МЕТОДЫ СОВРЕМЕННОЙ МАТЕМАТИКИ

Выпуск 4

ВВЕДЕНИЕ В ВЫСШУЮ АЛГЕБРУ И АНАЛИТИЧЕСКУЮ ГЕОМЕТРИЮ

МСМ

В. А. Артамонов

Факториал

МЕТОДЫ СОВРЕМЕННОЙ МАТЕМАТИКИ

Выпуск 4

В. А. Артамонов

ВВЕДЕНИЕ В ВЫСШУЮ АЛГЕБРУ И АНАЛИТИЧЕСКУЮ ГЕОМЕТРИЮ

Москва
ФАКТОРИАЛ ПРЕСС
2007

УДК 512.64, 514.0
ББК 22.143, 22.151
А 86

Артамонов В. А.

А 86 Введение в высшую алгебру и аналитическую геометрию / В. А. Артамонов — М.: Факториал Пресс, 2007. — 128 с.: ил. — (Методы современной математики; Вып. 4).
ISBN 5-88688-085-2

Книга представляет собой курс лекций, неоднократно читавшийся автором для студентов первого курса факультета наук о материалах МГУ. В неё вошли такие разделы как системы линейных уравнений, матрицы, определители, линейные пространства, евклидовы и унитарные пространства, квадрики, многочлены, линейные операторы, поверхности второго порядка.

Для студентов естественных факультетов высших учебных заведений.

УДК 512.64, 514.0
ББК 22.143, 22.151

Серия

МЕТОДЫ СОВРЕМЕННОЙ МАТЕМАТИКИ

Выпуск 4

Научное издание

Вячеслав Александрович Артамонов

ВВЕДЕНИЕ В ВЫСШУЮ АЛГЕБРУ И АНАЛИТИЧЕСКУЮ ГЕОМЕТРИЮ

Формат 60×90/16. Усл. печ. л. 8. Бумага офсетная №1. Гарнитура литературная. Подписано к печати 14.09.2006. Тираж 700 экз. Заказ № 4224.

Издательство «Факториал Пресс», 117449, Москва, а/я 331; ЛР ИД № 00316 от 22.10.99.

e-mail: press@factorialco.com http://factorialco.com

Отпечатано с готовых диапозитивов издательства «Факториал Пресс» в ППП типографии «Наука» Академиздатцентра «Наука» РАН. 121099, Москва Г-99, Шубинский пер., 6.

ISBN 5-88688-085-2

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Эта книга представляет собой курс лекций по высшей алгебре и аналитической геометрии, неоднократно читавшихся автором для студентов первого курса Факультета наук о материалах МГУ. В настоящее время курс рассчитан на два семестра. В данное издание включен материал первого семестра, который содержит следующие разделы:

решение систем линейных уравнений;
операции над матрицами;
теория определителей;
комплексные числа;
многочлены и их корни;
рациональные дроби;
векторные пространства;
ранг матрицы;
плоскости и способы их задания;
билинейные и полуторалинейные функции;
евклидовы и эрмитовы пространства;
линейные операторы;
симметрические и ортогональные операторы;
классификация поверхностей второго порядка.

В программу второго семестра входят основы теории групп и их представлений. Этот материал достаточно подробно изложен в учебном пособии [2]. При рассмотрении геометрических вопросов изложение ведется, как правило, для пространств произвольной размерности. Это позволяет студенту естественнее воспринимать материал и проще с методической точки зрения. В ряде случаев рассмотрение ведется только для пространств размерности 2 или 3. При этом ясно, как исследовать общий случай. Таким образом, отобранный материал составляет основу фундаментального современного математического образования. Он необходим специалистам, работающим в различных областях современного естествознания, в первую очередь химикам и физикам.

Содержащийся в лекциях материал приведен с полными доказательствами и снабжен примерами. Изложение ясное и доступное студентам нематематических специальностей, на которых и рассчитано данное издание. Следует отметить, что предыдущее издание [1] курса лекций значительно переработано с учётом опыта преподавания и в связи с изменением учебного плана.

Читателям, интересующимся дополнительными вопросами линейной алгебры, можно порекомендовать учебники, указанные в списке литературы.

Издание подготовлено при поддержке Факультета наук о материалах МГУ и грантов РФФИ 06-01-00037, поддержки научных школ НШ 5666.2006.1. Автор выражает благодарность декану ФНМ академику РАН Ю. Д. Третьякову, коллективу факультета, а также профессорам Е. В. Майкову и А. А. Михалёву за полезные обсуждения и критические замечания, способствовавшие улучшению текста.

ГЛАВА 1

ЛИНЕЙНЫЕ УРАВНЕНИЯ И МАТРИЦЫ

Одной из важных задач алгебры является нахождение всех решений систем линейных уравнений. В этой главе мы рассмотрим метод Гаусса решения систем уравнений, называемый часто методом исключения неизвестных. Он был предложен немецким математиком XIX века Гауссом.

§1. Метод Гаусса

Рассмотрим прямоугольную систему линейных уравнений

$$\begin{cases} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{cases}, \qquad (1.1)$$

коэффициенты которой a_{ij}, b_j заданы. Наша цель — найти все решения и указать алгоритм для нахождения всех решений.

Определение 1.1. *Решением* системы (1.1) называется такой набор чисел $(\alpha_1, \dots, \alpha_n)$, что для всех $i = 1, \dots, m$ выполнены равенства $a_{i1}\alpha_1 + \dots + a_{in}\alpha_n = b_i$.

Мы будем различать следующие виды систем (1.1). Система

несовместна, если она не имеет решения;

совместна, если она имеет решение;

неопределенна, если она имеет более одного решения;

определенна, если она имеет единственное решение.

Скажем, что две системы линейных уравнений вида (1.1) *эквивалентны*, если они имеют одинаковые множества решений.

Мы будем совершать ряд простейших преобразований системы (1.1), сохраняющих множества решений. Заметим, что все информация о системе (1.1) содержится в её таблице её коэффициентов.

Матрицей системы (1.1) называется прямоугольная таблица

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \tag{1.2}$$

коэффициентов при неизвестных в системе.

Расширенной матрицей системы (1.1) называется прямоугольная таблица

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \tag{1.3}$$

коэффициентов при неизвестных и свободных членов.

Иногда расширенную матрицу системы (1.1) обозначают через

$$\left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right).$$

Следующие преобразования системы (1.1) (и её (расширенной) матрицы) называются *элементарными*:

◇ к одному уравнению (к одной строке) прибавить другое уравнение (другую строку), умноженное (умноженную) на произвольное число;

♡ умножить одно уравнение (одну строку) на ненулевое число.

Упражнение. Доказать, что любые две строки матрицы можно поменять местами, совершив 4 элементарных преобразования.

Теорема 1.2. *Элементарные преобразования переводят систему* (1.1) *в эквивалентную систему.*

Доказательство. Предположим, что мы совершаем преобразование типа $\diamondsuit$, именно, к i-ому уравнению прибавляем j-ое, умноженное на α. Пусть $(\beta_1, \ldots, \beta_n)$ — решение исходной системы (1.1). Все уравнения новой системы, кроме i-го, не изменились. Если мы подставим набор $(\beta_1, \ldots, \beta_n)$ в i-ое уравнение новой системы, то получим

$$\begin{aligned}(a_{i1} + \alpha a_{j1})\beta_1 + \ldots + (a_{in} + \alpha a_{jn})\beta_n &= \\ = (a_{i1}\beta_1 + \ldots + a_{in}\beta_n) + \alpha(a_{j1}\beta_1 + \ldots + a_{jn}\beta_n) &= b_i + \alpha b_j.\end{aligned}$$

Таким образом, $(\beta_1, \ldots, \beta_n)$ является решением новой системы. Поскольку исходная системы (1.1) получается из новой системы элементарным преобразованием прибавлением к i-ому уравнению j-го, умноженного на $-\alpha$, то, аналогично, каждое решение новой системы является решением исходной системы. $\square$

Будем приводить матрицу системы к наиболее простому — ступенчатому виду.

Матрица (1.3) называется *ступенчатой*, если

1) ниже нулевой строки находятся нулевые строки;

2) первый ненулевой элемент каждой строки равен 1;

3) если первый ненулевой элемент i-ой строки расположен в k_i-ом столбце, то

   а) $k_{i+1} > k_i$;

   б) все остальные элементы k_i-го столбца равны нулю.

Теорема 1.3. *Каждая матрица конечным числом элементарных преобразований строк приводится к ступенчатому виду.*

Доказательство. Пусть матрица A имеет вид (1.3). Если $A = 0$, то она уже имеет ступенчатый вид.

Пусть $A \neq 0$. Будем вести доказательство индукцией по числу строк m. Без ограничения общности можно считать, что в первом столбце есть ненулевой элемент a_{i1}. Если $i > 1$ и $a_{11} = 0$, то к первой строке можно прибавить i-ую и добиться, чтобы $a_{11} \neq 0$. Далее умножив 1-ую строку на a_{11}^{-1} можно считать, что $a_{11} = 1$. Отсюда следует утверждение теоремы при $m = 1$.

Пусть $m > 1$, и для $m - 1$ теорема доказана. Для каждого $i > 1$ вычтем из i-ой строки первую строку, умноженную на a_{i1}. В новой матрице все коэффициенты $a_{i1} = 0$, $i > 1$.

Рассмотрим подматрицу B в A, получающуюся отбрасыванием первой строки. По индукции можно считать, что матрица B имеет ступенчатый вид. Пусть в матрице B первые ненулевые элементы расположены в столбцах с номерами $1 < k_2 < k_3 < \dots$ Вычтем из первой строки 2-ю строку, умноженную на a_{1,k_2}, 3-ю строку, умноженную на a_{1,k_3}, и т. д. В результате получается ступенчатая матрица. □

Пусть матрица системы (1.1) имеет ступенчатый вид. Назовем неизвестную x_i *главной*, если в некотором уравнении все коэффициенты при $x_1, \dots, x_{i-1}$ равны нуля, а коэффициент при x_i отличен от нуля (и потому равен 1). Все остальные неизвестные назовем *свободными*.

Применим теоремы 1.2, 1.3 к исследованию системы (1.1). В силу указанных теорем можно считать, что расширенная матрица система (1.1) имеет ступенчатый вид.

Пусть её последняя ненулевая строка имеет вид

$$(0, \dots, 0, 1). \tag{1.4}$$

Это означает, что системы (1.1) содержит несовместное уравнение $0x_1 + \dots + 0x_n = 1$. Поэтому в этом случае исходная система несовместна.

Пусть в A нет строки (1.4). Предположим для простоты, что переменные $x_1, \dots, x_r$ главные, а $x_{r+1}, \dots, x_n$ свободные. Тогда система имеет вид

$$\begin{cases} x_1 \qquad\qquad + a_{1,r+1}x_{r+1} + \ldots + a_{1n}x_n = b_1 \\ \quad x_2 \qquad + a_{2,r+1}x_{r+1} + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\quad x_r + a_{r,r+1}x_{r+1} + \ldots + a_{rn}x_n = b_r \end{cases}.$$

Перенося свободные переменные в правую часть, получаем выражение главных неизвестных через свободные

$$\begin{cases} x_1 = b_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n \\ x_2 = b_2 - a_{2,r+1}x_{r+1} - \ldots - a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r = b_r - a_{r,r+1}x_{r+1} - \ldots - a_{rn}x_n \end{cases}.$$

Таким образом, придавая свободным неизвестным произвольные значения, мы однозначно находим значения главных неизвестных. Итак, система совместна, и, если есть свободные неизвестные, то система неопределенна. Если все неизвестные главные, то система определенна.

Разберём следующий пример. Пусть задана система уравнений

$$\begin{cases} 2x_1 - 2x_2 + 3x_3 = 3 \\ -x_1 + x_2 - x_3 = 1 \end{cases}.$$

Составим расширенную матрицу системы

$$\left(\begin{array}{ccc|c} 2 & -2 & 3 & 3 \\ -1 & 1 & -1 & 1 \end{array}\right).$$

и приведём её с ступенчатому виду. Для этого к первой строке прибавим вторую и получим

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ -1 & 1 & -1 & 1 \end{array}\right).$$

Далее ко второй строке прибавим первую

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 4 \\ 0 & 0 & 1 & 5 \end{array}\right).$$

Затем из первой строки вычтем вторую

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 5 \end{array}\right).$$

Таким образом, исходная система эквивалентная ступенчатой системе

$$\begin{cases} x_1 - x_2 = -1 \\ \qquad x_3 = 5 \end{cases}.$$

Главными являются неизвестные x_1, x_3. Неизвестная x_2 свободная. Поэтому система совместна и неопределённа, а её ответ записывается в виде

$$\begin{cases} x_1 = -1 + x_2 \\ x_3 = 5 \end{cases}.$$

Определение. Система (1.1) *однородна*, если все её свободные члены нулевые, т. е. $b_1 = \ldots = b_m = 0$.

Предложение 1.4. *Если в однородной системе число неизвестных n больше числа уравнений m, то система неопределенна.*

Доказательство. Приведем систему уравнений к ступенчатому виду. При этом мы получим снова однородную систему, в которой число главных неизвестных не превосходит числа ненулевых уравнений исходной системы. Поэтому не все неизвестные главные, и им можно придать ненулевые значения. Тем самым получается ненулевое решение системы. □

§ 2. Матрицы и операции над ними

При решении систем уравнений методом Гаусса возникла необходимость работы с матрицами. В этом разделе мы рассмотрим основным операции над матрицами. Введем следующие обозначения. Через $\mathrm{Mat}(n \times m)$ будем обозначать множество всех матриц с n строками и m столбцами. Если $A \in \mathrm{Mat}(n \times m)$, то мы будем также писать $A = A_{n\times m}$. Если $A_{n\times m} = (a_{ij})$, $B_{n\times m} = (b_{ij})$, то через сумму $A+B$ обозначим матрицу того же размера, в которой на месте (i, j) стоит сумма $a_{ij} + b_{ij}$ соответствующих элементов слагаемых. Кроме того, $\lambda A_{n\times m} = (\lambda a_{ij})$.

Предложение 1.5. *Пусть* $A, B, C \in \mathrm{Mat}(n \times m)$ *и* λ, ν — *числа. Тогда справедливы следующие* 8 *аксиом векторного пространства*:

1) $A + B = B + A$;

2) $A + (B + C) = (A + B) + C$;

3) *если* 0 — *нулевая матрица (все её коэффициенты равны нулю), то* $A + 0 = A$ *для любой матрицы* A;

4) *для любой матрицы* A *существует такая матрица* $-A$, *что* $A + (-A) = 0$;

5) $\lambda(A + B) = \lambda A + \lambda B$;

6) $(\lambda + \nu)A = \lambda A + \nu A$;

7) $(\lambda\nu)A = \lambda(\nu A)$;

8) $1A = A$.

Доказательство. Рассмотрим первое утверждение. Если

$$A = (a_{ij}), \quad B = (b_{ij}),$$

то

$$A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A.$$

Остальные утверждения доказываются аналогично. □

Определение 1.6. Пусть заданы две матрицы

$$A_{n\times m} = (a_{ij}), \quad C_{m\times k} = (c_{st}).$$

Тогда их произведение $D = AC \in \mathrm{Mat}(n \times k)$ определяется как матрица, в которой на месте (i, j) стоит элемент

$$d_{is} = a_{i1}d_{1s} + \ldots + a_{im}d_{ms} \tag{1.5}$$

для всех $i = 1, \ldots, n$, $s = 1, \ldots, k$.

Например,

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

С помощью матричного умножения удобно в компактной форме записывать системы линейных уравнений. Именно, система (1.1) имеет вид $AX = b$, где A — матрица (1.2) системы (1.1),

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

— столбцы, составленные из неизвестных и из свободных членов, соответственно.

Предложение 1.7. *Умножение матриц ассоциативно, т. е.*

$$(AC)F = A(CF)$$

для любых матриц

$$A \in \mathrm{Mat}(n \times m), \quad C \in \mathrm{Mat}(m \times k), \quad F \in \mathrm{Mat}(k \times l).$$

Доказательство. Пусть

$$A = A_{n\times m} = (a_{ij}), \quad C = C_{m\times k} = (c_{st}), \quad F = F_{k\times l} = (f_{tq}).$$

Если D — матрица из определения 1.6, то по (1.5) на месте (i, q) в матрице $DF = (AC)F$ стоит элемент

$$\sum_{\alpha=1}^{k} d_{i,\alpha} f_{\alpha,q} = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{n} a_{i,\beta} c_{\beta,\alpha} f_{\alpha,q}\,. \tag{1.6}$$

С другой стороны, если

$$CF = U = (u_{\beta,q}) \in \mathrm{Mat}(m \times l),$$

то на месте (i,q) в матрице $AU = A(CF)$ стоит элемент

$$\sum_{\beta=1}^{n} a_{i,\beta}u_{\beta,q} = \sum_{\beta=1}^{n}\sum_{\alpha=1}^{k} a_{i,\beta}c_{\beta,\alpha}f_{\alpha,q}\,. \tag{1.7}$$

Из (1.6), (1.7) вытекает утверждение. □

Предложение 1.8. *Справедливы равенства:*

1) $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

2) $A(B+C) = AB + AC, \quad (A+U)V = AV + UV$.

Доказательство. Докажем, например, второе утверждение. Пусть $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. Тогда на месте (i,j) в матрице $A(B+C)$ строит элемент

$$\sum a_{ik}(b_{kj} + c_{kj}) = \sum a_{ik}b_{kj} + \sum a_{ik}c_{kj}\,,$$

который равен элементу, стоящему на то же месте в матрице $AB + AC$. Так как размеры матриц $A(B+C)$ и $AB + AC$ совпадают, то они равны. Аналогично проверяются остальные утверждения. □

Следом $\mathrm{tr}(A)$ матрицы $A = (a_{ij}) \in \mathrm{Mat}(n)$ называется сумма $a_{11} + \ldots + a_{nn}$ её элементов, стоящих на главной диагонали.

Предложение. *Пусть* $A, B \in \mathrm{Mat}(n)$. *Тогда*

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Доказательство. Пусть $A = (a_{ij}), B = (b_{ij})$. Тогда на месте (i,i) в матрице AB стоит $\sum_{j=1}^{n} a_{ij}b_{ji}$, откуда

$$\mathrm{tr}(AB) = \sum_{i,j=1}^{n} a_{ij}b_{ji}\,.$$

Аналогично,

$$\mathrm{tr}(BA) = \sum_{s,t=1}^{n} b_{st}a_{ts} = \sum_{s,t=1}^{n} a_{ts}b_{st} = \mathrm{tr}(AB).$$ □

Символ Кронекера δ_{ij} равен 1, если $i = j$, и 0, если $i \neq j$. *Единичная* матрица $E = E_n \in \mathrm{Mat}(n)$ — это матрица, в которой на месте (i, j) стоит символ Кронекера δ_{ij}.

Предложение 1.9. *Пусть* $A \in \mathrm{Mat}(n \times m)$. *Тогда*

$$E_n A = A = A E_m\,.$$

Доказательство. Пусть $A = (a_{ij})$. Тогда на месте (i, j) в матрице $E_n A$ стоит $\sum_{k=1}^{n} \delta_{ik} a_{kj} = \delta_{ii} a_{ij} = a_{ij}$, т. е. $E_n A = A$. □

Транспонированной матрицей ${}^tA = A^* \in \mathrm{Mat}(m \times n)$ для матрицы $A \in \mathrm{Mat}(n \times m)$ называется матрица, в которой на месте (i, j) стоит элемент a_{ji} матрицы A.

Предложение 1.10. *Справедливы следующие равенства*:

$${}^t(A + B) = {}^tA + {}^tB, \quad {}^t(\lambda A) = \lambda\, {}^tA, \quad {}^t(AC) = {}^tC\, {}^tA.$$

Доказательство. Докажем, например, последнее утверждение. В матрице ${}^t(AC)$ на месте (i, j) стоит

$$\sum_k a_{jk} c_{ki} = \sum_k c_{ki} a_{jk}\,,$$

т. е. элемент, стоящей на том же месте в матрице ${}^tC\,{}^tA$.

Аналогично доказываются остальные утверждения. □

Матричной единицей $E_{ij} \in \mathrm{Mat}(n \times m)$ называется матрица E_{ij}, в которых на месте (s, t) стоит элемент $\delta_{si}\delta_{tj}$. Другими словами в E_{ij} на месте (i, j) стоит 1, и все остальные элементы равны 0.

Упражнение. Доказать, что ${}^tE_{ij} = E_{ji}$ и если $A = (a_{ij})$, то

$$A = \sum_{i,j} a_{ij} E_{ij}\,.$$

Предложение 1.11. $E_{ij} E_{rs} = \delta_{jr} E_{is}\,.$

Доказательство. На месте (u, v) в $E_{ij}E_{rs}$ стоит элемент

$$\sum_p (\delta_{ui}\delta_{pj})(\delta_{pi}\delta_{vs}) = \begin{cases} 1, & u = i = p = j,\ v = s; \\ 0, & \text{в противном случае.} \end{cases}$$

Отсюда вытекает утверждение. □

Следствие 1.12. *Пусть* $A = (a_{rs}) \in \mathrm{Mat}(n \times m)$. *Тогда*

$$E_{ij}A = a_{j1}E_{i1} + \ldots + a_{jm}E_{im}\,,$$
$$AE_{ij} = a_{1i}E_{1j} + \ldots + a_{nj}E_{ni}\,.$$

Теорема 1.13. *Чтобы в матрице* $A \in \mathrm{Mat}(n \times m)$ *к* i*-ой строке прибавить* j*-ую, умноженную на* α*, нужно рассмотреть матрицу* $(E_n + \alpha E_{ij})A$.

Доказательство. В силу предложениий 1.8, 1.11, а также упражнения 1.9, и следствия 1.12 получаем

$$(E_n + \alpha E_{ij})A = E_nA + \alpha E_{ij}A = A + \alpha(a_{j1}E_{i1} + \ldots + a_{jm}E_{im}) =$$

$$= \sum_{rs} a_{rs}E_{rs} + (\alpha a_{j1})E_{i1} + \ldots + (\alpha a_{jm})E_{im} =$$

$$= \sum_{r \neq i,s} a_{rs}E_{rs} + \sum_{i,s} (a_{is} + \alpha a_{js})E_{is}\,. \qquad \square$$

Следствие. *Чтобы в матрице* $A \in \mathrm{Mat}(n \times m)$ *к* i*-ому столбцу прибавить* j*-ый, умноженный на* α *нужно рассмотреть матрицу* $A(E_m + \alpha E_{ji})$.

Обозначим через $D_i(\alpha) = E_n + (\alpha - 1)E_{ii} \in \mathrm{Mat}(n)$ диагональную матрицу, в которой i-ый диагональный элемента равен α, а остальные диагональные элементы равны 1.

Теорема 1.14. *Пусть задана матрица* $A \in \mathrm{Mat}(n \times m)$. *Тогда матрица* $D_i(\alpha)A$ *получается из* A *умножением* i*-ой строки на* α*. Матрица* $AD_i(\alpha)$ *получается из* A *умножением* i*-го столбца на* α.

