принимающая в каждой точке отрезка $[0, 1]$ наименьшее (наибольшее) из значений, принимаемых в этой точке функциями $f(x)$ и $g(x)$.

В любой структуре S для всякого конечного множества ее элементов $a_1, \ldots, a_n$ существуют такие элементы c, d, что $c \leqslant a_i \leqslant d$ $(i = 1, \ldots, n)$ и если $c' \leqslant a_i$ $(i = 1, \ldots, n)$, $c' \in S$ (или $d' \geqslant a_i$, $i = 1, \ldots, n$, $d' \in S$), то $c' \leqslant c$ (соответственно $d' \geqslant d$); c называется *пересечением*, d — *объединением* элементов $a_1, \ldots, a_n$. Если пересечение и объединение существуют для каждого множества элементов структуры S (может быть, и бесконечного), то структура S называется *полной*.

Всякая конечная структура является полной. Примерами полных структур являются также структура всех подгрупп данной группы (пример 1) структура всех подколец (§ 2, п. 3) данного кольца (где для подколец A, B считается, что $A \leqslant B$, если A содержится в B).

Если в структуре S имеется такой элемент a_0, что для всякого $b \in S$ имеет место $a_0 \leqslant b$, то a_0 называется *нулем* структуры S и обозначается через 0. Если в S имеется такой элемент a_1, что для любого $b \in S$ имеет место $b \leqslant a_1$, то a_1 называется *единицей* структуры и обозначается через 1. В любой структуре, состоящей из конечного числа элементов, нулем является пересечение, а единицей — объединение всех элементов структуры. В структуре всех подгрупп данной группы нулем является единичная подгруппа (§ 1, п. 4), единицей — вся группа. В структуре примера 4 предыдущего пункта ни нуля, ни единицы нет.

Подмножество (см. введение) структуры, содержащее вместе с любыми двумя элементами их пересечение и их объединение, называется *подструктурой* данной структуры. Очевидно, всякая подструктура сама является структурой относительно частичной упорядоченности, заданной во всей структуре. Но подмножество структуры может быть структурой относительно этой же частичной упорядоченности и в том случае, когда оно подструктурой не является.

Пример 2. Структура всех нормальных делителей (§ 1, п. 7) произвольной группы G (где для нормальных делителей H, K считается, что $H \leqslant K$, если H содержится в K) является подструктурой структуры всех ее подгрупп.

Пример 3. Структура всех подгрупп группы (пример 1) не является подструктурой структуры всех подмножеств множества всех элементов этой группы.

Две структуры называются *изоморфными*, если они изоморфны как частично упорядоченные множества (п. 1). Всякое частично упорядоченное множество, изоморфное структуре, само является структурой.

Если структура состоит из конечного числа элементов, то для нее можно построить *диаграмму*: элементы структуры изображаются точками плоскости так, что если $b > a$, то точка, изображающая b, лежит выше, чем точка, изображающая a; точки, соответствующие элементам, из которых один покрывает другой (п. 1), соединяются отрезками.

Пример 4. Пусть группа G есть прямое произведение (§ 1, п. 10) двух своих циклических подгрупп 2-го порядка (§ 1, п. 4): $G = \{a\} \times \{b\}$. Тогда структура подгрупп группы G имеет диаграмму, изображенную на рис. 1, где $A = \{a\}$, $B = \{b\}$, $C = \{ab\}$, E — единичная подгруппа (§ 1, п. 4).

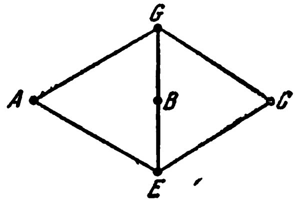


Рис. 1.

Второе определение структуры. Множество S называется структурой, если в нем определены две алгебраические операции (§ 1, п. 1), умножение и сложение, ставящие в соответствие всякой паре элементов $a, b \in S$ их произведение ab и их сумму $a + b$, причем эти операции коммутативны и ассоциативны (т. е. $ab = ba$, $a + b = b + a$, $a(bc) = (ab)c$, $a + (b + c) = (a + b) + c$), удовлетворяют при всяком $a \in S$ условиям $aa = a$, $a + a = a$ и связаны между собой условием: если $ab = a$, то $a + b = b$, и обратно.

Это определение структуры эквивалентно приведенному выше. Для доказательства нужно пересечение и объединение элементов a, b (в смысле первого определения) считать соответственно их произведением и суммой и, наоборот, нужно считать, что $a \leqslant b$, если $ab = a$ (в смысле второго определения).