Доказательство. Пусть $A = (a_{ij})$. В матрице $D_i(\alpha)$ на месте (s, t) стоит $\delta_{st} + (\alpha - 1)\delta_{is}\delta_{it}$. Поэтому в матрице $D_i(\alpha)A$ на месте (p, q) стоит

$$\sum_r (\delta_{pr} + (\alpha - 1)\delta_{ip}\delta_{ir})a_{rq} = \sum_r \delta_{pr}a_{rq} + (\alpha - 1)\sum_r \delta_{ip}\delta_{ir}a_{rq} =$$

$$= a_{pq} + (\alpha - 1)\delta_{ip}a_{iq} = \begin{cases} a_{pq}, & i \neq q; \\ \alpha a_{iq}, & p = i. \end{cases}$$

$\square$

Элементарными матрицами называются квадратные матрицы вида $E_n + \alpha E_{ij}$ при $i \neq j$ и вида $D_i(\beta)$ при $\beta \neq 0$. Таким образом, в теоремах 1.13 и 1.14 доказано, что каждое элементарное преобразование строк или столбцов матрицы эквивалентно умножению этой матрицы на соответствующую элементарную матрицу слева или справа.

ГЛАВА 2

ПЕРЕСТАНОВКИ И ИХ ЗНАКИ

В ряде важных случаев решение систем линейных уравнений можно вычислять с помощью определителей. Для изучения определителей над потребуется ряд утверждений о знаках перестановок.

Перестановкой степени n называется упорядоченный набор

$$\sigma = (i_1, \ldots, i_n) \tag{2.1}$$

чисел $(1, 2, \ldots, n)$. Через S_n обозначается множество всех перестановок степени n.

Предложение. *Число всех перестановок равно* $n!$.

Доказательство. Пересчитаем все перестановки вида (2.1). Число i_1 принимает произвольные n значений. Если i_1 фиксировано, то i_2 принимает любое значение, отличное от i_1. Поэтому i_2 принимает $(n-1)$ значение и т. д. $\square$

Инверсией в перестановке (2.1) называется такая пара i_s, i_t, что $s < t$ и $i_s > i_t$. *Знаком* $(-1)^\sigma$ перестановки σ называется число $(-1)^M$, где M — число инверсий в σ из (2.1). Подстановка *четна* (*нечетна*), если её знак $+1(-1)$. *Транспозицией* называется перестановка двух элементов перестановки.

Теорема 2.1. *Транспозиция меняет знак перестановки.*

Доказательство. Пусть задана перестановка

$$\sigma = (i_1, \dots, i_k, \dots, i_r, \dots, i_n) \in S_n .$$

Сравним число инверсий этой перестановки и перестановки

$$\tau = (i_1, \dots, i_{k-1}, i_r, i_{k+1}, \dots, i_{r-1}, i_k, i_{r+1}, \dots, i_n).$$

Предположим сначала, что $k = r - 1$. В этом случае взаимное расположение любой пары (i_d, i_m), отличной от пары (i_k, i_r) в σ и τ одинаково и потому они дают одинаковый вклад в подсчет общего числа инверсий. Из пар (i_k, i_r), (i_r, i_k) одна дает инверсию, а другая нет. Отсюда вытекает утверждение.

Пусть $k < r - 1$. Тогда переставляя i_k последовательно с элементами $i_{k+1}, \dots, i_{r-1}, i_r$ мы будем менять знак $r - k$ раз. Затем элемент i_r будем последовательно переставлять с $i_{r-1}, \dots, i_{k+1}$. При этом получим перестановку τ изменив знак ещё $r - k - 1$ раз. Таким образом, при переходе от σ к τ мы меняли знак $(r - k) + (r - k - 1) = 2(r - k) - 1$ раз. □

Следствие. *Число всех четных перестановок степени $n \geqslant 2$ равно числу нечетных и равно $\frac{n!}{2}$.*

Доказательство. Поменяем местами во всех перестановках (2.1) первые два индекса i_1, i_2. По теореме каждая четная перестановка становится четной и наоборот, причем разные перестановки переходят в разные. Поэтому число четных перестановок не больше числа четных и наоборот. Следовательно, число тех и других совпадает и равно $\frac{n!}{2}$. □

Следствие 2.2. *Рассмотрим две матрицы*

$$\begin{pmatrix} 1 & \dots & n \\ i_1 & \dots & i_n \end{pmatrix}, \quad \begin{pmatrix} j_1 & \dots & j_n \\ 1 & \dots & n \end{pmatrix},$$

строки которых являются перестановками степени n. Предположим, что, переставляя столбцы, первую матрицу можно перевести во вторую. Тогда знаки перестановок $(i_1, \dots, i_n)$ и $(j_1, \dots, j_n)$ совпадают.

Доказательство. По теореме 2.1 при любой перестановке столбцов произведение знаков перестановок верхней и нижней строк не меняется. При этом чётны в начальной перестановке — первая строка, в последней — нижняя строка. Следовательно, указанное постоянное произведение равно, с одной стороны, знаку $(i_1, \ldots, i_n)$, а, с другой стороны, — знаку $(j_1, \ldots, j_n)$. $\square$

ГЛАВА 3

ОПРЕДЕЛИТЕЛИ, ОБРАТНАЯ МАТРИЦА

В этой главе мы рассматриваем теорию определителей и её приложения для вычисления обратных матриц и нахождения решения систем линейных уравнений.

§ 1. Определители

Определителем $\det A = |A|$ квадратной матрицы

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \tag{3.1}$$

называется число

$$\det A = \sum_{(i_1,\dots,i_n)\in S_n} (-1)^{(i_1,\dots,i_n)} a_{1,i_1} \cdots a_{n,i_n}\,. \tag{3.2}$$

Теорема 3.1. *Пусть* $A = (a_{ij})$ — *верхнетреугольная матрица, т. е.* $a_{ij} = 0$ *при* $i > j$. *Тогда*

$$\det A = a_{11} \cdots a_{nn}\,.$$

Доказательство. В матрице A элемент $a_{ij} = 0$, если $i > j$. Таким образом, если в определителе $\det A$ произведение $a_{1,i_1} \cdots a_{n,i_n}$ отлично от нуля, то $1 \leqslant i_1$, $2 \leqslant i_2, \dots, n \leqslant i_n$. Учитывая, что индексы $i_1, \dots, i_n$ различны, получаем $n = i_n$, $i_{n-1} = n-1, \dots, i_1 = 1$. □

Теорема 3.2 (Простейшие свойства определителя). *Если одна из строк матрицы A является линейно комбинацией двух строк, то $\det A$ является линейной комбинацией определителей соответствующих матриц.*

Доказательство. Пусть s-ая строка является указанной линейной комбинацией, т. е. $a_{sj} = \alpha a'_{sj} + \beta a''_{sj}$ для всех $j = 1, \ldots, n$. Тогда

$$\begin{aligned}
&\det A = \\
&= \sum (-1)^{(i_1,\ldots,i_n)} a_{1,i_1} \ldots a_{s-1,i_{s-1}} (\alpha a'_{sj} + \beta a''_{sj})\, a_{s+1,i_{s+1}} \ldots a_{n,i_n} = \\
&= \alpha \sum (-1)^{(i_1,\ldots,i_n)} a_{1,i_1} \cdots a_{s-1,i_{s-1}} a'_{sj} a_{s+1,i_{s+1}} \cdots a_{n,i_n} + \\
&+ \beta \sum (-1)^{(i_1,\ldots,i_n)} a_{1,i_1} \cdots a_{s-1,i_{s-1}} a''_{sj} a_{s+1,i_{s+1}} \cdots a_{n,i_n}\,,
\end{aligned}$$

где суммы берутся по всем перестановкам $(i_1, \ldots, i_n) \in S_n$. □

Следствие 3.3. *Элементарные преобразования строк типа ♡ из главы 1 приводят к умножению определителя матрицы на указанное число.*

Теорема 3.4. *Если в A две строки равны, то* $\det A = 0$.

Доказательство. Пусть в A k-ая и j-ая строки равны, т. е. $a_{ks} = a_{js}$ для всех s, где $k < j$. В (3.2) входит произведение

$$a_{1,i_1} \cdots a_{k,i_k} \cdots a_{j,i_j} \cdots a_{n,i_n} \tag{3.3}$$

со знаком $(-1)^{(i_1,\ldots,i_k,\ldots,i_j,\ldots,i_n)}$. Заметим, что в (3.2) входит также произведение

$$a_{1,i_1} \cdots a_{k,i_j} \cdots a_{j,i_k} \cdots a_{n,i_n} \tag{3.4}$$

со знаком $(-1)^{(i_1,\ldots,i_j,\ldots,i_k,\ldots,i_n)} = -(-1)^{(i_1,\ldots,i_k,\ldots,i_j,\ldots,i_n)}$ в силу теоремы 2.1. По условию произведения (3.3), (3.4) равны. □

Теорема 3.5. *При элементарных преобразованиях ◇ из главы 1 определитель не меняется.*

Доказательство. Выделим в матрице A строки A_i и A_j с номерами $i < j$,

$$A = \begin{pmatrix} \dots \\ A_i \\ \dots \\ A_j \\ \dots \end{pmatrix}.$$

После преобразования $\diamondsuit$ из главы 1 получаем матрицу

$$\begin{pmatrix} \dots \\ A_i + \lambda A_j \\ \dots \\ A_j \\ \dots \end{pmatrix}.$$

Ее определитель по теореме 3.2 и теореме 3.4 равен

$$\begin{vmatrix} \dots \\ A_i + \lambda A_j, \\ \dots \\ A_j \\ \dots \end{vmatrix} = \begin{vmatrix} \dots \\ A_i \\ \dots \\ A_j \\ \dots \end{vmatrix} + \lambda \begin{vmatrix} \dots \\ A_j \\ \dots \\ A_j \\ \dots \end{vmatrix} = \det A.$$

□

Следствие. *При перестановке строк матрицы её определитель меняет знак.*

Доказательство. Рассмотрим вспомогательную матрицу B, в которой i-ая и j-ая строки равны сумме $A_i + A_j$ i-ой и j-ой строк исходной матрицы A,

$$B = \begin{pmatrix} \dots \\ A_i + A_j \\ \dots \\ A_i + A_j \\ \dots \end{pmatrix}.$$

Тогда по теореме 3.2

$$\det B = \begin{vmatrix} \cdots \\ A_i + A_j \\ \cdots \\ A_i + A_j \\ \cdots \end{vmatrix} = \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_i + A_j \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_i + A_j \\ \cdots \end{vmatrix} =$$

$$= \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_i \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_i \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_j \\ \cdots \end{vmatrix} = \begin{vmatrix} \cdots \\ A_i \\ \cdots \\ A_j \\ \cdots \end{vmatrix} + \begin{vmatrix} \cdots \\ A_j \\ \cdots \\ A_i \\ \cdots \end{vmatrix}.$$

Отсюда вытекает утверждение. □

Теорема. *Определитель матрицы не меняется при транспонировании.*

Доказательство. В (3.2) входит произведение (3.3) со знаком $(-1)^{(i_1,\ldots,i_n)}$. Это же произведение входит в $\det({}^tA)$ со знаком $(-1)^{(j_1,\ldots,j_n)}$. Остается воспользоваться следствием 2.2. □

Следствие. *Все свойства строк в* $\det A$ *справедливы и для столбцов.*

Теорема 3.6. *Пусть* $A \in \mathrm{Mat}(n)$ *элементарными преобразованиями строк приводится к ступенчатой матрице* S. *Тогда определители* A *и* S *отличаются на ненулевой множитель. В частности, если* $|A| = 0$, *то последняя строка в* S *является нулевой; если* $|A| \neq 0$, *то* $S = E_n$ *— единичная матрица.*

Доказательство. При приведении к ступенчатому виду S некоторые строки умножались на ненулевые числа. Следовательно, в результате $|A|$ умножился на ненулевое число. Заметим, что ступенчатая матрица S является треугольной. Если $|A| = 0$, то и $|S| = 0$. Отсюда по теореме 3.1 в S последняя строка нулевая. Если же $|A| \neq 0$, то и $|S| \neq 0$, т. е. $S = E_n$ в силу определения ступенчатой матрицы. □

Следствие 3.7. *Пусть* $A \in \mathrm{Mat}(n)$. *Следующие условия эквивалентны:*

1) *матрица A элементарными преобразованиями строк приводится к единичной матрице E_n;*

2) $|A| \neq 0$.

Теорема 3.8. *Определитель произведения квадратных матриц одинакового размера равен произведению определителей сомножителей.*

Доказательство. С помощью элементарных преобразований строк приведём матрицу A к ступенчатой матрице S. По теореме 3.6 $|A| = \alpha|S|$, где $\alpha \neq 0$. В силу теорем 1.13 и 1.14 $S = C_1 \cdots C_k A$ для некоторых элементарных матриц $C_1, \ldots, C_k$. Из ассоциативности умножениями матриц следует. что

$$SC = (C_1 \cdots C_k A)C = (C_1 \cdots C_k)(AC),$$

т. е. SC получается из AC теми же элементарными преобразованиями строк. Поэтому $|AC| = \alpha|SC|$.

Если $|A| = 0$, то в S последняя строка нулевая. Но тогда и в SC последняя строка нулевая. Следовательно, $|SC| = 0$. Таким образом, $|AC| = \alpha|SC| = 0 = |A||C|$.

Пусть $|A| \neq 0$. По теореме 3.6 $S = E$ и $|A| = \alpha|S| = \alpha$. Кроме того, $|AC| = \alpha|SC| = \alpha|EC| = \alpha|C| = |A||C|$. □

Теорема 3.9 (Определитель с углом нулей). *Если заданы матрицы* $A \in \mathrm{Mat}(n)$, $C \in \mathrm{Mat}(n \times m)$, $B \in \mathrm{Mat}(m)$, *то*

$$\left|\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right| = (\det A)\,(\det B)\,. \tag{3.5}$$

Доказательство. Приведем матрицы A и B к ступенчатому виду.

$$A \rightsquigarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \dots & a_{2,n-1} & a_{2,n} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix},$$

$$B \rightsquigarrow \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1,m-1} & b_{1m} \\ 0 & b_{22} & \dots & b_{2,m-1} & b_{2,m} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & b_{mm} \end{pmatrix}.$$

По теоремам 3.5, 3.1 и следствию 3.3 получаем, что

$$\det A = \lambda a_{11} \cdots a_{nn}, \quad \det B = \mu b_{11} \cdots b_{mm},$$

где λ, μ возникают из-за применения преобразований типа $\heartsuit$. Те же преобразования, примененные к матрице (3.5), приводят её к верхнетреугольному виду с элементами

$$a_{11}, \dots, a_{nn}, b_{11}, \dots, b_{mm}$$

по главной диагонали. При этом определитель матрицы (3.5) равен $\lambda\mu a_{11} \cdots a_{nn} b_{11} \cdots b_{mm} = (\det A)(\det B)$. □

Следствие. *Определитель Вандермонда*

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i).$$

Доказательство. Случай $n = 2$ очевиден. Пусть для $n - 1$ утверждение верно. Вычтем из каждой строки, начиная снизу, предыдущую, умноженную на x_1. Получаем

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & x_2 - x_1 & \dots & x_n - x_1 \\ 0 & x_2^2 - x_2x_1 & \dots & x_n^2 - x_nx_1 \\ \hdotsfor{4} \\ 0 & x_2^{n-1} - x_2^{n-2}x_1 & \dots & x_n^{n-1} - x_n^{n-2}x_1 \end{vmatrix} =$$

$$= \begin{vmatrix} x_2 - x_1 & \dots & x_n - x_1 \\ x_2^2 - x_2x_1 & \dots & x_n^2 - x_nx_1 \\ \hdotsfor{3} \\ x_2^{n-1} - x_2^{n-2}x_1 & \dots & x_n^{n-1} - x_n^{n-2}x_1 \end{vmatrix} =$$

$$= (x_2 - x_1) \cdots (x_n - x_1) \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_2 & x_3 & \dots & x_n \\ x_2^2 & x_3^2 & \dots & x_n^2 \\ \hdotsfor{4} \\ x_2^{n-2} & x_3^{n-2} & \dots & x_n^{n-2} \end{vmatrix} =$$

$$= (x_2 - x_1) \cdots (x_n - x_1) \prod_{2 \leqslant i < j \leqslant n} (x_j - x_i) = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i).$$

□

Минором M_{ij}, $1 \leqslant i, j \leqslant n$, матрицы (3.1) называется определитель матрицы размера $n - 1$, получающейся вычёркиванием i-ой строки и j-го столбца. *Алгебраическим дополнением* A_{ij} называется число

$$a_{ij} = (-1)^{i+j} M_{ij}\,.$$

Теорема 3.10 (Разложение определителя по строке и столбцу). *Предположим, что задана матрица A из* (3.1). *Тогда любого $i = 1, \dots, n$ имеем*

$$\det A = a_{i1}A_{i1} + \ldots + a_{in}A_{in}\,.$$

Доказательство. Имеем

$$(a_{i1}, \ldots, a_{in}) = \sum_{j=1}^{n}(0, \ldots, 0, a_{ij}, 0, \ldots, 0).$$

По теореме 3.2 можно считать, что i-ая строка имеет вид

$$(0, \ldots, 0, a_{ij}, 0, \ldots, 0).$$

Поместив эту строку на первое место, переставляя ее со всеми предыдущими, мы умножим определитель матрицы на $(-1)^{i-1}$. Затем, переставляя столбцы, мы умножим определитель матрицы на $(-1)^{j-1}$ и поместим j-й столбец на первое место. Итак, определитель матрицы умножится на $(-1)^{i+j}$ и по теореме 3.9 он станет равным $a_{ij}M_{ij}$. □

Следствие 3.11 (Фальшивое разложение). *Если $i \neq j$, то*

$$a_{i1}A_{j1} + \ldots + a_{in}A_{jn} = 0.$$

Доказательство. Воспользоваться теоремой 3.4 для вспомогательной матрицы, получающейся из A заменой j-ой строки на i-ую. □

Пусть задана квадратная матрица $A = (a_{ij}) \in \mathrm{Mat}(n)$. Через $\widehat{A} \in \mathrm{Mat}(n)$ обозначим матрицу, в которой на месте (i, j) стоит алгебраическое дополнение A_{ji}.

Теорема 3.12. *Если $A \in \mathrm{Mat}(n)$, то*

$$A\widehat{A} = \widehat{A}A = |A|E_n.$$

Доказательство. На месте (i, j) в $A\widehat{A}$ стоит

$$a_{i1}A_{j1} + \ldots + a_{in}A_{jn} = 0.$$

Остается воспользоваться теоремой 3.10 и следствием 3.11. □

§ 2. Обратная матрица. Матричные уравнения

Определение 3.13. Матрица $A^{-1} \in \mathrm{Mat}(n)$ называется *обратной* к матрице $A \in \mathrm{Mat}(n)$, если $AA^{-1} = A^{-1}A = E_n$.

Из предложения 1.11 вытекает

Предложение 3.14. *Если $i \neq j$, $\beta \neq 0$, то*

$$(E + \alpha E_{ij})^{-1} = E - \alpha E_{ij}, \quad D_i(\beta)^{-1} = D_i(\beta^{-1}).$$

Предложение. *Если A^{-1} существует, то она единственна.*

Доказательство. Пусть заданы две матрицы B, C, обратные к матрице A. Тогда $AC = E = BA$, откуда $B = B(AC) = (BA)C = C$. □

Теорема 3.15. *Пусть $A \in \mathrm{Mat}(n)$. Обратная матрица A^{-1} существует тогда и только тогда, когда $\det A \neq 0$.*

Доказательство. Пусть матрица A^{-1} существует. Из определения 3.13 и теоремы 3.8 следует, что $1 = \det A \det(A^{-1})$ и поэтому $\det A \neq 0$.

Обратно, пусть $\det A \neq 0$. Рассмотрим матрицу

$$B = \frac{1}{|A|}\widehat{A} = (b_{ij}), \text{ где } b_{ij} = \frac{A_{ji}}{\det A}. \tag{3.6}$$

По теореме 3.12 $BA = AB = E$, т. е. $B = A^{-1}$. □

Теорема 3.16 (Теорема Крамера). *Квадратная система линейных уравнений $AX = b$ определенна тогда и только тогда, когда $\det A \neq 0$. Если $\det A \neq 0$, то решение*

$$x_i = \frac{\det A'_i}{\det A},$$

где матрица A'_i получается из матрицы A заменой i-го столбца на b.

Доказательство. Система $AX = b$ определенна тогда и только тогда, когда все её неизвестные главные. Это эквивалентно тому, что ступенчатый вид A является единичной матрицей. В силу теореме 3.7 это эквивалентно тому, что $\det A \neq 0$.

Пусть $\det A \neq 0$. По теореме 3.15 существует A^{-1}. Умножим уравнение на A^{-1} и получим $X = A^{-1}b$. Подставляя (3.6) и пользуясь разложением A_i' по i-му столбцу, завершаем доказательство. □

Формула (3.6) позволяет вычислять элементы A^{-1}. Её удобно применять для квадратных матриц размера 2. Именно, если

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

то

$$A^{-1} = \begin{pmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Укажем другой способ вычисления A^{-1}. Для этого нужно решить матричное уравнение $AX = E_n$. Рассмотрим общий случай решения матричного уравнения $AX = B$, где $A \in \mathrm{Mat}(n)$, причем $\det A \neq 0$, и $X, B \in \mathrm{Mat}(n \times m)$.

Теорема 3.17. *Расширенную матрицу $(A \mid B)$ элементарными преобразованиями строк приведем к ступенчатому виду $(E \mid C)$. Тогда $C = X$.*

Доказательство. В силу теорем 1.13, 1.14 умножим уравнение $AX = B$ на некоторые элементарные обратимые матрицы (см. предложение 3.14) $Z_1 \cdots Z_k$ и получим уравнение

$$Z_1 \cdots Z_k AX = Z_1 \cdots Z_k B,$$

где по условию $Z_1 \cdots Z_k A = E$. Отсюда $X = Z_1 \cdots Z_k B = C$. □

Заметим, что в условии теоремы 3.17 решение уравнения $AX = B$ единственно, так как оно имеет вид $X = A^{-1}B$.

Рассмотрим пример решения матричного уравнения

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

где X — матрица размера 2×3. Заметим, что матрица A невырожденна. Составим расширенную матрицу

$$\left(\begin{array}{cc|ccc} 2 & -1 & 0 & -1 & 0 \\ 1 & -1 & 1 & -1 & 3 \end{array}\right)$$

и приведем её к ступенчатому виду. Для этого из первой строки вычтем вторую и получим

$$\left(\begin{array}{cc|ccc} 1 & 0 & -1 & 0 & -3 \\ 1 & -1 & 1 & -1 & 3 \end{array}\right).$$

Далее из второй строки вычтем первую и получим

$$\left(\begin{array}{cc|ccc} 1 & 0 & -1 & 0 & -3 \\ 0 & -1 & 2 & -1 & 6 \end{array}\right).$$

Затем вторую строку умножаем на -1

$$\left(\begin{array}{cc|ccc} 1 & 0 & -1 & 0 & -3 \\ 0 & 1 & -2 & 1 & -6 \end{array}\right).$$

Получаем ответ

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

Эти соображения позволяют вычислять обратные матрицы. Найдем, например, A^{-1} для матрицы

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

Для этого решаем матричной уравнение $AX = E$, где X — квадратная матрица размера 2, а E — единичная матрица размера 2. Составляем вспомогательную матрицу

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{array}\right).$$

В построенной матрице ко второй строке прибавим первую. Получаем

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right).$$

Далее к первой строке прибавим вторую. Получаем

$$\left(\begin{array}{cc|cc} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right).$$

Итак,

$$A^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

ГЛАВА 4

КОМПЛЕКСНЫЕ ЧИСЛА

Комплексные числа являются одним из основных объектов в современной математике и её приложениях в физике, особенно в квантовой механике, в электродинамике, гидродинамике и т. д. В этой главе приводятся основы теории комплексных чисел: операции над ними, геометрическая интерпретация, тригонометрическая форма, возведение в степень и извлечение корней. Предлагаемый способ введения комплексных чисел важен тем, что аналогичным способом можно задать кватернионы и октанионы, которые играют важную роль в квантовой механике.

§ 1. Действия с комплексными числами

Комплексными числами называются пары вещественных чисел $z = (a, b)$ с операциями сложения и умножения

$$\begin{aligned} (a,b) + (a_1,b_1) &= (a + a_1, b + b_1), \\ (a,b)(a_1,b_1) &= (aa_1 - bb_1, ab_1 + a_1b). \end{aligned} \tag{4.1}$$

Через $\mathbb{C}$ обозначается множество всех комплексных чисел. Отождествим вещественное число a с парой $(a, 0)$. Из (4.1) видно, что при этом отождествлении сумма переходит в сумму, а произведение — в произведение. Таким образом, $\mathbb{R} \subset \mathbb{C}$. В дальнейшем мы будем отождествлять комплексное число $(a, 0)$ с вещественным числом a.

Положим $i = (0,1)$. Из (4.1) следует, что $i^2 = (-1,0) = -1$. Кроме того, $(a,b) = a + bi$. Такое представление комплексного числа называется *алгебраической формой комплексного числа.*

Модулем числа $z = a + bi$ называется неотрицательное вещественное число $|z| = \sqrt{a^2 + b^2}$. Комплексное число $\overline{z} = a - bi$ называется *комплексно сопряженным* к z.

Упражнение. Доказать, что $z\overline{z} = |z|^2$ для любого комплексного числа z.

Сопоставим комплескному числу $z = (a,b)$ вещественную матрицу

$$\Psi(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a, b \in \mathbb{R}. \tag{4.2}$$

Несложно проверяется, что если $a \in \mathbb{R}$, то $\Psi(a) = aE$, где E — единичная матрица размера 2.

Используя определение (4.1) легко проверяется

Предложение 4.1. *Отображение Ψ биективно отображает $\mathbb{C}$ на множество всех вещественных матриц вида* (4.2)*, причем*

$$\Psi(z + z_1) = \Psi(z) + \Psi(z_1), \ \Psi(zz_1) = \Psi(z)\Psi(z_1),$$
$$\Psi(\overline{z}) = {}^t\Psi(z), \ |z| = \sqrt{\det(\Psi(z))}.$$

В частности, для комплексных чисел справедливы следующие равенства 1)–8) *из предложения* 1.5*, свойства* 1)–2) *из предложения* 1.8 *и свойства ассоциативности умножения из предложения* 1.7*. При этом $|zz_1| = |z||z_1|$. Кроме того, умножение комплексных чисел коммутативно и для любого ненулевого комплексного числа z существует обратное число $z^{-1} = \frac{1}{|z|^2}\overline{z}$ такое, что $zz^{-1} = 1$.*

Отметим, что из предложениий 1.10 и 4.1 следует, что

$$\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}, \quad \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}.$$

§ 2. Тригонометрическая форма комплексного числа

Представим комплексное число $z = a+bi$ в виде вектора на плоскости $\mathbb{R}^2$ с координатами a, b. Тогда $|z|$ — это длина этого вектора z. При этом сумме двух комплексных чисел соответствует сумма векторов, сопоставляемых слагаемым.

Упражнение. Доказать, что $|z_1| - |z_2| \leqslant |z_1 + z_2| \leqslant |z_1| + |z_2|$, для любых $z_1, z_2 \in \mathbb{C}$.

Аргументом ненулевого комплексного числа z называется угол φ между z и положительным направлением оси OX. Аргумент ненулевого комплексного числа определен с точностью до слагаемого $2\pi n, n \in \mathbb{Z}$. Если $z = a + bi \neq 0$, то

$$\cos\varphi = \frac{x}{|z|}, \quad \sin\varphi = \frac{y}{|z|}.$$

Получаем *тригонометрическую форму комплексного числа*

$$z = |z|(\cos\varphi + i\sin\varphi). \tag{4.3}$$

Предложение 4.2. *Аргумент произведения комплексных чисел равен сумме аргументов множителей.*

Доказательство. Пусть z из (4.3) и $z_1 = |z_1|(\cos\psi + i\sin\psi)$. По формулам приведения

$$zz_1 = |z||z_1|\,(\cos(\varphi+\psi) + i\sin(\varphi+\psi))\,. \qquad \square$$

Например,

$$(1+i\sqrt{3})(1+i) = 2\left(\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)\cdot\sqrt{2}\left(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right) =$$

$$= 2\sqrt{2}\cdot\left(\cos\frac{\pi}{3} + i\frac{\pi}{3}\right)\left(\cos\frac{\pi}{4} + i\frac{\pi}{4}\right) = 2\sqrt{2}\cdot\left(\cos\frac{7\pi}{12} + i\frac{7\pi}{12}\right).$$

Предложение 4.2 позволяет ввести следующее обозначение. Если $a \in \mathbb{R}$, то $\exp(i\alpha) = \cos\alpha + i\sin\alpha$, и поэтому тригонометрическая форма числа z имеет вид

$$z = |z|\exp(i\varphi). \tag{4.4}$$

Следствие 4.3 (Формула Моавра). *Пусть z из* (4.4) *и $n \in \mathbb{Z}$. Тогда*

$$z^n = |z|^n \exp(ni\varphi).$$

Рассмотрим вопрос об извлечении комплексных корней. Если n — натуральное число, то *корнем n-ой* степени из комплексного числа z называется такое комплексное число t, что $t^n = z$.

Найдем все корни степени n из z. Пусть z из (4.4), и $t = |t|\exp(i\psi)$. Тогда $z = t^n = |t|^n \exp(ni\psi)$. Отсюда $|z| = |t|^n$, т. е. $|t| = \sqrt[n]{|z|}$. Кроме того, $n\psi \equiv \varphi \pmod{2\pi m}$, $m \in \mathbb{Z}$. Следовательно,

$$\psi = \frac{\varphi + 2\pi m}{n}, \quad m = 0, 1, \ldots, n-1.$$

Итак,

$$t = \sqrt[n]{|z|}\left(\cos\frac{\varphi + 2\pi m}{n} + i\sin\frac{\varphi + 2\pi m}{n}\right),\ 0 \leqslant m \leqslant n-1.$$

Поэтому, если $z \neq 0$, то имеется n различных корней степени n из 1. В частности, если $z = 1$, то m-ый корень из 1 имеет вид

$$\varepsilon_m = \cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n}.$$

Все эти корни расположены в вершинах правильного n-угольника, вписанного в единичную окружность с центром в 0. Например, все корни степени 3 из 1 имеют вид

$$\varepsilon_0 = 1, \quad \varepsilon_1 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2},$$
$$\varepsilon_2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

Все корни степени 4 из 1 имеют вид

$$\varepsilon_0 = 1, \quad \varepsilon_1 = \cos\frac{2\pi}{4} + i\sin\frac{2\pi}{4} = i,$$
$$\varepsilon_2 = \cos\frac{4\pi}{4} + i\sin\frac{4\pi}{4} = -1, \quad \varepsilon_3 = \cos\frac{6\pi}{4} + i\sin\frac{6\pi}{4} = -i.$$

ГЛАВА 5

МНОГОЧЛЕНЫ ОТ ОДНОЙ ПЕРЕМЕННОЙ

Эта глава посвящена рассмотрению свойств комплексных многочленов: деление многочленов, нахождение корней, разложение рациональных дробей.

§ 1. Многочлены от одной переменной

Комплексным многочленом называется множество всех формальных конечных сумм

$$f = a_0 + a_1 X + \ldots + a_n X^n, \tag{5.1}$$

где a_i — комплексные числа. Если все коэффициенты a_j являются вещественными числами, то многочлен f называется *вещественным* Множество всех комплексных многочленов обозначается через $\mathbb{C}[X]$, а множество всех вещественных многочленов через $\mathbb{R}[X]$.

Суммой многочленов f из (5.1) и

$$g = b_0 + b_1 X + \ldots \tag{5.2}$$

называется многочлен

$$f + g = (a_0 + b_0) + (a_1 + b_1) X + \ldots,$$

произведением многочленов f из (5.1) и g из (5.2) называется многочлен $fg = c_0 + c_1X + \ldots$, где для любого $k \geqslant 0$

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}\,.$$

Непосредственные вычисления доказывают

Предложение. *Сложение и умножение многочленов ассоциативно и коммутативно, т. е.*

$$f + g = g + f, \quad fg = gf,$$
$$f + (g + h) = (f + g) + h, \quad f(gh) = (fg)h$$

для всех $f, g, h \in C[X]$. *Кроме того, выполнен закон дистрибутивности* $f(g + h) = fg + fh$. *При этом сумма и произведение вещественных многочленов снова вещественны.*

Пусть f из (5.1), причем $a_n \neq 0$. *Степенью* $\deg f$ называется число n. Коэффициент a_n называется *старшим коэффициентом* многочлена f, коэффициент a_0 называется *свободным коэффициентом* многочлена f. Многочлены нулевой степени и нулевой многочлен называются *константами*.

Предложение. *Старший (свободный) коэффициент произведения многочленов равен произведению старших (свободных) коэффициентов сомножителей.*

Доказательство непосредственно вытекает из определения умножения в $\mathbb{C}[X]$.

Следствие 5.1. *Если* $f, g \in \mathbb{C}[X] \setminus 0$, *то* $\deg(fg) = \deg f + \deg g$. *Если* $f, g, f + g \neq 0$, *то*

$$\deg(f + g) \leqslant \max(\deg f, \deg g).$$

§ 2. Деление многочленов

Теорема 5.2 (Деление с остатком). *Пусть* $f, g \in \mathbb{C}[X]$, *причем* $g \neq 0$. *Тогда существуют и притом единственные такие* $q, r \in \mathbb{C}[X]$, *что*

1) $f = qg + r$;

2) $r = 0$ *или* $\deg r < \deg g$.

Если $f, g \in \mathbb{R}[X]$, *то и* $q, r \in \mathbb{R}[X]$.

Доказательство.

Существование. Если $f = 0$ или $\deg f < \deg g$, то полагаем $q = 0, r = g$. Пусть $\deg f = n$, и для меньших степеней теорема доказана. Предположим, что a_n, b_m — старшие коэффициенты f, g. Тогда степень $h = f - \frac{a_n}{b_m}X^{n-m}g$ меньше n. По индукции $h = q'g + r'$, откуда

$$f = h + \frac{a_n}{b_m}X^{n-m}g = (q' + \frac{a_n}{b_m}X^{n-m})g + r'.$$

Единственность. Пусть

$$f = qg + r = q'g + r',$$

где $\deg r, \deg r' < \deg g$, если они ненулевые. Тогда $g(q - q') = r' - r$. Если $q - q' \neq 0$, то $r - r' \neq 0$. Отсюда по следствию 5.1

$$\deg g > \deg(r - r') = \deg(g(q - q')) = \deg g + \deg(q - q') \geqslant \deg g.$$

$\square$

Многочлен q из условия теоремы 5.2 называется *частным*, а многочлен r — *остатком* f при делении на g. Многочлен g *делит* f, если $r = 0$; в этом случае пишут $g \mid f$.

Наибольшим общим делителем многочленов $f_1, \ldots, f_m$, не все из которых равны нулю, называется такой многочлен d, что

1) $d \mid f_i$, $i = 1, \ldots, m$;

2) если $d' \in \mathbb{C}[X]$, и $d' \mid f_i$, $i = 1, \ldots, m$, то $d' \mid d$.

Наибольший общий делитель многочленов $f_1, \ldots, f_m$, обозначается либо $(f_1, \ldots, f_m)$.

Предложение 5.3. *Наибольший общий делитель многочленов $(f_1, \ldots, f_m)$ определен однозначно, с точностью до множителя нулевой степени (ненулевой константы). Кроме того,*

$$(f_1, \ldots, f_m) = (f_1, (f_2, \ldots, f_m)).$$

Изложим *алгоритм Евклида* нахождения наибольшего общего делителя двух многочленов f, g, $g \neq 0$. Будем последовательно делить с остатком:

$$\begin{aligned}
f &= q_1 g + r_1, & \deg r_1 &< \deg g; \\
g &= q_2 r_1 + r_2, & \deg r_2 &< \deg r_1; \\
r_1 &= q_3 r_2 + r_3, & \deg r_3 &< \deg r_2; \\
&\ldots\ldots\ldots\ldots\ldots\ldots & &\ldots\ldots\ldots\ldots\ldots\ldots \\
r_{k-1} &= q_{k+1} r_k + r_{k+1}, & \deg r_{k+1} &< \deg r_k; \\
r_k &= q_{k+2} r_{k+1}.
\end{aligned} \tag{5.3}$$

Отметим, что в (5.3) число k существует, поскольку степени остатков $r_1, r_2, \ldots$ убывают.

Теорема 5.4. $(f, g) = r_{k+1}$.

Доказательство. Из последнего равенства в (5.3) вытекает, что $r_{k+1} \mid r_k$ и т. д. Поднимаясь вверх получаем, что $r_{k+1} \mid g$, $r_{k+1} \mid f$.

Обратно, если $d' \mid f$, $d' \mid g$, то из первого равенства в (5.3) получаем, что $d' \mid r_1$, а из второго, что $d' \mid r_2$. Продвигаясь вниз, получаем, что $d' \mid r_{k+1}$. □

Следствие. *Если f, g — вещественные многочлены, то и их наибольший общий делитель со старшим коэффициентом 1 является вещественным многочленом.*

Пусть $d = (f_1, \ldots, f_m)$. Тогда существуют такие многочлены $u_1, \ldots, u_m$, что $d = u_1 f_1 + \ldots + u_m f_m$.

Доказательство. В силу предложения 5.3 можно считать, что $m = 2$, и $f_1 = f$, $f_2 = g$. Остается воспользоваться алгоритмом (5.3), где $d = r_{k+2}$. □

Многочлены $f_1, \ldots, f_m$ *взаимно просты*, если их наибольший общий делитель равен 1.

Следующее предложение вытекает из предыдущего следствия.

Предложение 5.5. *Многочлены $f_1, \ldots, f_m$ взаимно просты тогда и только тогда, когда существуют такие многочлены $u_1, \ldots, u_m$, что $1 = u_1 f_1 + \ldots + u_m f_m$.*

§3. Корни многочленов

Пусть c — комплексное число и f — многочлен из (5.1). Положим $f(c) = a_0 + a_1 c + \ldots + a_n c^n$. Элемент c является *корнем* f, если $f(c) = 0$.

Теорема 5.6 (Теорема Безу). *Пусть c — комплексное число и $f \in \mathbb{C}[X]$. Тогда остаток от деления f на $X - c$ равен $f(c)$. В частности, c является корнем f тогда и только тогда, когда $(X - c) \mid f$.*

Доказательство. Деля f с остатком на $X - c$ получаем, что $f = g(X-c)+r$, где r — комплексное число. Отсюда $f(x) = r$. □

Отметим, что если c — комплексное число, то деля любой многочлен f последовательно с остатком на $X - c$, получаем для f разложение Тейлора

$$f = b_0 + b_1(X - c) + \ldots + b_n(X - c)^n. \tag{5.4}$$

Изложим *схему Горнера* для быстрого вычисления коэффициентов b_i в (5.4) по коэффициентам многочлена f. Для этого необходимо совершить один шаг деления на $X - c$. Пусть f из (5.1) и

$$f = g(X - c) + r, \quad r \in \mathbb{C}, \tag{5.5}$$

где $g = s_0 + s_1X + \ldots + s_{n-1}X^{n-1}$. Подставляя в (5.5) получаем

$$a_0 + a_1X + \ldots + a_nX^n = (X - c)(s_0 + s_1X + \ldots + s_{n-1}X^{n-1}) + r.$$

Приравнивая коэффициенты при одинаковых степенях X, получаем

$$\begin{aligned} a_n &= s_{n-1} \\ a_{n-1} &= s_{n-2} - cs_{n-1} \\ &\cdots\cdots\cdots\cdots \\ a_1 &= s_0 - cs_1 \\ a_0 &= r - cs_0 \quad . \end{aligned}$$

Отсюда

$$\begin{aligned} s_{n-1} &= a_n \\ s_{n-2} &= a_{n-1} + cs_{n-1} \\ &\cdots\cdots\cdots\cdots \\ s_0 &= a_1 + cs_1 \\ r &= a_0 + cs_0 \quad . \end{aligned} \tag{5.6}$$

Формулы (5.6) позволяют быстро за n умножений вычислить $r = f(c)$. Результаты этих вычислений обычно записываются в виде таблицы

	a_n	a_{n-1}	$\ldots$	a_1	a_0
c	$a_0 = s_{n-1}$	s_{n-2}	$\ldots$	s_1	r

. (5.7)

Деля далее g с остатком на $X - c$ получаем, получаем формулу Тейлора. Результаты этих вычислений запишем в таблицу

	a_n	a_{n-1}	$\ldots$	a_1	a_0
c	$a_n = s_{n-1}$	s_{n-2}	$\ldots$	s_1	$r = b_0$
c	$a_n = t_{n-2}$	t_{n-3}	$\ldots$	$t_0 = b_1$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
c	$a_n = u_1$	$u_0 = b_{n-1}$			
c	$a_n = b_n$				

,

где b_i из (5.4).

Например, напишем разложение многочлена $f = X^3 - 3X + 2$ по степеням $x + 2$. Заполняем первую строку таблицы коэффициентами многочлена f. Получаем

	1	0	−3	2
−2				

.

Затем заполняем вторую строку в соответствии с формулами (5.6). Получаем

	1	0	−3	2
−2	1	−2	1	0

.

Далее заполняем остальные строки,

	1	0	−3	2
−2	1	−2	1	0
−2	1	−4	9	
−2	1	−6		
−2	1			

.

Получаем формулу Тейлора $f = (X+2)^3 - 6(X+2)^2 + 9(X+2)$.

Кратностью корня c многочлена f называется кратность неприводимого множителя $X - c$ в f. Другими словами, кратность равна k, если $(X-c)^k \mid f$, но $(X-c)^{k+1} \nmid f$.

Это определение согласовано с теоремой Безу 5.6.

Теорема 5.7. *Пусть f — комплексный многочлен ненулевой степени. Если k — кратность корня c в f, то кратность c в f' равна $k - 1$.*

Доказательство. Имеем $f = (X-c)^k g$, где $g(c) \neq 0$. Тогда

$$\begin{aligned} f' &= k(X-c)^{k-1}g + (X-c)^k g' = \\ &= (X-c)^{k-1}(kg + (X-c)g'). \end{aligned} \tag{5.8}$$

Следовательно, $(X-c)^{k-1} \mid f'$. Если $(X-c)^k \mid f'$, то

$$(X-c) \mid \left(kg + (X-c)g'\right),$$

в силу (5.8), и поэтому $(X-c) \mid kg$, что невозможно. □

Следствие 5.8. *Пусть задан ненулевой многочлен f. Тогда корнями $d = (f, f')$ являются только кратные корни f. Их кратности в d на 1 меньше, чем в f.*

Упражнение. Следующие условия эквивалентны:

1) кратность корня c многочлена f равна k;

2) в формуле Тейлора

$$f = b_k(X-c)^k + b_{k+1}(X-c)^{k+1} + \ldots + b_n(X-c)^n, \quad b_k \neq 0.$$

Если $c_1, \ldots, c_m$ — различные корни многочлена f с кратностями $k_1, \ldots, k_m$, то

$$f = (X-c_1)^{k_1} \cdots (X-c_m)^{k_m} g,$$

где многочлен g не имеет корней. Поэтому справедливо

Предложение 5.9. *Сумма числа корней (с кратностями) не превосходит степени многочлена.*

Упражнение 5.10. Элемент c является кратным корнем (корнем кратности $\geqslant 2$) тогда и только тогда, когда c общий корень многочлена и его производной.

§ 4. Интерполяция

Пусть заданы различные комплексные числа $a_0, a_1, \ldots, a_n$, и числа $b_0, b_1, \ldots, b_n$. Рассмотрим *интерполяционный многочлен Лагранжа*

$$L = \sum_{i=0}^{n} b_i \frac{(X-a_0)\ldots(X-a_{i-1})(X-a_{i+1})\ldots(X-a_n)}{(a_i-a_0)\ldots(a_i-a_{i-1})(a_i-a_{i+1})\ldots(a_i-a_n)}.$$

Теорема. $\deg L \leqslant n$ *и* $L(a_j) = b_j$ *для всех* $j = 0, \ldots, n$.

Доказательство.

$$L(a_j) = \sum_{i=0}^{n} b_i \frac{(a_j-a_0)\ldots(a_j-a_{i-1})(a_j-a_{i+1})\ldots(a_j-a_n)}{(a_i-a_0)\ldots(a_i-a_{i-1})(a_i-a_{i+1})\ldots(a_i-a_n)} = b_j.$$

□

Теорема. *Существует в точности один многочлен* L *степени не выше* n *с условием* $L(a_j) = b_j$ *для всех* $j = 0, \ldots, n$.

Доказательство. Предположим, что найдутся два многочлена L_1, L_2 с этим условием. Тогда $G = L_1 - L_2 \neq 0$, причем $\deg G \leqslant n$, и элементы $a_0, a_1, \ldots, a_n$ является его корнями. Это противоречит предложению 5.9. □

§ 5. Корни многочленов над $\mathbb{C}$ и $\mathbb{R}$

Теорема 5.11 (Основная теорема алгебры). *Комплексный многочлен положительной степени от одной переменной имеет комплексный корень.*

Следствие 5.12. *В частности, каждый многочлен* $f \in \mathbb{C}[X]$ *однозначно представляется в виде*

$$f = a(X-c_1)^{k_1}\ldots(X-c_m)^{k_m}, \quad a \in \mathbb{C},$$

где $c_1, \ldots, c_m$ *различные корни* f *кратностей* $k_1, \ldots, k_m$.

Доказательство. Применить теоремы 5.11, 5.6. □

Предложение 5.13. *Пусть* $f \in \mathbb{R}[X]$ *и* $c \in \mathbb{C}$ — *корень* f. *Тогда* $\overline{c}$ *также корень* f.

Доказательство. Пусть f имеет вид (5.1). Тогда

$$f(\overline{c}) = a_0 + a_1\overline{c} + \ldots + a_n\overline{c}^n = \overline{a_0} + \overline{a_1c} + \ldots + \overline{a_n}\overline{c^n} =$$

$$= \overline{a_0 + a_1c + \ldots + a_nc^n} = \overline{f(c)} = 0.$$ □

Упражнение. Пусть $f \in \mathbb{R}[X]$ и $c \in \mathbb{C}$ — корень f кратности k. Тогда $\overline{c}$ также корень f кратности k.