3. Дистрибутивные и дедекиндовы структуры. Если в структуре S сложение и умножение связаны законом дистрибутивности, т. е. для любых трех элементов $a, b, c \in S$

$$(a + b)c = ac + bc, \tag{4.4}$$

то структура S называется *дистрибутивной*.

Закон дистрибутивности в структуре может быть выражен также в форме $(a + c)(b + c) = ab + c$.

Структура S тогда и только тогда является дистрибутивной, когда в ней из равенств $ac = bc$, $a + c = b + c$ (a, b, $c \in S$) вытекает равенство $a = b$.

Пример 1. Всякая структура, являющаяся линейно упорядоченным множеством (п. 1), дистрибутивна (например, структура всех целых чисел, где $a \leqslant b$ означает, что число a меньше или равно числу b). Множества примеров 1, 3, 4 п. 1 также являются дистрибутивными структурами. Структура всех подгрупп данной группы G дистрибутивна тогда и только тогда, когда G есть или циклическая группа, или объединение возрастающей последовательности циклических групп (§ 1, пп. 4, 5).

Всякая подструктура (п. 2) дистрибутивной структуры дистрибутивна. Структура, двойственная (п. 1) дистрибутивной, также дистрибутивна.

Структура тогда и только тогда не является дистрибутивной, когда она содержит подструктуру (п. 2), которая имеет вид одной из двух диаграмм, изображенных на рис. 2 и 3.

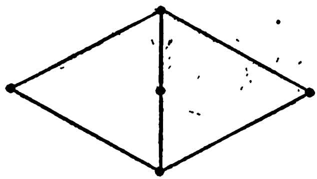


Рис. 2.

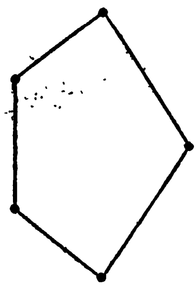


Рис. 3.

Если в структуре S равенство (4.4) выполнено при условии, что $a \leqslant c$ (а без этого условия может и не выполняться), то структура S называется *дедекиндовой* (или *модулярной*). Структура S является дедекиндовой тогда и только тогда, когда для любых элементов $a, b, c \in S$ выполнено равенство $(ac+b)c = ac+bc$. Другое необходимое и достаточное условие дедекиндовости структуры S: для любых элементов $a, b, c \in S$ из равенства $ac = bc$, $a+c = b+c$ и соотношения $a \leqslant b$ должно следовать равенство $a = b$.

Пример 2. Структура всех нормальных делителей произвольной группы является дедекиндовой.

Пример 3. Дедекиндовой является всякая структура, имеющая диаграмму, изображенную на рис. 2.

Всякая подструктура дедекиндовой структуры и структура, двойственная (п. 1) дедекиндовой, сами являются дедекиндовыми структурами.

Структура S тогда и только тогда не является дедекиндовой, когда она содержит подструктуру, имеющую вид диаграммы, изображенной на рис. 3.

4. Булевы алгебры. Структура S, содержащая 0 и 1 (п. 2), называется *структурой с дополнениями*, если для всякого ее элемента a существует такой элемент $b \in S$, что $a \cap b = 0$, $a \cup b = 1$. Дистрибутивная структура (п. 3) с дополнениями называется *булевой алгеброй*.

Структура всех подмножеств данного множества (п. 2, пример 1) всегда является булевой алгеброй. Обратно, всякая конечная булева алгебра изоморфна (п. 2) структуре всех подмножеств некоторого множества. Произвольная булева алгебра изоморфна подструктуре структуры всех подмножеств некоторого множества.

В булевой алгебре дополнение к каждому элементу единственно; дополнение к пересечению (объединению) двух элементов равно объединению (пересечению) дополнений к каждому из этих элементов. В любой булевой алгебре S справедлив бесконечный дистрибутивный закон: если для бесконечного множества элементов $\ldots, a_\alpha, \ldots$ из S в S существует объединение $\bigcup\limits_\alpha a_\alpha$ и для некоторого элемента $a \in S$ в S существует объединение $\bigcup\limits_\alpha (a \cap a_\alpha)$, то

$$a \cap \left(\bigcup_\alpha a_\alpha\right) = \bigcup_\alpha (a \cap a_\alpha).$$

Подалгеброй булевой алгебры называется ее подструктура (п. 2), содержащая вместе с каждым элементом a его дополнение a'.