Теорема 5.14. *Каждый вещественный многочлен* f *однозначно представляется в виде произведения ненулевой константы* a, *множителей* $(X - c_j)^{k_j}$ *и множителей* $(X^2 + p_kX + q_k)^{l_k}$, *где* $a \in \mathbb{R}$ *и* c_j *различные вещественные корни* f *кратности* k_j, *а многочлены* $(X^2 + p_kX + q_k)^{l_k} \in \mathbb{R}[X]$ *различны и не имеют вещественных корней. В частности, вещественный многочлен нечетной степени имеет вещественный корень.*

Доказательство. Вещественный многочлен f над комплексными числами разлагается в произведение многочленов $X - c_j$, где c_j — его комплексные корни по теореме Безу 5.6. В силу предложения 5.13 среди множителей первой степени встречается и множитель $X - \overline{c}_j$, причем

$$g = (X - c_j)(X - \overline{c}_j) = X^2 - (c_j + \overline{c}_j) + |c_j|^2.$$

Заметим, что если $c_j = a + bi$, где $a, b \in \mathbb{R}$, то $c_j + \overline{c}_j = (a + bi) + (a - bi) = 2a \in \mathbb{R}$. Таким образом, g является вещественным многочленом, не имеющим вещественных корней, если $c_j \in \mathbb{C} \setminus \mathbb{R}$. Отсюда вытекают все остальные утверждения □

Приведем оценку модулей корней комплексного многочлена.

Теорема. *Пусть* $f \in \mathbb{C}[X]$ *имеет вид* (5.1). *Предположим, что* $a_{n-1} = \ldots = a_{k+1} = 0$, *и* $a_n, a_k \neq 0$, *где* $0 \leqslant k < n$. *Положим* $B = \max(|a_k|, |a_{k-1}|, \ldots, |a_0|)$. *Если* $z \in \mathbb{C}$ — *корень* f, *то*

$$|z| < 1 + \sqrt[n-k]{\frac{B}{|a_n|}}.$$

Доказательство. Пусть

$$|z| \geqslant 1 + \sqrt[n-k]{\frac{B}{|a_n|}}.$$

Достаточно показать, что $|f(z)| > 0$. Действительно,

$$\begin{aligned}
|f(z)| &= |a_n z^n + a_k z^k + \ldots + a_1 z + a_0| \geqslant \\
&\geqslant |a_n z^n| - |a_k z^k + \ldots + a_1 z + a_0| \geqslant \\
&\geqslant |a_n||z|^n - \sum_{j=0}^{k} |a_j||z|^j \geqslant |a_n||z|^n - B\sum_{j=0}^{k} |z|^j = \\
&= |a_n||z|^n - B\frac{|z|^{k+1} - 1}{|z| - 1} = \\
&= \frac{|a_n|}{|z| - 1}\left[|z|^n(|z| - 1) - \frac{B}{|a_n|}|z|^{k+1} + \frac{B}{|a_n|}\right] > \\
&> \frac{|a_n|}{|z| - 1}\left[|z|^n(|z| - 1) - \frac{B}{|a_n|}|z|^{k+1}\right] = \\
&= \frac{|a_n||z|^{k+1}}{|z| - 1}\left[|z|^{n-k-1}(|z| - 1) - \frac{B}{|a_n|}\right].
\end{aligned}$$

По условию $|z| - 1 \geqslant \alpha = \sqrt[n-k]{\dfrac{B}{|a_n|}} > 0$. Так как $1 + \alpha > \alpha$, то

$$\begin{aligned}
|z|^{n-k-1}(|z| - 1) - \frac{B}{|a_n|} &\geqslant (1 + \alpha)^{n-k-1}\alpha - \alpha^{n-k} = \\
&= \alpha[(1 + \alpha)^{n-k-1} - \alpha^{n-k-1}] > 0.
\end{aligned}$$

□

Приведем алгоритм Штурма приближенного нахождения вещественных корней многочленов с вещественными коэффициентами. Заметим, что если $f \in \mathbb{R}[X]$, то по следствию 5.8 многочлен $\frac{f}{(f,f')} \in \mathbb{R}[X]$ имеет те же корни, что и f, но их кратности становятся равными 1, т. е. $\frac{f}{(f,f')}$ не имеет кратных корней. Поэтому необходимо уметь вычислять корни многочленов, не имеющих кратных корней.

Теорема (Штурм). *Пусть $f \in \mathbb{R}[X]$ не имеет кратных корней. Построим последовательность Штурма $f_0 = f$, и $f_1 = f'$. Далее каждое для любого $i \geqslant 2$ через f_i обозначим остаток со знаком «—» от деления f_{i-2} на f_{i-1}, т. е.*

$$f_{i-2} = g_{i-1} f_{i-1} - f_i. \tag{5.9}$$

Пусть $u < v$ — вещественные числа, не являющиеся корнями f. Для любого вещественного числа a через $W(a)$ обозначим число перемен знаков в ряду

$$f_0(a), f_1(a), \ldots, f_n(a), \quad n = \deg f. \tag{5.10}$$

Тогда число вещественных корней в $[u, v]$ равно $W(u) - W(v)$.

Доказательство. Нам потребуется ряд лемм.

Лемма 5.15. *Если $f_i(c) = 0$, то $f_{i+1}(c) \neq 0$. Кроме того, если $i > 0$, то $f_{i-1}(c) = -f_{i+1}(c)$.*

Доказательство. Если бы $f_i(c) = f_{i+1}(c) = 0$, то по (5.9) $f(c) = f'(c) = 0$, т. е. c — кратный корень f, что невозможно в силу теоремы 5.7 (см. также упражнение 5.10).

Второе утверждение вытекает из (5.9). □

Лемма 5.16. *Пусть $f_i(c) = 0$, где $i > 0$. Тогда существует такое $\varepsilon > 0$, что $f_{i-1}(x), f_{i+1}(x)$ имеют разные знаки при $x \in (c - \varepsilon, c + \varepsilon)$.*

Доказательство. Воспользоваться непрерывностью многочленов f_j для всех j. □

Завершим доказательство теоремы. Пусть $c \in (u, v)$ не является корнем f. Рассмотрим поведение функции $W(x)$ в окрестности c. Если все $f_j(c) \neq 0$, то в некоторой окрестности c все $f_j(c)$ имеют постоянный знак. Поэтому $W(x)$ постоянно в этой окрестности c. Если же некоторое $f_j(c) = 0, j \geqslant 1$, то в некоторой окрестности c по лемме 5.16 функции $f_{j-1}(x), f_{j+1}(x)$ имеют постоянный знак, и поэтому число перемен знаков в ряду

$$f_{j-1}(x), \quad f_j(x), \quad f_{j+1}(x) \tag{5.11}$$

постоянно. Беря пересечение этих окрестностей c для всех j, получаем окрестность c, в которой $W(x)$ постоянно.

Пусть $f(c) = 0$. По лемме 5.15 имеем $f_1(c) \neq 0$. Пусть $f_1(c) > 0$. Тогда в некоторой окрестности U точки c функция $f(x)$ возрастает. Пусть $f(x) < 0, f(y) > 0$, при

$$x, y \in U, \quad x < c < y. \tag{5.12}$$

Если $f_j(c) \neq 0$, то можно считать, что $f_j(x)$ имеет постоянный знак в U. Если $f_j(c) = 0, j > 1$, то можно считать, что в ряду (5.11) число перемен знаков постоянно. Следовательно, для числа $W(a)$ перемен знаков в ряду (5.10) получаем $W(x) - W(y) = 1$, если x, y из (5.12).

Пусть $f_1(c) < 0$. Тогда в некоторой окрестности U точки c функция $f(x)$ убывает. Пусть $f(x) > 0, f(y) < 0$, при

$$x, y \in U, \quad x < c < y. \tag{5.13}$$

Если $f_j(c) \neq 0$, то можно считать, что $f_j(x)$ имеет постоянный знак в U. Если $f_j(c) = 0, j > 1$, то можно считать, что в ряду (5.11) число перемен знаков постоянно. Следовательно, для числа $W(a)$ перемен знаков в ряду (5.10) получаем $W(x) - W(y) = 1$, если x, y из (5.13). □

§ 6. Рациональные дроби

Рассмотрим множество отношений $\frac{a}{b}$, где a, b — комплексные (вещественные) многочлены от одной переменной, причем $b \neq 0$. Скажем, что два отношения $\frac{a}{b}$, $\frac{c}{d}$ равны, если $ad = bc$.

Предложение. *Это отношение является отношением эквивалентности.*

Доказательство. Прямая проверка показывает, что если $\frac{a}{b} = \frac{c}{d}$ и $\frac{c}{d} = \frac{u}{v}$, то $\frac{a}{b} = \frac{u}{v}$. Кроме того, если $\frac{a}{b} = \frac{c}{d}$, то $\frac{c}{d} = \frac{a}{b}$, и $\frac{a}{b} = \frac{a}{b}$. □

Упражнение. Доказать, что следующие дроби равны $\frac{ab}{ac} = \frac{b}{c}$.

Через $\mathbb{C}(X)$ (соответственно, через $\mathbb{R}(X)$) обозначим множество классов равных комплексных (вещественных) дробей.

Введем в $\mathbb{C}(X)$ (в $\mathbb{R}(X)$) сложение и умножение по правилу

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \quad \frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

Предложение. *Определение сложения и умножения корректно. При этом сложение и умножение коммутативны, ассоциативны, выполнены законы дистрибутивности, и у каждого ненулевой дроби есть обратная.*

Доказательство состоит в непосредственном вычислении.

Предложение. *Зададим отображение $\varphi : \mathbb{C}[X] \to \mathbb{C}(X)$ по правилу $\varphi(a) = \frac{a}{1}$. Тогда φ инъективно, и $\varphi(a+b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$.*

Итак, можно считать, что $\mathbb{C}[X] \subset \mathbb{C}(X)$ и $\mathbb{R}[X] \subset \mathbb{R}(X)$.

Множество $\mathbb{C}(X)$ (соответственно, $\mathbb{R}(X)$) называется *полем рациональных комплексных функций* (соответственно, *полем рациональных вещественных функций*). *Правильной дробью* называется дробь $\frac{f}{g} \in \mathbb{C}(X)$ (или $\mathbb{R}(X)$), где $\deg f < \deg g$. Комплексная дробь вида $\frac{a}{(X-c)^k} \in \mathbb{C}(X)$, где $a, c \in \mathbb{C}$, $a \neq 0$, $k \geqslant 1$, называется *простейшей комплексной дробью*. Вещественная дробь называется *простейшей вещественной дробью*, если она имеет вид либо $\frac{a}{(X-c)^k} \in \mathbb{R}(X)$, где $a, c \in \mathbb{R}$, $a \neq 0$, $k \geqslant 1$, либо $\frac{aX+b}{(X^2+pX+q)^k} \in \mathbb{R}(X)$, где $a, b, p, q \in \mathbb{R}$, $ax + b \neq 0$, $k \geqslant 1$, причем $X^2 + pX + q$ не имеет вещественных корней.

Теорема. *Любая комплексная (вещественная) рациональная дробь является суммой многочлена и простейших дробей.*

Доказательство. Нам потребуется

Лемма. *Пусть $\frac{f}{g}$ либо комплексная, либо вещественная дробь, и g разложено в произведение $g = uv$ двух комплексных или вещественных взаимно простых многочленов u, v. Тогда*

$$\frac{f}{g} = \frac{h}{u} + \frac{t}{v},$$

где h, t — комплексные (вещественные) многочлены.

Доказательство. Имеем $ua + vb = 1$ для некоторых многочленов a, b. Отсюда $f = uaf + vbf$, и

$$\frac{f}{g} = \frac{uaf}{uv} + \frac{vbf}{uv} = \frac{af}{v} + \frac{bf}{u}. \qquad \square$$

В силу следствия 5.12 и теоремы 5.14 можно считать, что дробь имеет вид $\frac{f}{p^k}$, где p — многочлен либо первой степени, либо вещественный многочлен второй степени, не имеющий вещественных корней. Покажем утверждение теоремы индукцией по $\deg f$. Если $\deg f < \deg p$, то эта дробь простейшая. Пусть $\deg f \geqslant \deg p$. Разделив f с остатком на p, получаем $f = qp + r$, причем либо $r = 0$, либо $\deg r < \deg p$. Тогда

$$\frac{f}{p^k} = \frac{qp + r}{p^k} = \frac{q}{p^{k-1}} + \frac{r}{p^k}.$$

Второе слагаемое является правильной дробью, а в первом степень числителя уменьшилась. Продолжая этот процесс получаем утверждение теоремы. $\square$

Найдем, например, представление правильной комплексной дроби

$$\frac{i - X}{(X^2 + 1)(X - 1)^2}$$

в виде суммы простейших дробей. Для этого разложим знаменатель X^2+1 на множители первой степени. Получаем

$$(X^2+1)(X-1)=(X-i)(X+i)(X-1).$$

Тогда

$$\frac{i-X}{(X^2+1)(X-1)^2}=\frac{A}{X+i}+\frac{B}{X-i}+\frac{C}{X-1}+\frac{D}{(X-1)^2}, \tag{5.14}$$

где A,B,C,D — комплексные числа. Приводя правую часть в (5.14) к общему знаменателю, равному $(X^2+1)(X-1)^2$, и приравнивая числители, получаем

$$\begin{aligned} i-X=A(X-i)(X-1)^2+B(X+i)(X-1)^2+\\ +C(X^2+1)(X-1)+D(X^2+1). \end{aligned} \tag{5.15}$$

Полагая последовательно $X=i$, $X=-i$, $X=1$, $X=0$ в (5.15), получаем

$$\begin{aligned} 0&=2iB(i-1)^2, & 2i&=-2iA(i+1)^2,\\ i-1&=2D, & i&=-iA+iB-C+D. \end{aligned}$$

Поэтому

$$A=-\frac{1}{(i+1)^2}=-\frac{1}{2i}=\frac{i}{2}, \quad B=0, \quad D=\frac{i-1}{2},$$
$$C=-i-iA+iB+D=-i+\frac{1}{2}+\frac{i-1}{2}=\frac{-2i+1+i-1}{2}=-\frac{i}{2}.$$

ГЛАВА 6

ЛИНЕЙНЫЕ ПРОСТРАНСТВА. РАНГ МАТРИЦЫ

В главе излагаются общие свойства линейных пространств. В терминах линейной зависимости можно определить понятие ранга матрицы, которое позволяет найти число свободных и главных неизвестных в произвольной системе линейных уравнений. Кроме того, в этой главе изучаются плоскости любой размерности в линейных пространствах и различные способы их задания. Показывается, что плоскости и только они являются решениями совместных систем линейных уравнений.

§ 1. Линейные пространства

Линейным или *векторным пространством* называется множество V, в котором определена операция сложения $x+y$ элементов из V, называемых *векторами*, и операция αx умножения вектора x на число α. При этом выполнены следующие аксиомы:

1) сложение ассоциативно, т. е. $(x+y)+z = x+(y+z)$ для всех $x, y, z \in V$;

2) сложение коммутативно, т. е. $x+y = y+x$ для всех $x, y \in V$;

3) существует такой элемент $0 \in V$, называемый *нулем*, что $x+0 = x$ для всех $x \in V$;

4) для любого $x \in V$ существует такой элемент $-x \in V$, называемый *противоположным* к x, что $x + (-x) = 0$;

5) если α, β — числа и $x \in V$, то $(\alpha\beta)x = \alpha(\beta x)$;

6) если α, β — числа и $x \in V$, то $(\alpha + \beta)x = \alpha x + \beta x$;

7) если α — число и $x, y \in V$, то $\alpha(x + y) = \alpha x + \alpha y$;

8) если $x \in V$, то $1x = x$.

Примеры. Примерами пространств являются прямоугольные матрицы $\mathrm{Mat}(n \times m)$, векторы плоскости $\mathbb{R}^2$, пространства $\mathbb{R}^3$, все функции на отрезке $[a, b]$ и т. д.

Предложение. *Выполняются следующие утверждения*:

1) *нулевой элемент единствен*;

2) *противоположный элемент определен однозначно*;

3) $0x = \alpha 0 = 0$;

4) $(-1)x = -x$.

Доказательство. Проверим, например, третье утверждение. По аксиомам линейного пространства $0x = (0 + 0)x = 0x + 0x$. Отсюда

$$\begin{aligned} 0 = 0x + (-0x) = (0x + 0x) + (-0x) = \\ = 0x + (0x + (-0x)) = 0x + 0 = 0x. \end{aligned} \qquad \square$$

Следствие. $(\alpha - \beta)x = \alpha x - \beta x, \quad \alpha(x - y) = \alpha x - \alpha y$.

Система векторов $x_1, \ldots, x_n$ линейного пространства *линейно зависима*, если

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0. \tag{6.1}$$

для некоторого ненулевого набора чисел $\alpha_1, \ldots, \alpha_n$ (т. е. не все эти числа равны нулю). Система векторов $x_1, \ldots, x_n$ линейного пространства *линейно независима*, если для любых чисел $\alpha_1, \ldots, \alpha_n$ из равенства (6.1) вытекает, что $\alpha_1 = \ldots = \alpha_n = 0$.

Ясно, что любая конечная система векторов линейного пространства либо линейно независима, либо линейно зависима.

Вектор x является *линейно комбинацией* системы векторов $x_1, \dots, x_n$ если существует такой набор чисел $\alpha_1, \dots, \alpha_n$, что

$$x = \alpha_1 x_1 + \dots + \alpha_n x_n.$$

Линейной оболочкой $\langle x_1, \dots, x_n \rangle$ набора векторов $x_1, \dots, x_n$ называется множество всех линейных комбинаций этой системы векторов. Несложно проверяется следующее

Предложение. *Если в системе векторов есть нулевой вектор, то система зависима. Если подсистема в системе линейно зависима, то и вся система зависима. Система зависима тогда и только тогда, когда один из её векторов является линейно комбинацией остальных.*

Теорема 6.1 (Критерий равенства определителя нулю). *Определитель квадратной матрицы A равен нулю тогда и только тогда, когда её строки (столбцы) линейно зависимы.*

Доказательство. Пусть столбцы A зависимы, и, например, последний столбец A является линейной комбинацией предыдущих. Вычтем из последнего столбца эту линейную комбинацию. Тогда определитель не меняется, но, с другой стороны, он равен 0.

Обратно, пусть $\det A = 0$. Система линейных однородных уравнений с матрицей ${}^tA$ совместна и не имеет единственного решения по теореме Крамера 3.16. Поэтому у нее есть ненулевой столбец решений Λ, откуда ${}^tA\Lambda = 0$. Это дает зависимость столбцов матрицы A с коэффициентами из столбца Λ. □

Теорема 6.2 (Основная лемма о линейной зависимости). *Предположим, что заданы две системы векторов*

$$\{a_1, \dots, a_n\} \text{ и } \{b_1, \dots, b_m\}, \tag{6.2}$$

причем каждый вектор $a_i \in \langle b_1, \dots, b_m \rangle$. Если $n > m$, то система $a_1, \dots, a_n$ линейно зависима.

Доказательство. По условию,

$$\begin{array}{lcl} a_1 & = & b_1\alpha_{11} + \ldots + b_m\alpha_{m1} \\ \multicolumn{3}{c}{\dotfill} \\ a_n & = & b_1\alpha_{1n} + \ldots + b_m\alpha_{mn} \end{array} \quad . \tag{6.3}$$

Рассмотрим однородную систему линейных уравнений

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \ldots + \alpha_{mn}x_n = 0 \end{cases} . \tag{6.4}$$

По предложению 1.4 у системы (6.4) имеется ненулевое решение $c_1, \ldots, c_n$. Из (6.3) и (6.4) следует, что $c_1a_1 + \ldots + c_na_n = 0$. □

Следствие 6.3. *Заданы две системы векторов* (6.2), *причем $a_i \in \langle b_1, \ldots, b_m \rangle$ для любого $i = 1, \ldots, n$. Если система $a_1, \ldots, a_n$ независима, то $n \leqslant m$. Если $\langle a_1, \ldots, a_n \rangle = \langle b_1, \ldots, b_m \rangle$ и обе системы независимы, то $n = m$.*

Базисом линейного пространства L называется система векторов $a_1, \ldots, a_n \in L$ со следующими свойствами:

1) векторы $a_1, \ldots, a_n$ независимы;

2) $L = \langle a_1, \ldots, a_n \rangle$.

Размерностью $\dim V$ пространства V называется число векторов в базисе V. Векторное пространство конечномерно, если его размерность конечна.

В силу следствия 6.3 число векторов в любом базисе постоянно, и поэтому размерность пространства определена однозначно.

Упражнение. Доказать, что матричные единицы E_{ij} образуют базис $\mathrm{Mat}(n \times m)$.

Теорема 6.4. *Любую независимую систему векторов конечномерного пространства можно дополнить до базиса.*

Доказательство. Пусть

$$\mathbf{e} = (e_1, \ldots, e_n) \tag{6.5}$$

— базис векторного пространства V, Пусть задана независимая система векторов $a_1, \ldots, a_k$. Если система векторов $a_1, \ldots, a_k, e_1$ зависима, то $e_1 \in \langle a_1, \ldots, a_n \rangle$. В этом случае переходим к рассмотрению вектора e_2. Если же эта система независима, первоначальную систему $a_1, \ldots, a_k$ заменяем на систему $a_1, \ldots, a_k, a_{k+1} = e_1$. Продолжая этот процесс, получаем такую независимую систему векторов

$$a_1, \ldots, a_k, e_{i_1}, \ldots, e_{i_s},$$

линейная оболочка L которой содержит базис $\mathbf{e}$ из (6.5). Отсюда $L = V$, и построенная система является базисом. □

Пусть $\mathbf{e}$ — базис векторного пространства V из (6.5), и $x \in V$. Тогда

$$x = x_1 e_1 + \ldots + x_n e_n = \mathbf{e}X, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

— разложение вектора x по базису $\mathbf{e}$. Набор X называется *системой (столбцом) координат* вектора x в базисе $\mathbf{e}$.

Предложение. *Система координат вектора в базисе определена однозначно.*

Доказательство. Пусть $\mathbf{e}$ — базис пространства из (6.5) и

$$x = x_1 e_1 + \ldots + x_n e_n = y_1 e_1 + \ldots + y_n e_n\,.$$

Тогда

$$0 = (x_1 - y_1)e_1 + \ldots + (x_n - y_n)e_n\,,$$

откуда

$$x_1 - y_1 = \ldots = x_n - y_n = 0$$

в силу независимости векторов $\mathbf{e}$. □

Предположим, что в векторном пространстве L выбраны два базиса $\mathbf{e}$ из (6.5) и $\mathbf{e}' = (e'_1, \dots, e'_n)$. Тогда для любых $i, j = 1, \dots$ $\dots, n$ имеем

$$e'_i = e_1 c_{1i} + \dots + e_n c_{ni}, \quad 1 \leqslant i \leqslant n,$$
$$e_j = e'_1 c'_{1j} + \dots + e'_n c'_{nj}, \quad 1 \leqslant j \leqslant n,$$

или

$$\mathbf{e}' = \mathbf{e}C, \quad \mathbf{e} = \mathbf{e}'C', \text{ где } C = (c_{ij}), \quad C' = (c'_{ij}) \in \mathrm{Mat}(n). \quad (6.6)$$

Матрица C (соответственно, C') из (6.6) называется *матрицей перехода* от $\mathbf{e}$ к $\mathbf{e}'$ (соответственно, от $\mathbf{e}'$ к $\mathbf{e}$).

Из (6.6) вытекает, что $\mathbf{e} = \mathbf{e}CC'$, и поэтому $CC' = E_n$. Аналогично, $C'C = E_n$. Поэтому $C' = C^{-1}$. Обратно, если у матрицы C есть обратная, и $\mathbf{e}' = \mathbf{e}C$, то $\mathbf{e}'$ — базис пространства. Поэтому справедливо

Предложение. *Матрицами перехода от одного базиса к другому являются обратимые матрицы и только они.*

Предложение 6.5. *Пусть* $\mathbf{e}, \mathbf{e}'$ *— два базиса пространства* L, *и вектор* $x \in L$ *имеет в этих базисах столбцы координат* X, X', *соответственно. Если* C *— матрица перехода от* $\mathbf{e}$ *к* $\mathbf{e}'$, *то* $X = CX'$.

Доказательство. В силу (6.6) $x = \mathbf{e}'X' = \mathbf{e}CX' = \mathbf{e}X$, откуда вытекает утверждение в силу единственности разложения по базису. □

Непустое подмножество L в линейном пространстве V называется *подпространством*, если из того, что $x, y \in L$ следует, что $x + y, \alpha x \in L$.

Упражнение. Доказать следующие утверждения:

1) Если L — подпространство в V, то $0 \in L$;

2) Если $x \in L$, то $-x \in L$;

3) Если $a_1, \ldots, a_k$ — векторы некоторого линейного пространства L, то линейная оболочка $\langle a_1, \ldots, a_k \rangle$ является подпространством L.

Пример 6.6. Рассмотрим однородную систему линейных уравнений $AX = 0$, $A \in \mathrm{Mat}(m \times n)$. Тогда все её решения образуют подпространство в пространстве столбцов $\mathrm{Mat}(n \times 1)$.

Теорема 6.7. *Пусть L — подпространство конечномерного пространства V. Тогда $\dim L \leqslant \dim V$. Если $\dim L = \dim V$, то $L = V$.*

Доказательство. Пусть $\mathbf{e}'$ — базис L, и $\mathbf{e}$ — базис V из (6.5). По следствию 6.3 число векторов в $\mathbf{e}'$ ($= \dim L$) не превосходит числа векторов в $\mathbf{e}$ ($= \dim L$).

Если $\dim L = \dim V$, то, присоединяя к $\mathbf{e}'$ любой вектор из $\mathbf{e}$, получаем зависимую систему. Поэтому любой вектор из $\mathbf{e}$ лежит в линейной оболочке $\mathbf{e}'$, т. е. в L. Отсюда $L = V$. □

Скажем, что пространство V является *прямой суммой* подпространств U, W (обозначение $V = U \oplus W$), если каждый вектор $x \in V$ обладает и притом единственным представлением $x = y + z$, где $y \in U$, $z \in W$.

Теорема 6.8. *Пусть $V = U \oplus W$. Тогда $U \cap W = 0$. Если V конечномерно, то $\dim V = \dim U + \dim W$.*

Доказательство. Предположим, что $x \in U \cap W$. Тогда для вектора x имеем два представления $x = x + 0 = 0 + x$ в виде суммы векторов из U и W. В силу единственности такого представления получаем $x = 0$.

Для доказательства второго утверждения выберем в подпространстве U базис $e_1, \ldots, e_k$, а в подпространстве W — базис $e_{k+1}, \ldots, e_n$. Покажем, что векторы $e_1, \ldots, e_n$ составляют базис $V = U \oplus W$.

Если $x \in V$, то $x = y + z$, где $y \in U, z \in W$. Поэтому

$$y = x_1 e_1 + \ldots + x_k e_k, \quad z = x_{k+1} e_{k+1} + \ldots + x_n e_n .$$

Отсюда $x = x_1 e_1 + \ldots + x_n e_n$.

Предположим теперь, что $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$. Тогда

$$x = \lambda_1 e_1 + \ldots + \lambda_k e_k = -(\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n) \in U \cap W.$$

По доказанному выше получаем, что $x = 0$, т. е.

$$\lambda_1 e_1 + \ldots + \lambda_k e_k = \lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n = 0,$$

откуда $\lambda_1 = \ldots = \lambda_n = 0$, поскольку $e_1, \ldots, e_k$ — базис подпространства U, и $e_{k+1}, \ldots, e_n$ — базис подпространства W. □

Упражнение. Доказать, что

1) если U — подпространство конечномерного пространства V, то существует такое подпространство W в V, что $V = U \oplus W$;

2) если U, W — подпространства в конечномерном пространстве V, причем $U \cap W = 0$ и $\dim U + \dim W = \dim V$, то $V = U \oplus W$.

Рангом линейного пространства называется максимальное число линейно независимых векторов этой системы.

Упражнение. Доказать, что ранг системы векторов равен размерности их линейной оболочки.

Предположим, что задана система векторов $a_1, \ldots, a_m$ в векторном пространстве V с базисом $\mathbf{e}$ из (6.5). Пусть задано разложение $a_i = e_1 a_{1i} + \ldots + e_n a_{ni}$ каждого вектора a_i, $1 \leqslant i \leqslant m$, по базису $\mathbf{e}$. Положим $A = (a_{rs}) \in \mathrm{Mat}(n \times m)$. Тогда

$$(a_1, \ldots, a_m) = \mathbf{e}A. \tag{6.7}$$

Опишем алгоритм решения следующей задачи:

необходимо найти максимальную независимую подсистему системы векторов $a_1, \ldots, a_m$ и выразить остальные векторы системы через выбранную подсистему.

Для решения этой задачи приведем матрицу A элементарными преобразованиями к ступенчатому виду B. Пусть, например,

$$B = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & b_{1,r+1} & \ldots & b_{1m} \\ 0 & 1 & 0 & \ldots & 0 & 0 & b_{2,r+1} & \ldots & b_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & b_{r,r+1} & \ldots & b_{rm} \end{pmatrix}. \tag{6.8}$$

Заметим, что системы однородных уравнений с матрицами A и B эквивалентны. Решениями этих систем являются такие столбцы $X = {}^t(x_1, \ldots, x_n)$, что $AX = BX = 0$. В силу (6.7) это эквивалентно тому, что $(a_1, \ldots, a_n)X = 0$. Из вида B вытекает, что для любого $i = 1, \ldots, r$, выполнены равенства

$$x_i + b_{i,r+1}x_{r+1} + \ldots + b_{i,n}x_n = 0. \tag{6.9}$$

При этом неизвестные $x_1, \ldots, x_n$ — главные, а остальные — свободные. Придадим одной свободной неизвестной x_j, $j > r$, значение 1, а всем остальным свободным неизвестным — нулевое значение. Тогда главные неизвестные принимают в силу (6.9) значения $x_1 = -b_{1j}, \ldots, x_r = -b_{rj}$. Отсюда $a_j = a_1 b_{1j} + \ldots + a_r b_{rj}$.

Если же все свободные неизвестные $x_{r+1}, \ldots, x_n$ принимают нулевое значение, то и значения главных неизвестных $x_1, \ldots, x_r$ в силу (6.9) равны 0. Отсюда следует, что первые r векторов $a_1, \ldots, a_r$ независимы.

§ 2. Ранг матрицы

Рангом $r(A)$ матрицы A называется максимальное число линейно независимых строк A. Другими словами, ранг матрицы — это ранг её системы строк или размерность линейной оболочки строк A.

Предположим, что в матрице A выделены r строк и столбцов. На их пересечении строит квадратная матрица размера r. Ее определитель называется *минором* M порядка r матрицы A. Любой минор порядка $r+1$, получающийся присоединением еще одной строки и столбца A, называется *окаймляющим* для M.

Теорема 6.9 (Теорема об окаймляющем миноре). *Ранг матрицы равен порядку ненулевого минора, все окаймляющие которого равны* 0.

Доказательство. Пусть для простоты указанный минор M размера r расположен в верхнем левом углу матрицы A. Присоединим к M i-ую строку и j-ый столбец. Получающийся окаймляющий минор по условию всегда равен 0 (включая случаи, когда либо $i < r$, либо $j < r$. Разложим этот минор по присоединенному столбцу

$$0 = a_{1j}A_{1,r+1} + \ldots + a_{rj}A_{r,r+1} + a_{ij}M.$$

Так как $M \neq 0$, то

$$a_{ij} = a_{1j}(-\frac{A_{1,r+1}}{M}) + \ldots + a_{rj}(-\frac{A_{r,r+1}}{M}). \qquad (6.10)$$

В (6.10) коэффициенты $-\frac{A_{s,r+1}}{M}$, $s = 1, \ldots, r$, не зависят от j. Поэтому объединяя равенства (6.10) для всех j, получаем, что i-ая строка является линейной комбинацией первых r строк A. По теореме 6.1 первые r строк A независимы. □

Теорема 6.10 (Теорема о ранге матрицы). *Ранг системы строк матрицы совпадает с рангом системы столбцов.*

Доказательство. Порядок ненулевого минора, у которого все окаймляющие миноры равны нулю не меняется при переходе к транспонированной матрице. Применить теорему 6.9. □

Теорема 6.11. *Ранг матрицы не меняется при элементарных преобразованиях строк и столбцов.*

Доказательство. Линейная оболочка строк матрицы не меняется при элементарных преобразованиях строк. Поэтому не меняется её размерность и, следовательно, ранг матрицы. Аналогично, при элементарных преобразованиях столбцов матрицы линейная оболочка столбцов не меняется. □

В примере 6.6 отмечено, что все решения однородной системы $AX = 0$ образует подпространство в пространстве всех столбцов. Найдем его размерность.

Теорема 6.12. *Пространство решений однородной системы $AX = 0$ с n неизвестными имеет размерность $n - r(A)$, где $r(A)$ — ранг матрицы A.*

Доказательство. В силу теоремы 1.2 и теоремы 6.11 можно считать, что матрица A приведена в ступенчатому виду B из (6.8), где $r = r(A)$ — ранг матрицы A. Тогда система уравнений имеет вид (6.9), где неизвестные $x_1, \ldots, x_r$ — главные, а остальные — свободные. Придадим свободной неизвестной x_j, $j > r$, значение 1, а остальным свободным неизвестным — нулевое значение. Тогда главные неизвестные принимают в силу (6.9) значения $x_1 = -b_{1j}, \ldots, x_r = -b_{rj}$. Получаем решение

$$e_j = (-b_{1j}, \ldots, -b_{rj}, 0, \ldots, 0, 1, 0, \ldots, 0), \quad j = r+1, \ldots, n,$$

где 1 стоит на j-ом месте. Нетрудно видеть, что построенные решения $e_{r+1}, \ldots, e_n$ независимы.

Пусть $a = {}^t(a_1, \ \ldots, \ a_n)$ — произвольное решение системы. Тогда $b = a - a_{r+1}e_{r+1} - \ldots - a_ne_n$ также является решением, у которого все значения свободных переменных нулевые. Из (6.9) видно, что вектор $b = 0$. Итак, любое решение системы является линейной комбинацией $e_{r+1}, \ldots, e_n$. Поэтому они составляют базис пространства решений. □

Теорема 6.13 (Теорема Кронекера-Капелли). *Система линейных уравнений $AX = b$ совместна тогда и только тогда, когда ранги матрицы системы A и расширенной матрицы $(A \mid b)$ совпадают.*

Доказательство. Пусть столбец Λ является решением системы. Тогда $A\Lambda = b$, т. е. линейные оболочки столбцов A и $(A \mid b)$ совпадают. Отсюда вытекает совпадение рангов этих матриц.

Обратно, пусть ранги матриц A и $(A \mid b)$ совпадают. Линейная оболочка столбцов A содержится в линейной оболочке $(A \mid b)$, причем их размерности совпадают. По теореме 6.7 эти оболочки совпадают. Поэтому b лежит в линейно оболочке столбцов A. □

Теорема 6.14. *Ранг произведения матриц не превосходит ранга множителей. Если один из множителей является обратимой матрицей, то ранг произведения равен другому множителю.*

Доказательство. Пусть заданы прямоугольные матрицы A, C, $D = AC$. Каждый столбец матрицы D является линейной комбинацией столбцов матрицы A, а каждая строка матрицы D является линейной комбинацией строк матрицы C. По теореме 6.2 (см. также теорему 6.7) получаем первое утверждение.

Пусть матрица A обратима. По доказанному $r(D) \leqslant r(C)$. Кроме того, $r(C) = r(A^{-1}D) \leqslant r(D)$. Отсюда $r(C) = r(D)$. □

§ 3. Плоскости

Пусть U — подпространство в пространстве V и $a \in V$. *Плоскостью* $\Pi = a + U$, *проходящей через точку* a *с направляющим подпространством* U, называется множество всех векторов (точек) вида $\Pi = a + U = \{a + u \mid u \in U\}$.

Если U — направляющее подпространство для плоскости Π, то $U = \{x - y \mid x, y \in \Pi\}$. В частности, направляющее подпространство U по плоскости Π определено однозначно. Поэтому *размерностью* плоскости Π называется размерность его направляющего пространства U.

Если $U = \langle a_1, \ldots, a_k \rangle$ является линейной оболочкой векторов $a_1, \ldots, a_k$, то Π состоит из всех векторов $a + a_1 t_1 + \ldots + a_k t_k$, где $t_1, \ldots, t_k$ — произвольные скаляры. Это представление точек плоскости называется *параметрическим заданием* Π. Если в пространстве задан базис $\mathbf{e}$ из (6.5) и

$$\begin{aligned} a &= \alpha_1 e_1 + \ldots + \alpha_n e_n, \\ a_1 &= \alpha_{11} e_1 + \ldots + \alpha_{1n} e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_k &= \alpha_{k1} e_1 + \ldots + \alpha_{kn} e_n, \end{aligned}$$

то плоскость Π состоит из всех векторов $x = x_1 e_1 + \ldots + x_n e_n$ с координатами

$$\begin{aligned} x_1 &= \alpha_1 + t_1 \alpha_{11} + \ldots + t_k \alpha_{k1}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= \alpha_n + t_1 \alpha_{1n} + \ldots + t_k \alpha_{kn} \end{aligned}$$

в базисе $\mathbf{e}$, где $t_1, \ldots, t_k$ — произвольные числа.

Теорема. *Все решения совместная систем линейных уравнений $AX = b$ образуют плоскость Π размерности $n - r(A)$. Его направляющее пространство состоит из всех решений однородной системы $AY = 0$.*

Доказательство. Зафиксируем частное решение X_0 системы $AX = b$. Тогда общее решение Z имеет вид $Z = X_0 + T$, где T — произвольное решение однородной системы $AY = 0$. Отсюда вытекает утверждение. □

Теорема. *Пусть в пространства V с базисом* $\mathbf{e}$ *из* (6.5) *задана плоскость $\Pi = a + U$ размерности d. Тогда существует такая система из $n - d$ линейных уравнений, что вектор $x = x_1e_1 + \ldots + x_ne_n \in \Pi$ в том и только в том случае, если столбец координат ${}^t(x_1, \ldots, x_n)$ является решением этой системы.*

Доказательство. Пусть $U = \langle f_1, \ldots, f_k \rangle$, где $k \geqslant d$. Предположим, что $f_i = e_1f_{1i} + \ldots + e_nf_{ni}$, где $i = 1, \ldots, k$ и $a = e_1c_1 + \ldots + e_nc_n$. Составим матрицу

$$D = \begin{pmatrix} f_{11} & \ldots & f_{1k}, & x_1 - c_1 \\ \multicolumn{4}{c}{\dotfill} \\ f_{n1} & \ldots & f_{nk}, & x_n - c_n \end{pmatrix}.$$

Вектор x лежит в Π тогда и только тогда, когда $x - a \in U$. Это эквивалентно тому, что последний столбец D является линейной комбинацией первых k столбцов, т. е. система линейных уравнений с расширенной матрицей D совместна. Для решения этой системы приведем матрицу D к ступенчатому виду. Предположим, что он имеет вид

$$D' = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \ldots & \ldots & l_1(x,a) \\ 0 & 1 & 0 & \ldots & 0 & 0 & \ldots & \ldots & l_2(x,a) \\ \multicolumn{9}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 0 & 1 & \ldots & \ldots & l_d(x,a) \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & l_{d+1}(x,a) \\ \multicolumn{9}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & l_n(x,a) \end{pmatrix},$$

где ранг d системы векторов $f_1, \dots, f_k$ равен $\dim U = \dim \Pi$. Здесь $l_i(x,p)$ — линейные выражения от координат x, a. Таким образом, по теореме Кронекера-Капелли 6.13 система совместна тогда и только тогда, когда

$$l_{d+1}(x,a) = \dots = l_n(x,a) = 0. \qquad \square$$

Следствие. *Непустое пересечение конечного числа плоскостей является плоскостью.*

Плоскость размерности 1 называется *прямой*. Плоскость размерности $n-1$ в n-мерном пространстве называется *гиперплоскостью*.

Пример. Пусть заданы две точки в $\mathbb{R}^n$,

$$A = (a_1, \dots, a_n) \neq B = (b_1, \dots, b_n).$$

Покажем, что существует и притом единственная прямая, проходящая через эти точки.

Действительно, прямая имеет вид $X = A + t(B - A)$. Для задания системой линейных уравнений напишем матрицу

$$\begin{pmatrix} b_1 - a_1 & x_1 - a_1 \\ \dots\dots & \dots\dots \\ b_n - a_n & x_t - a_n \end{pmatrix}.$$

Так как точки A, B различны, то первый столбец отличен от нуля. Пусть $b_i \neq a_i$. Необходимо, чтобы ранг этой матрицы равнялся 1. Это эквивалентно тому, что все миноры

$$\begin{vmatrix} b_i - a_i & x_i - a_i \\ b_j - a_j & x_j - a_j \end{vmatrix} = 0, \quad j = 1, \dots, i-1, i+1, \dots, n.$$

Если же нужно провести плоскость Π наименьшей размерности, проходящую через заданные точки $A_0, \dots, A_m$ то нужно рассмотреть плоскость Π, проходящую через точку A_0 с направляющим подпространством, являющимся линейной оболочкой $\langle A_1 - A_0, \dots, A_m - A_0 \rangle$.

Скажем, что плоскость $\Pi_1 = A_1 + U_1$ *параллельна плоскости* $\Pi_2 = A_2 + U_2$, обозначение $\Pi_1 \parallel \Pi_2$, если $\Pi_1 \cap \Pi_2 = \varnothing$ и $U_1 \subseteq U_2$.

Теорема. *Две гиперплоскости* Π_1, Π_2, *заданные, соответственно, уравнениями*

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i, \quad i = 1, 2, \tag{6.11}$$

параллельны тогда и только тогда, когда выполнено одно из следующих эквивалентных условий:

(1) $\Pi_1 \cap \Pi_2 = \varnothing$;

(2) *справедливы соотношения*

$$\frac{a_{11}}{a_{21}} = \cdots = \frac{a_{1n}}{a_{2n}} \neq \frac{b_1}{b_2}. \tag{6.12}$$

Доказательство. Из параллельности плоскостей вытекает условие (1). Обратно, пусть выполнено условие (1). Тогда система из двух уравнений (6.11) несовместна, и поэтому по теореме Кронекера-Капелли 6.13

$$r\begin{pmatrix} a_{11}, \ldots, a_{1n} \\ a_{21}, \ldots, a_{2n} \end{pmatrix} < r\begin{pmatrix} a_{11}, \ldots, a_{1n}, b_1 \\ a_{21}, \ldots, a_{2n}, b_n \end{pmatrix} \leqslant 2. \tag{6.13}$$

По условию хотя бы одни из коэффициентов $a_{1i}, a_{2j} \neq 0$. Следовательно, в (6.13) ранг первой матрицы равен 1, а второй 2. Отсюда вытекает (6.12).

Обратно, если выполнено (6.12), то в (6.13) ранг первой матрицы равен 1, а второй — 2. Это означает, что две гиперплоскости, заданные уравнениями (6.11), не пересекаются по теореме Кронекера-Капелли 6.13. Кроме того, направляющие пространства U_i задаются пропорциональными однородными уравнениями

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

Поэтому $U_1 = U_2$ и из условия (2) вытекает параллельность плоскостей. $\square$

ГЛАВА 7

ЕВКЛИДОВЫ И ЭРМИТОВЫ ПРОСТРАНСТВА

В этой главе мы рассматриваем геометрию евклидовых и эрмитовых пространств произвольной конечной размерности. Излагается процесс ортогонализации Грама-Шмидта, вычисляются расстояния от вектора до плоскости, угол между вектором и плоскостью, векторное произведение.

§1. Полуторалинейные функции

Комплексно-значная (вещественно-значная) функция $b(x, y)$, аргументами которой являются векторы x, y комплексного (вещественного) векторного пространства V называется *полуторалинейной (билинейной)*, если

1) $b(\alpha x_1 + \beta x_2, y) = \alpha b(x_1, y) + \beta b(x_2, y)$;

2) $b(x, \alpha y_1 + \beta y_2) = \overline{\alpha} b(x, y_1) + \overline{\beta} b(x, y_2)$

   (соответственно, $b(x, \alpha y_1 + \beta y_2) = \alpha b(x, y_1) + \beta b(x, y_2)$)

для всех $x_1, x_2, x, y_1, y_2, y \in V$ и всех скаляров α, β. Здесь $\overline{\alpha}, \overline{\beta}$ — комплексное сопряжение. В случае вещественного пространства для вещественных скаляров α, β имеем $\overline{\alpha} = \alpha$, $\overline{\beta} = \beta$, т. е. в вещественном случае все сводится только к билинейным функциям.

Матрицей Грама A полуторалинейной (билинейной) функции $b(x,y)$ в базисе (6.5) называется матрица

$$A = \begin{pmatrix} b(e_1,e_1), & \dots & b(e_1,e_n) \\ \dots & \dots & \dots \\ b(e_n,e_1), & \dots & b(e_n,e_n) \end{pmatrix} \in \mathrm{Mat}(n).$$

Нетрудно видеть, что если

$$x = \mathbf{e}X, \quad y = \mathbf{e}Y, \quad \text{где} \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

то

$$b(x,y) = \sum_{i,j=1}^{n} x_i b(e_i,e_j)\overline{y}_j = {}^tXA\overline{Y}. \tag{7.1}$$

Предположим, что мы перешли к новому базису $\mathbf{e}' = \mathbf{e}C$ с помощью матрицы перехода C. В этом случае i-ый столбец C_i матрицы C является координатами i-го вектора e'_i из базиса $\mathbf{e}'$. Следовательно, $b(e'_i, e'_j) = {}^tC_iA\overline{C}_j$ в силу (7.1). Таким образом, матрица A' функции $b(x,y)$ в базисе $\mathbf{e}'$ имеет вид

$$A' = {}^tCA\overline{C}. \tag{7.2}$$

Полуторалинейная (билинейная) функция $b(x,y)$ называется *эрмитовой (симметричной)*, если $b(x,y) = \overline{b(y,x)}$ для всех векторов $x, y \in V$.

Упражнение. Доказать, что если $b(x,y)$ — эрмитова (симметричная) функция, то $b(x,x) \in \mathbb{R}$.

Предложение 7.1. *Предположим, что полуторалинейная (билинейная) функция $b(x,y)$ в некотором базисе $\mathbf{e}$ имеет матрицу A. Следующие условия эквивалентны:*

1) *функция $b(x,y)$ эрмитова (симметрична);*

2) ${}^tA = \overline{A}$.

Доказательство. Ясно, что из условия 1) вытекает условие 2). Обратно, если $\overline{A} = {}^tA$, то в силу (7.1)

$$b(y,x) = {}^tYA\overline{X} = {}^t({}^tYA\overline{X}) = {}^t\overline{X}\,\overline{A}Y =$$
$$= \overline{({}^tXA)}Y = \overline{{}^tXA\overline{Y}} = \overline{b(x,y)}. \qquad \square$$

Эрмитова (симметрическая) функция $b(x,y)$ *положительно определена*, если $b(x,x) > 0$ при $x \neq 0$.