Если в булевой алгебре назвать произведением двух элементов их пересечение, а суммой элементов a, b назвать элемент $(a \cap b') \cup (a' \cap b)$ (a' и b' — дополнения соответственно к элементам a и b), то получится коммутативное кольцо с 1 (§ 2, п. 1), для каждого элемента a которого $2a = 0$ и $a^2 = a$ (так называемое *булево кольцо* с единицей).

Обратно, если в булевом кольце с единицей считать $a \geqslant b$, если $b = ab$, то получится булева алгебра, где $a \cap b = ab$ (в смысле умножения, определенного в кольце) и $a \cup b = a + b - ab$.

БИБЛИОГРАФИЯ

1. Александров П. С., Введение в теорию групп, М., Учпедгиз, 1938.
2. Биркгоф Г., Теория структур, М., ИЛ, 1952.
3. Бохер М., Введение в высшую алгебру, М., Гостехиздат, 1933.
4. Бронштейн И. Н. и Семендяев К. А., Справочник по математике для инженеров и учащихся втузов, Гостехиздат, изд. 7, 1957.
5. Ван-дер-Варден Б. Л.,Современная алгебра, ч. 1 и 2, Гостехиздат, 1947.
6. Гантмахер Ф. Р., Теория матриц, Гостехиздат, 1953.
7. Гантмахер Ф. Р. и Крейн М. Г., Осцилляционные матрицы и ядра и малые колебания механических систем, Гостехиздат, изд. 2, 1950.
8. Гельфанд И. М., Лекции по линейной алгебрс, Гостехиздат, изд. 2, 1951.
9. Джекобсон Н., Теория колец, ИЛ, 1947.
10. Каган В. Ф., Основания теории определителей, Государственное издательство Украины, Одесса, 1922.
11. Курош А. Г., Курс высшей алгебры, изд. 6, Физматгиз, 1959.
12. Курош А. Г., Теория групп, изд. 2, Гостехиздат, 1953.
13. Ляпин Е. С., Курс высшей алгебры, изд. 2, Учпедгиз, 1955.
14. Мальцев А. И., Основы линейной алгебры, изд. 2, Гостехиздат, 1956.
15. Мейман Н. Н., Некоторые вопросы расположения корней полиномов, УМН, 4:6 (34), 1949.
16. Окунев Л. Я., Высшая алгебра, изд. 4, Гостехиздат, 1949.
17. Окунев Л. Я., Кольцо многочленов и поле рациональных функций, Энциклопедия элементарной математики, кн. 2, Гостехиздат, 1951.
18. Окунев Л. Я., Основы современной алгебры, Учпедгиз, 1941.
19. Понтрягин Л. С., Непрерывные группы, изд. 2, Гостехиздат, 1954.
20. Проскуряков И. В., Сборник задач по линейной алгебре, Гостехиздат, 1957.
21. Смирнов В. И., Курс высшей математики, т. I, изд. 16, Гостехиздат, 1956.
22. Современная математика для инженеров, под редакцией Беккенбаха Э. Ф., ИЛ, 1958.

23. Сушкевич А. К., Основы высшей алгебры, изд. 4, Гостехиздат, 1941.
24. Фаддеев Д. К. и Соминский И. С., Сборник задач по высшей алгебре, изд. 2—6, Гостехиздат, 1949—1955.
25. Фаддеев Д. К. и Фаддеева В. Н., Вычислнтельные методы линейной алгебры, Физматгиз, 1960.
26. Фрезер Р., Дункан В., Коллар А., Теория матриц и ее приложения к дифференциальным уравнениям и динамике, ИЛ, 1950.
27. Чеботарев Н. Г., Основы теории Галуа, ч. 1, ОНТИ, 1934.
28. Чеботарев Н. Г., Теория Галуа, ОНТИ, 1936.
29. Чеботарев Н. Г. и Мейман Н. Н., Проблема Рауса — Гурвица для полиномов и целых функций, Труды Матем. ин-та им. В. А. Стеклова, т. XXVI, изд. АН СССР, 1949.
30. Чезаро Э., Элементарный учебиик алгебраического анализа и исчисления бескоиечно малых, ч. 1, ОНТИ, 1936.
31. Шапиро Г. М., Высшая алгебра, изд. 4, Учпедгиз, 1938.
32. Шилов Г. Е., Введение в теорию линейных простраиств, Гостехиздат, 1952.
33. Шмидт О. Ю., Абстрактная теория групп, изд. 2, ОНТИ, 1933.
34. Математнческий анализ (функции, пределы, ряды, цепные дроби), СМБ, М., Физматгиз, 1961.

УКАЗАТЕЛЬ ОБОЗНАЧЕНИЙ

АЛФАВИТНЫЙ УКАЗАТЕЛЬ