Теорема 7.2 (Критерий Сильвестра). *Пусть задана эрмитова (симметричная) функция $b(x,y)$ в n-мерном пространстве имеет в некотором базисе матрицу Грама $B = (b_{ij})$. Следующие условия эквивалентны:*

1) *функция $b(x,y)$ положительно определена,*

2) *для любого $k = 1,\dots,n$ положителен минор*

$$\begin{vmatrix} b_{11} & \dots & b_{1k} \\ \dots & \dots & \dots \\ b_{k1} & \dots & b_{kk} \end{vmatrix}.$$

Доказательство будет приведено в §7.3.

Отметим, что эта теорема часто применяется для нахождения максимальных значений функций от нескольких переменных.

§2. Квадратичные функции

Если задана эрмитова (билинейная симметричная) функция $b(x,y)$, то по ней определяется *квадратичная функция* $q(x) = b(x,x)$.

Предложение 7.3. *Билинейная симметричная (эрмитова) функция по квадратичной функции восстанавливается однозначно.*

Доказательство. Нетрудно видеть, что для билинейной функции получаем

$$b(x,y) = \frac{1}{2}\left[b(x+y,x+y) - b(x,x) - b(y,y)\right] = \\ = \frac{1}{2}\left[q(x+y) - q(x) - q(y)\right],$$

для эрмитовой функции

$$4b(x,y) = b(x+y,x+y) - b(x-y,x-y) + \\ + ib(x+iy,x+iy) - ib(x-iy,x-iy) = \\ = q(x+y) - q(x-y) + iq(x+iy) - ib(x-iy). \quad \square$$

Матрицей квадратичной функции $q(x)$ в базисе $\mathbf{e}$ называется матрица Грама A соответствующей билинейной симметрической или эрмитовой функции $b(x,y)$ в этом базисе. Из предыдущего предложения вытекает корректность этого определения.

Из (7.1) вытекает, что если $A = (a_{ij})$ и $q(x)$ получена из билинейной функции, то

$$q(x) = b(x,x) = {}^tXAX = \sum_{i,j=1}^{n} a_{ij}x_ix_j = \\ = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1\leqslant i<j\leqslant n} a_{ij}x_ix_j. \tag{7.3}$$

§ 3. Скалярные произведения

Пусть в вещественном (комплексном) векторном пространстве V задана эрмитова положительно определенная функция, которая называется *скалярным произведением*. Вещественное (комплексное) пространство V со скалярным произведением называется *евклидовым (эрмитовым) пространством*.

Примеры. Пусть $V = \mathbb{R}^3$, где $(x,y) = |x||y|\cos\varphi$ или $V = C_{[0,1]}$, где $(f,g) = \int\limits_{i=0}^{n} f(x)\overline{g(x)}dx$.

Два ненулевых вектора x, y евклидова (эрмитова) пространства *перпендикулярны, или ортогональны*, если $(x, y) = 0$ (обозначение $x \perp y$). Система ненулевых векторов $a_1, \dots, a_k$ из евклидова (эрмитова) пространства E *ортогональна*, если $a_i \perp a_j = 0$ при всех $1 \leqslant i \neq j \leqslant k$. Система векторов $a_1, \dots, a_k$ из евклидова (эрмитова) пространства E *ортонормирована*, если $(a_i, a_j) = \delta_{ij}$ при всех $1 \leqslant i \neq j \leqslant k$. *Матрицей Грама* системы векторов $b_1, \dots, b_k$ из евклидова (эрмитова) пространства E, как и выше, называется матрица

$$G(b_1, \dots, b_k) = \begin{pmatrix} (b_1, b_1) & \dots & (b_1, b_k) \\ \dots & \dots & \dots \\ (b_k, b_1) & \dots & (b_k, b_k) \end{pmatrix}. \tag{7.4}$$

Длиной вектора x называется $\|x\| = \sqrt{(x, x)}$.

Упражнение. Доказать, что $\|x\| \geqslant 0$ и $\|x\| = 0$ тогда и только тогда, когда $x = 0$.

Если задан в евклидовом (эрмитовом) пространстве E задан базис $\mathbf{e}$ из (6.5) то для $x = x_1e_1 + \ldots + x_ne_n$, $y = y_1e_1 + \ldots + y_ne_n$, из (7.1) следует, что

$$(x, y) = (x_1, \dots, x_n)G(e_1, \dots, e_n) \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}. \tag{7.5}$$

Таким образом, с помощью матрицы Грама удобно считать скалярные произведения.

Отметим, что базис $\mathbf{e}$ из (6.5) в евклидовом (эрмитовом) пространстве E ортонормирован тогда и только тогда, когда матрица Грама $G(e_1, \dots, e_n) = E_n$. В частности, если базис $\mathbf{e}$ ортонормирован и $x = x_1e_1 + \ldots + x_ne_n$, $y = y_1e_1 + \ldots + y_ne_n$, то по (7.5)

$$(x, y) = x_1\overline{y_1} + \ldots + x_n\overline{y_n}. \tag{7.6}$$

Предложение 7.4. *Ортогональная система векторов независима.*

Доказательство. Пусть задана ортогональная система векторов $a_1,\ldots,a_k$ из евклидова (эрмитова) пространства E, и $\lambda_1 a_1 + \ldots + \lambda_k a_k = 0$. Тогда для любого $i = 1,\ldots,k$

$$\begin{aligned} 0 = (0, a_i) = (\lambda_1 a_1 + \ldots + \lambda_k a_k, a_i) = \\ = \lambda_1(a_1, a_i) + \ldots + \lambda_k(a_k, a_i) = \lambda_i(a_i, a_i). \end{aligned} \tag{7.7}$$

Так как $a_i \neq 0$, то по определению $(a_i, a_i) > 0$ Следовательно, из (7.7) получаем $\lambda_i = 0$ для любого $i = 1,\ldots,k$. □

§ 4. Процесс ортогонализации и матрица Грама

Пусть задана система независимая векторов $a_1,\ldots,a_k$ из евклидова (эрмитова) пространства E. Построим ортогональную системы векторов $b_1,\ldots,b_k \in E$ со следующим свойством:

$$b_1 = a_1, \text{ и } \langle b_1, \ldots b_s\rangle = \langle a_1, \ldots a_s\rangle \text{ для всех } s \leqslant k. \tag{7.8}$$

Первый шаг очевиден. Пусть для некоторого $t < k$ уже построены векторы $b_1,\ldots b_t$ со свойством (7.8) для всех $s \leqslant t$. Положим

$$b_{t+1} = a_{t+1} - \frac{(a_{t+1}, b_1)}{(b_1, b_1)} b_1 - \ldots - \frac{(a_{t+1}, b_t)}{(b_t, b_t)} b_t. \tag{7.9}$$

Несложная проверка показывает, что $(b_{t+1}, b_j) = 0$ для всех $1 \leqslant j \leqslant t$. При этом выполнено условие (7.8) для всех $s \leqslant t + 1$.

Замечание 7.5. Заметим, что если исходная система векторов $a_1,\ldots,a_k$ ортогональна, то $a_1 = b_1,\ldots,a_k = b_k$.

Пусть $a_1,\ldots,a_k$ — независимая система векторов в E и система $b_1,\ldots,b_k$ получена из нее процессом ортогонализации. Из (7.9) вытекает, что

$$(b_1,\ldots,b_k) = (a_1,\ldots,a_k) \begin{pmatrix} 1 & * & \ldots & * & * \\ 0 & 1 & \ldots & * & * \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & * \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}. \tag{7.10}$$

Теорема 7.6. *Определитель матрицы Грама не меняется при процессе ортогонализации.*

Доказательство. В силу (7.2) и (7.10) получаем, что

$$G(b_1,\ldots,b_k)=\begin{pmatrix}1&0&\ldots&0&0\\ *&1&\ldots&0&0\\ \ldots&\ldots&\ldots&\ldots&\ldots\\ *&*&\ldots&1&0\\ *&*&\ldots&*&1\end{pmatrix}G(a_1,\ldots,a_k)\times$$

$$\times\begin{pmatrix}1&\bar{*}&\ldots&\bar{*}&\bar{*}\\ 0&1&\ldots&\bar{*}&\bar{*}\\ \ldots&\ldots&\ldots&\ldots&\ldots\\ 0&0&\ldots&1&\bar{*}\\ 0&0&\ldots&0&1\end{pmatrix}.$$

В силу теоремы 3.8 об определителе произведения матриц и теоремы 3.1 получаем, что $\det G(b_1,\ldots,b_k)=\det G(a_1,\ldots,a_k)$. □

Теорема 7.7. *В любом евклидовом (эрмитовом) пространстве E существует ортонормированный базис.*

Доказательство. Пусть $a_1,\ldots,a_n$ — произвольный базис в E. Применяя процесс ортогонализации получаем ортогональный базис $b_1,\ldots,b_n$ в E в силу предложения 7.4. Остается положить $e_i=\frac{b_i}{\|b_i\|}$. □

Доказательство теоремы 7.2 (Сильвестра). Пусть выполнено условие 1) и эрмитова функция $b(x,y)$ положительно определена и в базисе $\mathbf{e}$ из (6.5) задана её матрица Грама $B=(b_{ij})$. Тогда пространство является эрмитовым (евклидовым), если в качестве скалярного произведения взять функцию $b(x,y)$.

Рассмотрим линейную оболочку $U=\langle e_1,\ldots,e_k\rangle$. По теореме 7.7 в U имеется ортонормированный базис $\mathbf{e}'$ с матрицей перехода $\mathbf{e}'C=(e_1,\ldots,e_k)$. Пусть B' — матрица минора из условия 2) теоремы 7.2. Тогда $B'={}^t C\overline{C}$ по (7.2), поскольку в базисе $\mathbf{e}'$ матрица Грама для $b(x,y)$ единичная. Отсюда

$$\det B'=(\det C)\left(\det\overline{C}\right)=(\det C)\overline{(\det C)}=|\det C|^2>0. \quad (7.11)$$

Обратно, пусть выполнено условие 2). Если размерность пространства равна 1, то для любого $x = x_1 e_1 \neq 0$ имеем $b(x,x) = |x_1|^2 b_{1,1} > 0$, и поэтому в этом случае теорема верна.

Пусть для $n-1$ теорема верна, и U — линейная оболочка векторов $e_1, \dots, e_{n-1}$. По индукции на U функция $b(x,y)$ положительно определена. Следовательно, по теореме 7.7 в U можно найти ортонормированный относительно скалярного произведения $b(x,y)$ базис $(e'_1, \dots, e'_{n-1})$. Тогда $\mathbf{e}' = (e'_1, \dots, e'_{n-1}, e_n)$ — базис всего пространства. Положим $\mathbf{e}' = \mathbf{e}C$, где C — матрица перехода. Тогда матрица Грама функции в базисе $\mathbf{e}'$ имеет вид

$$ {}^tCB\overline{C} = \begin{pmatrix} E_{n-1} & {}^t\overline{d} \\ d & r \end{pmatrix}, $$

где $d = (d_1, \dots, d_{n-1})$ — строка длины $n-1$ и r — число. Положим

$$ g_i = \begin{cases} e'_i, & i < n, \\ e_n - d_1 e'_1 - \dots - d_{n-1} e'_{n-1}, & i = n. \end{cases} $$

Тогда векторы $\mathbf{g} = (g_1, \dots, g_n)$ образуют базис пространства, причем

$$ \mathbf{g} = \mathbf{e}'D, \quad D = \begin{pmatrix} E_{n-1} & -{}^td \\ 0 & 1 \end{pmatrix}. $$

В базисе $\mathbf{g}$ матрица Грама функции $b(x,y)$ имеет вид

$$ \begin{aligned} {}^tD\,{}^tCB\overline{CD} &= \begin{pmatrix} E_{n-1} & 0 \\ -d & 1 \end{pmatrix} \begin{pmatrix} E_{n-1} & {}^t\overline{d} \\ d & r \end{pmatrix} \begin{pmatrix} E_{n-1} & -{}^t\overline{d} \\ 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} E_{n-1} & 0 \\ 0 & r - d\,{}^t\overline{d} \end{pmatrix}. \end{aligned} \tag{7.12} $$

Как и в (7.11), в силу условия 2) получаем

$$ \begin{aligned} \det\left({}^t(CD)B\overline{CD}\right) &= (\det B)\,|\det C|^2\,|\det D|^2 = \\ &= r - d\,{}^t\overline{d} = r - |d_1|^2 - \dots - |d_{n-1}|^2 > 0. \end{aligned} $$

Пусть $x = x_1 g_1 + \dots + x_n g_n$. Тогда

$$ b(x,x) = (x_1, \dots, x_n) \begin{pmatrix} E_{n-1} & 0 \\ & r - d\,{}^t\overline{d} \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix} = $$

$$= |x_1|^2 + \ldots + |x_{n-1}|^2 + |x_n|^2(r - d^t\overline{d}) > 0$$

при $x \neq 0$. □

Теорема 7.8. *Пусть $e_1, \ldots, e_k$ — ортонормированная система векторов из евклидова (эрмитова) пространства E. Тогда её можно дополнить до ортонормированного базиса в E*

.

Доказательство. Система векторов $e_1, \ldots, e_k$ независима по предложению 7.4. Поэтому по теореме 6.4 можно дополнить до базиса $e_1, \ldots, e_k, a_{k+1}, \ldots, a_n$ пространства E. Применяя к этому базису процесс ортогонализации в силу замечания 7.5 получаем ортогональный базис $b_1 = e_1, \ldots, b_k = e_k$, $b_{k+1}, \ldots, b_n$. Остается положить $e_i = \frac{b_i}{\|b_i\|}$ при $i = k+1, \ldots, n$. □

Теорема 7.9. *Пусть задан ортонормированный базис* **e** *из* (6.5) *в евклидовом (эрмитовом) пространстве E, и задана система векторов*

$$a_i = e_1 a_{1i} + \ldots + e_n a_{ni}, \quad i = 1, \ldots, k.$$

Тогда

$$G(a_1, \ldots, a_k) = \begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \ldots & \ldots & \ldots \\ a_{1k} & \ldots & a_{nk} \end{pmatrix} \begin{pmatrix} \overline{a_{11}} & \ldots & \overline{a_{1k}} \\ \ldots & \ldots & \ldots \\ \overline{a_{n1}} & \ldots & \overline{a_{kn}} \end{pmatrix}.$$

Доказательство. На месте (i,j) в произведении матриц из условия теоремы стоит $a_{i1}\overline{a_{1j}} + \ldots + a_{in}\overline{a_{nj}} = (a_i, a_j)$ в силу (7.6). Отсюда вытекает утверждение. □

Теорема 7.10. *Если векторы $a_1, \ldots, a_k$ независимы, то*

$$\det G(a_1, \ldots, a_k) > 0.$$

Если векторы $a_1, \ldots, a_k$ зависимы, то $\det G(a_1, \ldots, a_k) = 0$.

Доказательство. Пусть векторы $a_1, \ldots, a_k$ независимы, и векторы $b_1, \ldots, b_k$, получены из них процессом ортогонализации. При этом определитель матрицы Грама не меняется. Отсюда

$$\det G(a_1, \ldots, a_k) = \det G(b_1, \ldots, b_k) =$$

$$= \begin{vmatrix} (b_1, b_1) & 0 & 0 & \ldots & 0 & 0 \\ 0 & (b_2, b_2) & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & (b_{k-1}, b_{k-1}) & 0 \\ 0 & 0 & 0 & \ldots & 0 & (b_k, b_k) \end{vmatrix} =$$

$$= (b_1, b_1) \cdots (b_k, b_k) > 0.$$

Пусть теперь векторы $a_1, \ldots, a_k$ зависимы. Тогда ранг матрицы

$$\begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \ldots & \ldots & \ldots \\ a_{1k} & \ldots & a_{nk} \end{pmatrix}$$

меньше k. В силу теорем 7.9 и 6.14 ранг матрицы Грама меньше k, и поэтому её определитель равен 0. □

Пусть в евклидовом пространстве E заданы векторы $a_0, \ldots, a_k$. *Параллелепипедом* с вершинами в $a_0, \ldots, a_k$ называется множество всех векторов евклидова пространства вида

$$a_0 + t_1(a_1 - a_0) + \ldots + t_k(a_k - a_0), \quad 0 \leqslant t_i \leqslant 1.$$

k-мерным объемом этого параллелепипеда называется число

$$\sqrt{\det G(a_1 - a_0, \ldots, a_k - a_0)}.$$

Из теоремы 7.6 следует, что объем не меняется при ортогонализации системы векторов $a_1 - a_0, \ldots, a_k - a_0$. Кроме того, если эти векторы зависимы, то объем нулевой, а если они независимы, то он положителен. Из инвариантности определителя матрицы Грама при процессе ортогонализации следует, что объем равен

$(k-1)$-мерному объему основания, умноженному на длину высоты, проведенной из оставшейся вершины к этому основанию. В частности, если $k=1$, то параллелепипед совпадает с отрезком $[a_0, a_1]$ и соответствующий одномерный объем равен длине этого отрезка.

Пусть в n-мерном евклидовом (эрмитовом) пространстве E с ортонормированных базисом $\mathbf{e}$ из (6.5) задан $(n-1)$ вектор

$$a_i = a_{i1}e_1 + \ldots + a_{in}e_n, \quad i = 1, \ldots, n-1.$$

Векторным произведением $b = [a_1, \ldots, a_{n-1}] \in E$ называется вектор

$$b = \begin{vmatrix} e_1 & \ldots & e_n \\ \overline{a_{11}} & \ldots & \overline{a_{1n}} \\ \ldots & \ldots & \ldots \\ \overline{a_{n-1,1}} & \ldots & \overline{a_{n-1,n}} \end{vmatrix} = e_1\overline{A_{11}} + \ldots + e_n\overline{A_{1n}}.$$

где

$$A_{1j} = (-1)^{1+j} \begin{vmatrix} a_{11} & \ldots & a_{1,i-1} & a_{1,i+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & \ldots & a_{n-1,i-1} & a_{n-1,i+1} & \ldots & a_{n-1,n} \end{vmatrix}$$

— соответствующее алгебраическое дополнение.

Заметим, что вектор $b = 0$ тогда и только тогда, когда

$$A_{11} = \ldots = A_{1n} = 0.$$

В силу теоремы 6.9 это выполнено тогда и только тогда, когда векторы $a_1, \ldots, a_{n-1}$ линейно зависимы.

Теорема. $(a_j, b) = 0$ *для всех* $j = 1, \ldots, n-1$. *Длина*

$$\|b\| = \sqrt{\det G(a_1, \ldots, a_{n-1})}.$$

Кроме того, если $a_1, \ldots, a_{n-1}$ *независимы, система*

$$b, a_1, \ldots, a_{n-1}$$

является базисом, ориентированным так же, как и исходный базис $\mathbf{e}$, *т. е. матрица перехода от* $\mathbf{e}$ *к* $b, a_1, \ldots, a_{n-1}$ *имеет положительный определитель.*

Доказательство. Имеем

$$(a_j, b) = a_{j1}A_{11} + \ldots + a_{jn}A_{1n} = \begin{vmatrix} a_{j1} & \ldots & a_{jn} \\ \multicolumn{3}{c}{\dotfill} \\ a_{j1} & \ldots & a_{jn} \\ \multicolumn{3}{c}{\dotfill} \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = 0.$$

Поэтому, учитывая теорему 7.9, получаем

$$\|b\|^2 \cdot \det G(a_1, \ldots, a_{n-1}) = \det G(b, a_1, \ldots, a_{n-1}) =$$

$$= \det \begin{pmatrix} \overline{A_{11}} & \ldots & \overline{A_{1n}} \\ a_{11} & \ldots & a_{1n} \\ \multicolumn{3}{c}{\dotfill} \\ a_{n-1,1} & \ldots & a_{n-1,n} \end{pmatrix} \det \begin{pmatrix} A_{11} & \overline{a_{11}} & \ldots & \overline{a_{n-1,1}} \\ \multicolumn{4}{c}{\dotfill} \\ A_{1n} & \overline{a_{1n}} & \ldots & \overline{a_{n-1,n}} \end{pmatrix} =$$

$$= \left| \det \begin{pmatrix} \overline{A_{11}} & \ldots & \overline{A_{1n}} \\ a_{11} & \ldots & a_{1n} \\ \multicolumn{3}{c}{\dotfill} \\ a_{n-1,1} & \ldots & a_{n-1,n} \end{pmatrix} \right|^2 = (|A_{11}|^2 + \ldots + |A_{1n}|^2)^2 =$$

$$= \|b\|^4.$$

Следовательно, если $b \neq 0$, то $\|b\|^2 = \det G(a_1, \ldots, a_{n-1})$. Кроме того, матрица перехода от $\mathbf{e}$ к $b, a_1, \ldots, a_{n-1}$ имеет вид

$$\begin{pmatrix} \overline{A_{11}} & a_{11} & \ldots & a_{n1} \\ \multicolumn{4}{c}{\dotfill} \\ \overline{A_{1n}} & a_{1n} & \ldots & a_{nn} \end{pmatrix}.$$

Ее определитель по доказанному выше равен $\|b\|^2 > 0$. □

Упражнение. Пусть e_1, e_2, e_3 — ортонормированный базис в трехмерном пространства. Доказать, что $[e_1, e_2] = e_3$, $[e_1, e_3] = -e_2$.

§ 5. Геометрия евклидовых (эрмитовых) пространств

Теорема 7.11 (Теорема Пифагора). *Пусть x, y — векторы из евклидова (эрмитова) пространства, причем $x \perp y$. Тогда*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Доказательство вытекает из определения длины вектора.

Ортогональным дополнением $U^{\perp}$ к подпространству U евклидова (эрмитова) пространства E называется множество

$$U^{\perp} = \{x \in E \mid x \perp y \quad \forall y \in U\}.$$

Теорема 7.12. *$U^{\perp}$ является подпространством и $E = U \oplus U^{\perp}$.*

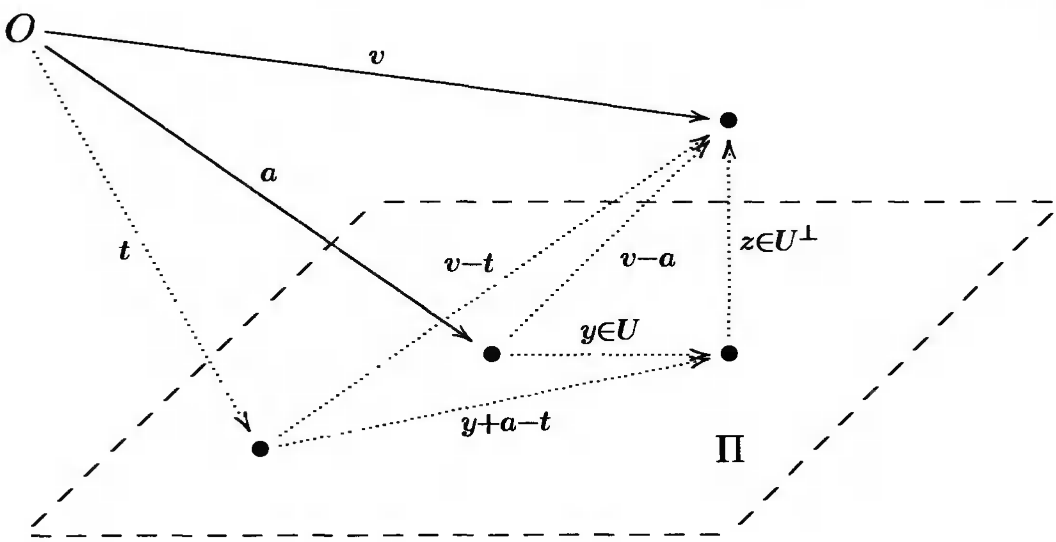
Доказательство. Пусть $e_1, \ldots, e_k$ — ортонормированный базис U. По теореме 7.8 систему $e_1, \ldots, e_k$ можно дополнить до ортонормированного базиса $e_1, \ldots, e_n$ в E.

Пусть $x = x_1 e_1 + \ldots + x_n e_n$. Тогда $x \in U^{\perp} \iff x_i = (x, e_i) = 0$ $\forall i = 1, \ldots, k$. Отсюда $U^{\perp} = \langle e_{k+1}, \ldots e_n \rangle$. □

Расстоянием $\rho(x, \Pi)$ от вектора x евклидова(эрмитова) пространства до плоскости Π называется $\inf\limits_{t \in \Pi} \|x - t\|$.

Теорема 7.13. *Пусть $\Pi = a + U$, и $v \in E$ имеет разложение $v - a = y + z$, где $y \in U$, $z \in U^{\perp}$. Тогда $\rho(v, \Pi) = \|z\|$.*

Доказательство. Имеем $t = a + u$, $u \in U$, т. е. $t - a = u \in U$.



Поэтому $v-t=a+y+z-t=(y+a-t)+z$, где $y+a-t \in U$, $z \in U^{\perp}$.

По теореме Пифагора 7.11

$$\rho(v,\Pi)=\inf_{t\in\Pi}\|v-t\|=\inf_{t\in\Pi}\|(y+a-t)+z\|=$$

$$=\inf_{t\in\Pi}\sqrt{\|y+a-t\|^2+\|z\|^2}=\|z\|. \quad \square$$

Изложим алгоритм нахождения расстояния от вектора v до плоскости $\Pi = a+\langle b_1,\ldots,b_k\rangle$. Предположим, что в некотором ортонормированном базисе $a=(a_1,\ldots,a_n)$, $b_i=(b_{1i},\ldots,b_{ni})$, $1\leqslant i\leqslant k$. Если $v-a=y+z$, $y\in U, z\in U^{\perp}$, то $y=t_1b_1+\ldots$ $\ldots+t_kb_k$. Отсюда для любого $i=1,\ldots,k$ получаем

$$(y,b_i)=(b_1,b_i)t_1+\ldots+(b_n,b_i)t_n.$$

С другой стороны,

$$(y,b_i)=(v-a-z,b_i)=(v-a,b_i).$$

Таким образом, получаем системы линейных уравнений

$$G(b_1,\ldots,b_k)\begin{pmatrix}t_1\\ \vdots\\ t_k\end{pmatrix}=\begin{pmatrix}(v-a,b_1)\\ \vdots\\ (v-a,b_k)\end{pmatrix}.$$

Решая эту систему, находим y и $z=v-a-y$.

В рассмотренном выше случае плоскость Π задается параметрически как $\Pi=a+\langle b_1,\ldots,b_k\rangle$. Предположим теперь, что плоскость Π задается в некотором ортонормированном базисе $\mathbf{e}$ системой линейных уравнений $f_1(x)+c_1=\ldots=f_m(x)+c_m=0$, где $f_i(x)=x_1a_{i1}+\ldots+x_na_{in}$, $i=1,\ldots,m$. В этом случае направляющее подпространство U для плоскости Π задается системой однородных уравнений $f_1(x)=\ldots=f_m(x)=0$. Рассмотрим векторы $d_1,\ldots,d_m$, где вектор d_i имеет в базисе $\mathbf{e}$ координаты

$$d_i=(\overline{a_{i1}},\ldots,\overline{a_{in}}),\quad i=1,\ldots,m.$$

В этом случае $f_i(x) = (x, d_i)$, т. е. систему однородных уравнений, задающих U можно интерпретировать как $(x, d_1) = \ldots$ $\ldots = (x, d_m) = 0$. Это означает, что $\langle d_1, \ldots, d_m \rangle = U^{\perp}$. Поэтому $z = t_1 d_1 + \ldots + t_m d_m$, откуда для любого $i = 1 \ldots, m$ получаем

$$(v - a, d_i) = (y + z, d_i) = (z, d_i) = t_1(d_1, d_i) + \ldots + t_m(d_m, d_i).$$

Так как $a \in \Pi$, то $(a, d_i) = f_i(a) = -c_i$. Поэтому $(v - a, d_i) =$ $= (v, d_i) - (a, d_i) = f_i(v_i) + c_i$. Следовательно, получаем систему линейных уравнений

$$G(d_1, \ldots, d_m) \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} = \begin{pmatrix} f_1(v) + c_1 \\ \vdots \\ f_m(v) + c_m \end{pmatrix}. \tag{7.13}$$

Например, пусть $a = (-1, 0, 3)$, $b_1 = (0, 1, -2)$, $b_2 = (2, -1, -2)$ и $v = (2, 1, 4)$. В этом случае $v - a = (3, 1, 1)$ и

$$G(b_1, b_2) = \begin{pmatrix} 5 & 3 \\ 3 & 9 \end{pmatrix}, \quad \begin{pmatrix} (v - a, b_1) \\ (v - a, b_2) \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}. \tag{7.14}$$

Таким образом, относительно t_1, t_2 получаем квадратную систему линейных уравнений с расширенной матрицей

$$\left(\begin{array}{cc|c} 5 & 3 & -1 \\ 3 & 9 & 3 \end{array}\right).$$

Деля второе уравнение на 3 и переставляя местами оба уравнения получаем

$$\left(\begin{array}{cc|c} 1 & 3 & 1 \\ 5 & 3 & -1 \end{array}\right).$$

Вычитая из второго уравнения первое, умноженное на 5, получаем

$$\left(\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & -12 & -6 \end{array}\right).$$

Далее делим второе уравнение на -12 и вычитаем из первого уравнения второе, умноженное на 3. Получаем

$$\left(\begin{array}{cc|c} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{array}\right).$$

Итак, $t_1 = -\frac{1}{2}$, $t_2 = \frac{1}{2}$. Поэтому

$$\begin{aligned} y &= -\frac{1}{2}b_1 + \frac{1}{2}b_2 = (1,-1,0), \\ z &= v - a - y = (2,2,1), \\ \|z\| &= \sqrt{4+4+1} = 3. \end{aligned}$$

Есть и другой способ явного вычисления указанного расстояния.

Теорема 7.14. *Пусть плоскость* Π *задана параметрически* $\Pi = a + \langle b_1, \ldots, b_k \rangle$, *причем векторы* $b_1, \ldots, b_k$ *независимы. Тогда расстояние от вектора* v *до* Π *равно*

$$\rho(v, \Pi) = \sqrt{\frac{\det G(b_1, \ldots, b_k, v-a)}{\det G(b_1, \ldots, b_k)}}. \tag{7.15}$$

Доказательство. Пусть $v - a = y + z$, где $y \in U$, $z \in U^{\perp}$. По теореме 7.6 определитель матрицы Грама не меняется при ортогонализации. Поэтому можно считать, что система векторов $b_1, \ldots, b_k$ ортогональна. При этом процессе из системы $b_1, \ldots, b_k, v - a$ получается набор векторов $b_1, \ldots, b_k, z$. По теореме 7.13 получаем, что $\rho(v, \Pi) = \|z\|$. Таким образом,

$$\sqrt{\frac{\det G(b_1, \ldots, b_k, v-a)}{\det G(b_1, \ldots, b_k)}} = \sqrt{\frac{\|b_1\|^2 \cdots \|b_k\|^2 \|z\|^2}{\|b_1\|^2 \cdots \|b_k\|^2}} = \\ = \sqrt{\|z\|^2} = \|z\| = \rho(v, \Pi). \qquad \square$$

Вычислим с помощью формулы (7.15) расстояние из предыдущего примера, где $b_1 = (0,1,-2)$, $b_2 = (2,-1,-2)$, $v - a = (3,1,1)$. Тогда $(v - a, v - a) = 11$ и по формуле (7.14)

$$\rho(v, \Pi) = \sqrt{\frac{\det G(b_1, \ldots, b_k, v-a)}{\det G(b_1, \ldots, b_k)}} =$$

$$= \sqrt{\frac{\begin{vmatrix} 5 & 3 & -1 \\ 3 & 9 & 3 \\ -1 & 3 & 11 \end{vmatrix}}{\begin{vmatrix} 5 & 3 \\ 3 & 9 \end{vmatrix}}} = \sqrt{\frac{18^2}{36}} = \sqrt{9} = 3.$$

Применим соображения из теоремы 7.14 для нахождения расстояния от вектора $v = (v_1, \ldots, v_n)$ до гиперплоскости $\alpha_1 x_1 + \ldots \ldots + \alpha_n x_n + \beta = 0$. Все координаты заданы в некотором ортонормированном базисе.

Теорема 7.15. *Указанное расстояние равно*

$$\frac{|\alpha_1 v_1 + \ldots + \alpha_n v_n + \beta|}{\sqrt{|\alpha_1|^2 + \ldots + |\alpha_n|^2}}.$$

Доказательство. В рассматриваемом выше примере $m = 1$ и вектор $d_1 = (\overline{\alpha_1}, \ldots, \overline{\alpha_n})$. Поэтому система уравнений (7.13) принимает вид

$$G(d_1)t_1 = (d_1, d_1)t_1 = \|d_1\|^2 t_1 = \alpha_1 v_1 + \ldots + \alpha_n v_n + \beta.$$

При этом $z = t_1 d_1$ имеет длину

$$\|z\| = |t_1|\|d_1\| = \frac{|\alpha_1 v_1 + \ldots + \alpha_n v_n + \beta|}{\|d_1\|}. \qquad \square$$

Предложение 7.16 (Неравенство Коши-Буняковского). *В евклидовом (эрмитовом) пространстве для любых векторов x, y выполнено неравенство $|(x, y)| \leqslant \|x\|\|y\|$, причем равенство имеет место тогда и только тогда, когда x, y пропорциональны.*

Доказательство. По теореме 7.10

$$G(x, y) = \begin{vmatrix} (x, x) & (x, y) \\ (x, y) & (y, y) \end{vmatrix} = \|x\|^2\|y\|^2 - (x, y)(y, x) =$$
$$= \|x\|^2\|y\|^2 - |(x, y)|^2 \geqslant 0,$$

причем равенство имеет место тогда и только тогда, когда x, y линейно зависимы. □

Угол φ между двумя ненулевыми векторами x, y евклидова пространства E определяется из условия

$$\cos\varphi = \frac{(x,y)}{\|x\|\|y\|}, \quad 0 \leqslant \varphi \leqslant \pi.$$

Это определение корректно в силу неравенства Коши-Буняковского. При этом $x \perp y \iff$ угол между ними равен $\frac{\pi}{2}$.

Пусть Π — плоскость в евклидовом пространстве E с направляющим пространством U, и $x \in E \setminus 0$. *Углом* между x и Π называется инфимум углов между x и всеми векторами из U.

Теорема 7.17. *Пусть Π — плоскость в евклидовом пространстве E с направляющим пространством U, и $x \in E \setminus 0$. Предположим, что $x = y + z$. Угол между x и Π совпадает с углом между x и y.*

Доказательство. Пусть $t \in U$. В силу неравенства Коши-Буняковского $(y,t) \leqslant \|t\|\|y\|$. Деля на $\|t\|\|x\|$ получаем

$$\frac{(y,t)}{\|x\|\,\|t\|} \leqslant \frac{\|y\|}{\|x\|} = \frac{(y,y)}{\|x\|\,\|y\|}.$$

Но $z \in U^{\perp}$, и поэтому $(y,t) = (x,t)$, $(y,y) = (x,y)$. Таким образом,

$$\cos\psi = \frac{(x,t)}{\|x\|\,\|t\|} \leqslant \frac{(x,y)}{\|x\|\,\|y\|} = \cos\varphi.$$ □

Матрица $Q \in \mathrm{Mat}(n,\mathbb{R})$ *ортогональна*, если $Q^{-1} = {}^tQ$. Матрица $U \in \mathrm{Mat}(n,\mathbb{C})$ *унитарна*, если $U^{-1} = \overline{{}^tU}$.

Теорема 7.18. *Пусть* $\mathbf{e}$ — *ортонормированный базис в евклидовом (эрмитовом) пространстве E, и $(a_1,\ldots,a_n) = \mathbf{e}Q$. Следующие условия эквивалентны:*

1) *матрица Q ортогональна (унитарна);*

2) $a_1,\ldots,a_n$ — *ортонормированный базис E.*

Доказательство. Пусть $Q = (q_{ij})$. Имеем

$$a_i = e_1 q_{1i} + \ldots + e_n q_{ni}, \quad i = 1, \ldots, n.$$

Поэтому в силу теоремы 7.9 имеем $G(a_1, \ldots, a_n) = {}^t Q\overline{Q}$. Следовательно, если Q матрица ортогональна (унитарна), то

$$G(a_1, \ldots, a_n) = E_n,$$

т. е. система $a_1, \ldots, a_n$ — ортонормированный базис E. Верно и обратное. □

Предложение 7.19. *Определитель унитарной матрицы равен по модулю* 1. *Произведение ортогональных (унитарных) матриц ортогонально (унитарно). Обратная к ортогональной (унитарной) матрице ортогональна (унитарна).*

Через $\mathrm{O}(n, \mathbb{R})$ обозначается множество всех $n \times n$ ортогональных матриц, а через $\mathrm{SO}(n, \mathbb{R})$ обозначается подмножество всех $n \times n$ ортогональных матриц с определителем 1. Через $\mathrm{U}(n, \mathbb{C})$ обозначается всех $n \times n$ унитарных матриц, а $\mathrm{SU}(n, \mathbb{C})$ обозначается подмножество всех $n \times n$ унитарных матриц с определителем 1.

Теорема 7.20. *Комплексная матрица U размера* 2 *принадлежит* $\mathrm{SU}(2, \mathbb{C})$ *тогда и только тогда, когда*

$$U = \begin{pmatrix} z_1 & -z_2 \\ \overline{z_2} & \overline{z_1} \end{pmatrix}, \quad z_1, z_2 \in \mathbb{C}, \quad |z_1|^2 + |z_2|^2 = 1. \tag{7.16}$$

В частности, $Q \in \mathrm{SO}(2, \mathbb{R})$ *тогда и только тогда, когда*

$$Q = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Доказательство. Пусть

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SU}(2, \mathbb{C}), \quad a, b, c, d \in \mathbb{C}, \quad ac - bd = 1.$$

Тогда

$$U^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = {}^t\overline{U} = \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix}.$$

Отсюда $\overline{a} = d$, $b = -\overline{c}$. Полагая $a = z_1$, $b = -z_2$, получаем (7.16), поскольку $1 = \det U = |z_1|^2 + |z_2|^2$. Непосредственная проверка показывает, что любая матрица вида (7.16) лежит в $SU(2, \mathbb{C})$.

В частности, если $U = Q \in SO(2, \mathbb{R})$, то $z_1, z_2 \in \mathbb{R}$, $z_1^2 + z_2^2 = 1$. Поэтому найдется такой угол φ, что $z_1 = \cos\varphi$, $z_2 = \sin\varphi$. Отсюда получаем последнее утверждение. $\square$

ГЛАВА 8

ЛИНЕЙНЫЕ ОПЕРАТОРЫ

Эта глава знакомит читателя с основой теории линейных операторов. Излагаются способы нахождения собственных векторов и собственных значений, канонический вид симметрических и ортогональных операторов. Следует отметить, что симметрические операторы являются одним из основных объектов в квантовой механике, поскольку они интерпретируются как физические величины, значения которых совпадают с собственными значениями этих операторов.

§ 1. Матрицы линейных операторов

Отображение $\mathcal{A} : V \to V$ линейного пространства V в себя называется *линейным оператором*, если $\mathcal{A}(\alpha x + \beta y) = \alpha \mathcal{A} x + \beta \mathcal{A} y$ для всех $x, y \in V$ и всех скаляров α, β.

Пусть $\mathbf{e}$ — базис в V из (6.5), и

$$\mathcal{A}(e_j) = e_1 a_{1j} + \ldots + e_n a_{nj}, \quad j = 1, \ldots, n. \tag{8.1}$$

Тогда $A = (a_{ij}) \in \mathrm{Mat}(n)$ называется *матрицей* оператора $\mathcal{A}$ в базисе $\mathbf{e}$.

Примеры. Приведем ряд примеров линейных операторов.

1) V — пространство многочленов, $\mathcal{A}(f) = \dfrac{df}{dx}$.

2) V — трехмерное пространство $\mathbb{R}^3$, $\mathcal{A}$ — поворот на некоторой угол вокруг фиксированной оси, проходящей через начало координат.

3) V — трехмерное пространство $\mathbb{R}^3$, $\mathcal{A}$ — растяжение (сжатие) по трем перпендикулярным осям со своими коэффициентами.

4) V — трехмерное пространство $\mathbb{R}^3$, $\mathcal{A}$ — симметрия относительно некоторой плоскости, проходящей через начало координат (центральная симметрия относительно начала координат).

Предложение 8.1. *Пусть $x = \mathbf{e}X \in V$ — вектор со столбцом координат X в базисе $\mathbf{e}$ из* (6.5). *Тогда столбец координат вектора $\mathcal{A}x$ в том же базисе $\mathbf{e}$ равен AX, где A — матрица $\mathcal{A}$ в базисе $\mathbf{e}$. Пусть $\mathbf{e}' = \mathbf{e}C$ — новый базис V с матрицей перехода C. Тогда матрица $\mathcal{A}$ в базисе $\mathbf{e}'$ равна $C^{-1}AC$. В частности,* $\det A = \det(C^{-1}AC)$.

Доказательство. Объединяя систему равенств (8.1) получаем

$$\mathcal{A}(\mathbf{e}) = \mathbf{e}A. \tag{8.2}$$

Поэтому $\mathcal{A}x = \mathcal{A}(\mathbf{e}X) = \mathcal{A}(\mathbf{e})X = \mathbf{e}AX$. Отсюда вытекает первое утверждение.

Докажем второе утверждение. Имеем

$$\mathcal{A}(\mathbf{e}') = \mathcal{A}(\mathbf{e}C) = \mathcal{A}(\mathbf{e})C = \mathbf{e}AC = \mathbf{e}'C^{-1}AC,$$

откуда по (8.2) получаем второе утверждение. □

Пусть $\mathcal{A}, \mathcal{B}$ — линейные операторы и λ — скаляр. Определим *сумму, произведение линейных операторов и умножение оператора на скаляр*, полагая для любого вектора x

$$(\mathcal{A} + \mathcal{B})(x) = \mathcal{A}(x) + \mathcal{B}(x), \quad (\mathcal{A}\mathcal{B})(x) = \mathcal{A}(\mathcal{B}(x)),$$
$$(\lambda\mathcal{A})(x) = \lambda(\mathcal{A}(x)).$$

Линейный оператор $\mathcal{E}$ *тождествен*, если $\mathcal{E}(x) = x$ для всех векторов x. Линейный оператор $\mathcal{O}$ *нулевой*, если $\mathcal{O}(x) = 0$ для всех векторов x.

Предложение 8.2. *Сумма и произведение линейных операторов является линейным оператором. Матрица суммы и произведения операторов равна сумме и произведению матриц слагаемых (сомножителей). Произведение оператора на скаляр является линейным оператором, матрица которого получается умножением матрицы оператора на этот скаляр. Пусть* $\mathbf{e}$ — *базис пространства из* (6.5), *и* A, B — *матрицы операторов* $\mathcal{A}$, $\mathcal{B}$ *в базисе* $\mathbf{e}$. *Тогда* $A + B$, AB, λA — *соответственно, матрицы операторов* $\mathcal{A} + \mathcal{B}$, $\mathcal{A}\mathcal{B}$, $\lambda\mathcal{A}$ *в базисе* $\mathbf{e}$.

Доказательство состоит в непосредственном вычислении.

В силу предложения 8.2 можно говорить о многочленах от линейных операторов. Именно, если задан многочлен

$$f = a_0 + a_1 X + \ldots + a_n X^n,$$

то полагаем

$$f(\mathcal{A}) = a_0 \mathcal{E} + a_1 \mathcal{A} + \ldots + a_n \mathcal{A}^n.$$

Пусть $\mathbf{e}$ — произвольный базис пространства из (6.5) и A — матрица оператора $\mathcal{A}$. Тогда в базисе $\mathbf{e}$ матрица оператора $f(\mathcal{A})$ равна $f(A)$.

Ядро $\ker \mathcal{A}$ — это множество всех таких векторов x, что $\mathcal{A}(x) = 0$. *Образ* $\operatorname{Im} \mathcal{A}$ линейного оператора $\mathcal{A}$ — это множество всех векторов вида $\mathcal{A}(y)$ для всех векторов y из линейного пространства.

Предложение 8.3. $\ker \mathcal{A}$, $\operatorname{Im} \mathcal{A}$ *являются подпространствами в* V. *Если* x — *вектор из* V *со столбцом координат* X *в базисе* $\mathbf{e}$ *из* (6.5), *то*

$$x \in \ker \mathcal{A} \iff AX = 0,$$

где A — *матрица* $\mathcal{A}$ *в базисе* $\mathbf{e}$.

Доказательство. Пусть $x, y \in \ker \mathcal{A}$. Тогда для любых скаляров α, β имеем

$$\mathcal{A}(\alpha x + \beta y) = \alpha(\mathcal{A}(x)) + \beta(\mathcal{A}(y)) = 0,$$

т. е. $\alpha x + \beta y \in \ker A$.

Поэтому $\ker \mathcal{A}$ является подпространством. Аналогично проверяется, что $\operatorname{Im} \mathcal{A}$ является подпространством. Последнее утверждение вытекает из предложения 8.1. $\square$

Следствие 8.4. $\dim \ker \mathcal{A} = \dim V - r(A)$.

Доказательство. Необходимо применить предложение 8.3 и теорему 6.12. $\square$

Предложение 8.5. $\dim \operatorname{Im} \mathcal{A} = r(A)$.

Доказательство. Пусть r — ранг матрицы A и предположим, что первые r столбцов матрицы A линейно независимы. Тогда последние столбцы A, начиная с $(r+1)$-го линейно выражаются через первые r столбцов. Это означает, что $\operatorname{Im} A$ совпадает с линейной оболочкой $\langle \mathcal{A}(e_1), \ldots, \mathcal{A}(e_r) \rangle$, причем векторы $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_r)$ независимы. Поэтому $\dim(\operatorname{Im} A) = r$. $\square$

Линейный оператор $\mathcal{A}^{-1}$ называется *обратный* к оператору $\mathcal{A}$, если $\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{E}$.

Теорема 8.6. *Пусть* $\mathbf{e}$ — *базис в* V *из* (6.5), *и* A — *матрица* $\mathcal{A}$ *в базисе* $\mathbf{e}$. *Тогда матрица* $\mathcal{A}^{-1}$ *в базисе* $\mathbf{e}$ *равна* A^{-1}. *Кроме того, следующие условия эквивалентны*:

1) *существует* $\mathcal{A}^{-1}$;

2) $\ker \mathcal{A} = 0$;

3) $\operatorname{Im} \mathcal{A} = V$;

4) $\det A \neq 0$.

Доказательство. Из предложения 8.2 вытекает, что матрица $\mathcal{A}^{-1}$ в базисе $\mathbf{e}$ равна A^{-1}. Поэтому условия 1) и 4) эквивалентны. Кроме того, в силу предложения 8.3 и предложения 8.5 условия 2) и 3) эквивалентны.

Если выполнено условие 1) и $x \in \ker \mathcal{A}$, то

$$x = \mathcal{A}^{-1}\mathcal{A}(x) = \mathcal{A}^{-1}(0) = 0.$$

Обратно, если выполнено условие 2), то по следствию 8.4 ранг матрицы A совпадает с размерностью пространства. Следовательно, $\det A \neq 0$ по теореме 6.1. □

§2. Инвариантные подпространства

Подпространство U в пространстве V *инвариантно* относительно линейного оператора $\mathcal{A}$, если $\mathcal{A}x \in U$ для всех $x \in U$.

Пример. Пусть V — пространство всех многочленов, U — подпространство всех многочленов степени не выше n, $\mathcal{A}$ — оператор дифференцирования. Тогда U инвариантно относительно $\mathcal{A}$.

Предложение. *Пусть $\mathcal{A}$ — линейный оператор в пространстве V, и $\mathbf{e}$ — базис V из* (6.5), *причем* $(e_1, \ldots, e_k)$ — *базис подпространства U. Пусть A — матрица $\mathcal{A}$ в базисе $\mathbf{e}$. Следующие условия эквивалентны:*

1) *U инвариантно относительно $\mathcal{A}$;*

2) $A = \left(\begin{array}{c|c} A' & A'' \\ \hline 0 & A''' \end{array}\right)$, *где A', A'', A''' — матрицы, имеющие, соответственно, размеры $k \times k$, $k \times (n-k)$, $(n-k) \times (n-k)$.*

Доказательство. Пусть U инвариантно. Тогда для $i = 1, \ldots$ $\ldots, k$ справедливо включение

$$\mathcal{A}(e_i) \in U = \langle e_1, \ldots, e_k \rangle. \tag{8.3}$$

Отсюда следует второе свойство.

Обратно, если A имеет указанный вид, то выполнено (8.3). Тогда $\mathcal{A}(x) \in U$ для любого $x = x_1 e_1 + \ldots + x_k e_k \in U$. □

Следствие. *Пусть $\mathcal{A}$ — линейный оператор в пространстве V, причем $V = U \oplus W$, где U, W — инвариантные подпространства в V. Пусть $\mathbf{e}$ — базис V из* (6.5), *причем* $(e_1, \ldots, e_k)$ — *базис подпространства U, а* $(e_{k+1}, \ldots, e_n)$ — *базис подпространства W. Тогда в базисе $\mathbf{e}$ матрица оператора $\mathcal{A}$ имеет вид*

$$A = \left(\begin{array}{c|c} A' & 0 \\ \hline 0 & A'' \end{array}\right), \quad A' \in \mathrm{Mat}(k), \quad A'' \in \mathrm{Mat}(n-k).$$

§ 3. Собственные векторы и собственные значения

Ненулевой вектор x называется *собственным* для линейного оператора $\mathcal{A}$ *с собственным значением* λ, если $\mathcal{A}(x) = \lambda x$.

Упражнение. Пусть $\mathcal{A}$ — линейный оператор в пространстве V. Доказать, что

1) ненулевой вектор $x \in V$ является собственным для $\mathcal{A}$ тогда и только тогда, когда $\langle a \rangle$ инвариантно относительно $\mathcal{A}$;

2) множество V_λ, состоящее из нуля и всех собственных векторов для оператора $\mathcal{A}$ с фиксированным собственным значением λ совпадает с подпространством $\ker(\mathcal{A} - \lambda\mathcal{E})$, где $\mathcal{E}$ — тождественный оператор; в частности, V_λ является подпространством.

Характеристическим многочленом матрицы A называется многочлен $\chi_A(t) = \det(A - tE)$.

Несложная проверка показывает, что

$$\chi_A(t) = (-t)^n + (-t)^{n-1}\operatorname{tr} A + \ldots + \det A. \tag{8.4}$$

Предложение. $\chi_{C^{-1}AC}(t) = \chi_A(t)$.

Доказательство.

$$\begin{aligned}\chi_{C^{-1}AC}(t) &= \det(C^{-1}AC - tE) = \det(C^{-1}(A - tE)C) = \\ &= \det C^{-1}\det(A - tE)\det C = \det(A - tE) = \\ &= \chi_A(t).\end{aligned}$$ □

Следствие 8.7. *Корректно определен характеристический многочлен $\chi_{\mathcal{A}}(t)$ линейного оператора $\mathcal{A}$.*

Доказательство. Воспользоваться предложением 8.1. □

Укажем способ вычисления собственных векторов и собственных значений. Пусть $\mathbf{e}$ — базис пространства V из (6.5), и A — матрица $\mathcal{A}$ в этом базисе. Пусть $x = \mathbf{e}X$ — собственный вектор $\mathcal{A}$ с собственным значением λ. Тогда $AX = \lambda X$, или

$$(A - \lambda E)X = 0. \tag{8.5}$$

Итак, система линейных однородных уравнений (8.5) с матрицей $A - \lambda E$ имеет ненулевое решение X. Это эквивалентно тому, что $\chi_A(\lambda) = \det(A - \lambda E) = 0$. Итак, справедлива

Теорема. *Собственными значениями $\mathcal{A}$ являются корни $\chi_{\mathcal{A}}(t)$ и только они.*

Если известно собственное значение λ оператора $\mathcal{A}$, то для нахождение собственного вектора $x = \mathbf{e}X$ с собственным значением λ нужно решить систему линейных однородных уравнений (8.5).

Теорема 8.8. *Пусть $\mathcal{A}$ — линейный оператор в пространстве V. Собственные векторы для $\mathcal{A}$ с разными собственными значениями линейно независимы.*

Доказательство. Пусть $\mathcal{A}(a_i) = \lambda_i a_i$, $i = 1, \ldots, k$, причем $\lambda_i \neq \lambda_j$ при $i \neq j$. Пусть

$$\mu_1 a_1 + \ldots + \mu_k a_k = 0. \tag{8.6}$$

Если $k = 1$, то утверждение очевидно, поскольку вектор $a_1 \neq 0$.

Пусть для $k - 1$ утверждение доказано. Применим в (8.6) оператор $\mathcal{A}$. Получаем

$$\lambda_1 \mu_1 a_1 + \ldots + \lambda_k \mu_k a_k = 0. \tag{8.7}$$

Вычтем из (8.7) равенство (8.6), умноженное на λ_k. Получаем

$$(\lambda_1 - \lambda_k)\mu_1 a_1 + \ldots + (\lambda_{k-1} - \lambda_k)\mu_{k-1} a_{k-1} = 0. \tag{8.8}$$

По индукции для любого $i = 1, \ldots, k-1$ имеем $(\lambda_i - \lambda_k)\mu_i = 0$, откуда $\mu_i = 0$, поскольку $\lambda_i \neq \lambda_k$. Возвращаясь к (8.6), получаем $\mu_k a_k = 0$ и $\mu_k = 0$. $\square$

Следствие 8.9. *Пусть все собственные значения $\lambda_1, \ldots, \lambda_n$ линейного оператора $\mathcal{A}$ различны. Тогда в пространстве существует базис, в котором матрица оператора имеет диагональный вид*

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 \\ \hdotsfor{5} \\ 0 & 0 & \ldots & \lambda_{n-1} & 0 \\ 0 & 0 & \ldots & 0 & \lambda_n \end{pmatrix}.$$

Доказательство. Пусть e_i — собственный вектор $\mathcal{A}$ с собственным значением λ_i. По теореме 8.8, $\mathbf{e}$ — базис пространства из (6.5). Остается воспользоваться определением матрицы оператора (8.1). □

Упражнение. Пусть λ — собственное значение линейного оператора $\mathcal{A}$, действующего в конечномерном пространстве. Тогда максимальное число независимых собственных векторов с собственным значением λ не превосходит кратности λ как корня характеристического многочлена $\chi_{\mathcal{A}}(t)$.

Если $\mathcal{A}$ — линейный оператор в конечномерном комплексном пространстве, то он всегда имеет собственный вектор. Действительно, характеристический многочлен этого оператора всегда имеет корень, которому соответствует собственный вектор.

Теорема 8.10. *Пусть $\mathcal{A}$ — линейный оператор в конечномерном вещественном пространстве V. Тогда оператор $\mathcal{A}$ имеет либо одномерное, либо двумерное инвариантное подпространство.*

Доказательство. Выберем в пространстве V произвольный базис и рассмотрим матрицу A линейного оператора $\mathcal{A}$. Так как все степени матрицы A находятся в n^2-мерном пространстве матриц, то они зависимы. Поэтому существует такой ненулевой многочлен $f(t) \in \mathbb{R}[t]$, что $f(A) = 0$. Отсюда $f(\mathcal{A}) = 0$. По теореме 5.14 $f(t) = p_1(t) \cdots p_m(t)$, где $p_i(t) \in \mathbb{R}[t]$ имеет степень $\leqslant 2$. Можно считать, что старшие коэффициенты всех $p_i(t)$ равны 1. Пусть $x \in E \setminus 0$, и $p_{m+1} = 1$. Тогда $f(\mathcal{A})x = 0$. Следовательно, существует такое $i = 1, \ldots, m$, что $y = p_{i+1}(\mathcal{A}) \cdots p_m(\mathcal{A})x \neq 0$, но $p_i(\mathcal{A})y = 0$.

Пусть $\deg p_i(t) = 1$, т. е. $p_i(t) = \alpha t + \beta$. Тогда $\mathcal{A}y = -\frac{\beta}{\alpha}y$, т. е. y — собственный вектор для $\mathcal{A}$, и поэтому одномерная линейная оболочка $\langle y \rangle$ является инвариантным подпространством.

Пусть $\deg p_i(t) = 2$, т. е. $p_i(t) = t^2 + \beta t + \gamma$. Остается показать, что подпространство $\langle y, \mathcal{A}y \rangle$ инвариантно относительно $\mathcal{A}$.

По условию $\mathcal{A}^2 y = -\beta \mathcal{A}y - \gamma y$. Поэтому, если $z = \mu y + \xi \mathcal{A}(y) \in \langle y, \mathcal{A}y \rangle$, то

$$\mathcal{A}(z) = \mu \mathcal{A}(y) + \xi \mathcal{A}^2(y) = -\xi\gamma y + (\mu - \xi\beta)\mathcal{A}(y). \qquad \square$$

§ 4. Симметрические операторы

В этом разделе рассматриваются линейные операторы в евклидовом (эрмитовом) пространстве E конечной размерности.

Предложение 8.11. *Пусть* e *ортонормированный базис в E из* (6.5). *Предположим, что $\mathcal{A}$ — линейный оператор в E с матрицей $A = (a_{ij})$ в базисе* e. *Тогда* $(e_i, \mathcal{A}(e_j)) = \overline{a_{ij}}$, $(\mathcal{A}(e_i), e_j) = a_{ji}$.

Доказательство. В силу (8.1) справедливо равенство

$$(e_i, \mathcal{A}(e_j)) = \sum_{s=1}^{n} \overline{a_{sj}}(e_i, e_s) = \overline{a_{ij}}.$$

Кроме того, $(\mathcal{A}(e_i), e_j) = \overline{(e_j, \mathcal{A}(e_i))} = a_{ji}$. $\square$

Линейный оператор $\mathcal{A}$ в E *симметричен*, или *самосопряжен*, если $(\mathcal{A}(x), y) = (x, \mathcal{A}(y))$ для всех $x, y \in E$.

Теорема 8.12. *Пусть* e *ортонормированный базис из* (6.5) *в евклидовом (эрмитовом) пространстве E, и $\mathcal{A}$ — линейный оператор в E с матрицей $A = (a_{ij})$ в базисе* e. *Следующие условия эквивалентны*:

1) *оператор $\mathcal{A}$ симметричен*;

2) *матрица A симметрична (эрмитова), т. е. $a_{ij} = \overline{a_{ji}}$ для всех i, j.*

Доказательство. Условие 1) влечет условие 2) по предложению 8.11. Обратно, пусть выполнено условие 2). Если $x = x_1e_1 + \ldots + x_ne_n$ и $y = y_1e_1 + \ldots + y_ne_n$, то

$$(\mathcal{A}(x), y) = \sum_{i,j} x_i\overline{y_j}(\mathcal{A}(e_i), e_j) = \sum_{i,j} x_i\overline{y_j}a_{ji}.$$

Аналогично, $(x, \mathcal{A}(y)) = \sum_{i,j} x_i\overline{y_j\overline{a_{ij}}}$. Из условия 2) вытекает условие 1). □

Предложение 8.13. *Собственные значения симметрического оператора в эрмитовом пространстве вещественны.*

Доказательство. Пусть $\mathcal{A}(x) = \lambda x$, $x \neq 0$. Тогда

$$\lambda(x, x) = (\mathcal{A}(x), x) = (x, \mathcal{A}(x)) = \overline{\lambda}(x, x).$$

Так как $(x, x) > 0$, то $\lambda = \overline{\lambda}$ и $\lambda \in \mathbb{R}$. □

В теореме 8.16 ниже мы покажем, что утверждение предложения 8.13 справедливо и для самосопряженного оператора в евклидовом пространстве.

Предложение 8.14. *Собственные векторы симметричного оператора с разными собственными значениями ортогональны.*

Доказательство. Пусть $\mathcal{A}(x) = \lambda x$, $\mathcal{A}(y) = \mu y$, где $\lambda \neq \mu$. Отсюда $\lambda(x, y) = (\mathcal{A}(x), y) = (x, \mathcal{A}(y)) = \mu(x, y)$, и поэтому $(x, y) = 0$. □

Предложение 8.15. *Пусть подпространство U в E инвариантно относительно симметричного оператора. Тогда $U^\perp$ также инвариантно относительно этого оператора.*

Доказательство. Пусть $x \in U^\perp$ и $y \in U$. Тогда $(\mathcal{A}(x), y) = (x, \mathcal{A}(y)) = 0$, поскольку $\mathcal{A}(y) \in U$ по предположению. Отсюда $\mathcal{A}(x) \in U^\perp$. □

Теорема 8.16. *Пусть $\mathcal{A}$ — симметричный оператор в E. Тогда в E существует собственный ортонормированный базис.*

Доказательство. Если $\dim E = 1$, то утверждение очевидно.

Пусть $n = \dim E \geqslant 2$. Покажем сначала, что в E существует собственный вектор для $\mathcal{A}$. Если это не так, то E — евклидово пространство. По теореме 8.10 в E в этом случае существует двумерное инвариантное подпространство U. Выберем в U произвольный ортонормированный базис $\mathbf{f} = (f_1, f_2)$. В этом базисе матрица оператора $\mathcal{A}\,|_U$ по теореме 8.12 имеет вид

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Тогда $\chi_A(t) = \det(A - tE) = t^2 - (a+c)t + (ac - b^2)$, причем дискриминант $(a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 \geqslant 0$. Поэтому у матрицы A есть собственное значение и в U существует собственный вектор e_1 длины 1.

Итак, во всех случаях в V для симметрического оператора $\mathcal{A}$ существует собственный вектор e_1 длины 1. В силу предложения 8.15 $\langle e_1 \rangle^{\perp}$ инвариантно и имеет размерность $n-1$. Остается воспользоваться индукцией по размерности пространства. □

Следствие 8.17. *Пусть A — симметричная (эрмитова) матрица. Тогда существует такая ортогональная (унитарная) матрица Q, что*

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_{n-1} & 0 \\ 0 & 0 & \ldots & 0 & \lambda_n \end{pmatrix}. \tag{8.9}$$

Теорема 8.18 (Приведение квадратичных функций к главным осям). *Пусть задана билинейная симметричная (эрмитова) функция $b(x, y)$ в евклидовом (эрмитовом) пространстве, имеющая в ортонормированном базисе матрицу Грама B. Тогда существует такой ортонормированный базис, в котором матрица функции $b(x, y)$ имеет диагональный вид, причем по главной диагонали стоят собственные значения матрицы B.*

Доказательство. Пусть $\mathbf{e}$ — произвольный ортонормированный базис евклидова (эрмитова) пространства E из (6.5), и B — матрица Грама функции $b(x,y)$ в этом базисе. Тогда ${}^t\overline{B} = B$ по предложению 7.1. Рассмотрим линейный оператор $\mathcal{B}$, который в базисе $\mathbf{e}$ задается матрицей $\overline{B}$. По теореме 8.12 оператор $\mathcal{B}$ симметричен. По теореме 8.16 существует собственный ортонормированный базис $\mathbf{f} = \mathbf{e}U$ для оператора $\mathcal{B}$, причем по теореме 7.18 матрица U ортогональна (унитарна). В то же время матрица $U^{-1}\overline{A}U$ диагональна, и на на главной диагонали расположены вещественный числа, являющиеся собственными значениями векторов базиса $\mathbf{f}$. В новом базисе $\mathbf{f}$ матрица Грама функции b имеет вид

$$ {}^tU\overline{BU} = \overline{U}^{-1}\overline{BU} = \overline{U^{-1}BU} = U^{-1}BU, $$

поскольку $U^{-1}BU$ — вещественная диагональная матрица. Отсюда вытекает утверждение теоремы. □

§ 5. Ортогональные и унитарные операторы

Линейный оператор $\mathcal{Q}$ в евклидовом (эрмитовом) пространстве E *ортогонален* (унитарен), если он сохраняет длины векторов, т. е. $\|\mathcal{Q}(x)\| = \|x\|$ для всех $x \in E$.

Предложение 8.19. *Произведение ортогональных (унитарных) операторов ортогонально (унитарно). Обратный к ортогональному (унитарному) оператору тоже ортогонален (унитарен). Собственные значения ортогонального (унитарного) оператора равны по модулю* 1.

Доказательство. Пусть $\mathcal{Q}(x) = \lambda x$. Тогда

$$ \|\mathcal{Q}(x)\| = |\lambda|\,\|x\| = \|x\|, $$

откуда $|\lambda| = 1$. □

Предложение. *Ортогональный (унитарный) оператор $\mathcal{Q}$ сохраняет скалярные произведения, т. е. $(x,y) = (\mathcal{Q}x, \mathcal{Q}y)$ для всех $x, y \in E$.*

Доказательство. Если Ω сохраняет скалярное произведение, то этот оператор сохраняет и длины векторов. Обратно, пусть линейный оператор Ω сохраняет длины векторов. Тогда совпадают квадратичные функции (x,x) и $(\Omega x, \Omega x)$, которые происходят из симметрических билинейных (эрмитовых) функций (x,y) и $(\Omega x, \Omega y)$. В силу предложения 7.3 получаем, что

$$(x,y) = (\Omega x, \Omega y). \qquad \square$$

Предложение 8.20. *Пусть Ω — ортогональный (унитарный) оператор в евклидовом (эрмитовом) пространстве E, и U — инвариантное подпространство в E. Тогда ортогональное дополнение $U^{\perp}$ также инвариантно относительно Ω.*

Доказательство. Оператор Ω имеет нулевое ядро и поэтому он отображает U на U в силу теоремы 8.6.

Пусть $x \in U^{\perp}$ и $y \in U$. Тогда существует такое $z \in U$, что $y = U(z)$. Следовательно, $(\Omega(x), y) = (\Omega(x), \Omega(z)) = (x,z) = 0$, поскольку $x \in U^{\perp}$, $z \in U$. $\square$

Предложение 8.21. *Собственные векторы ортогонального оператора с разными собственными значениями ортогональны.*

Доказательство. Пусть $\Omega(x) = \lambda x$, $\Omega(y) = \omega y$, причем $\lambda \neq \omega$. Тогда $(x,y) = (\Omega(x), \Omega(y)) = \lambda\overline{\omega}(x,y)$. Так как $|\omega| = 1$ по предложению 8.19, то $\overline{\omega} = \omega^{-1}$. Поэтому $\lambda\overline{\omega} = \lambda\omega^{-1} \neq 1$ по предположению. Отсюда получаем $(x,y) = 0$. $\square$

Предложение. *Пусть* $\mathbf{e}$ *— ортонормированный базис в E из* (6.5), *и Ω — линейный оператор в E с матрицей Q. Следующие условия эквивалентны:*

1) *оператор Ω ортогонален (унитарен);*

2) *матрица Q ортогональна (унитарна).*

Доказательство. В силу определения матрицы линейного оператора имеем $\Omega(\mathbf{e}) = \mathbf{e}Q$. Если оператор Ω ортогонален (унитарен), то $\Omega(\mathbf{e}) = \mathbf{e}Q$ — ортонормированный базис, причем Q —

матрица перехода от базиса $\mathbf{e}$ к базису $\mathcal{Q}(\mathbf{e})$. Следовательно, матрица Q ортогональна (унитарна) по теореме 7.18.

Обратно, пусть матрица Q ортогональна (унитарна). Тогда $\mathcal{Q}(\mathbf{e}) = \mathbf{e}Q$ — ортонормированный базис. Если $x = x_1e_1 + \ldots + x_ne_n \in E$, то $\mathcal{Q}(x) = x_1\mathcal{Q}(e_1) + \ldots + x_n\mathcal{Q}(e_n)$, откуда

$$\|x\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2} = \|\mathcal{Q}(x)\|. \qquad \square$$

Теорема 8.22. *Пусть $\mathcal{Q}$ — ортогональный оператор в двумерном евклидовом пространстве E. Если* $\det \mathcal{Q} = 1$, *и e_1, e_2 — произвольный ортонормированный базис E, то в этом базисе матрица Q этого оператора имеет вид*

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Если $\det \mathcal{Q} = -1$, *то существует ортонормированный базис, в котором матрица оператора $\mathcal{Q}$ имеет вид*

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

Доказательство. Случай $\det \mathcal{Q} = 1$ вытекает из теоремы 7.20.

Пусть $\det \mathcal{Q} = -1$. Из вида характеристического многочлена (8.4) получаем, что

$$\chi_{\mathcal{Q}}(t) = t^2 - t(\operatorname{tr} Q) - 1$$

имеет положительный дискриминант $(\operatorname{tr} Q)^2 + 4$. Следовательно, оператор $\mathcal{Q}$ имеет два различных вещественных собственных значения, которые по предложению 8.19 равны ± 1. Остается применить предложение 8.21. $\square$

Теорема 8.23. *Пусть $\mathcal{Q}$ — ортогональный оператор в трехмерном евклидовом пространстве E. Тогда в E существует ортонормированный базис, в котором матрица $\mathcal{Q}$ имеет вид*

$$\begin{pmatrix} \det \mathcal{Q} & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Доказательство. Характеристический многочлен

$$\chi_{\mathcal{Q}}(t) = -t^3 + \ldots + \det \mathcal{Q}$$

степени 3 по теореме 5.14 имеет вещественный корень $\lambda = \pm 1$. Пусть e — соответствующий собственный вектор длины 1. Тогда $E = \langle a \rangle \oplus \langle a \rangle^{\perp}$, и $\dim \langle a \rangle^{\perp} = 2$. При этом плоскость $W = \langle a \rangle^{\perp}$ инвариантна по предложению 8.20. Если $\det \mathcal{Q}\,|_W = 1$, то по теореме 8.22 нужно взять произвольный ортонормированный базис e_2, e_3 в W и получить искомый базис e_1, e_2, e_3 в E.

Если $\det \mathcal{Q}\,|_W = -1$, то по теореме 8.22 в W существует базис e_2, e_3, где $\mathcal{Q}(e_2) = e_2$, $\mathcal{Q}(e_3) = -e_3$. Если $\lambda = 1$, то базис e_3, e_1, e_2 — искомый. Если $\lambda = -1$, то базис e_2, e_1, e_3 — искомый. □

Теорема. *Пусть $\mathcal{Q}$ — ортогональный оператор в E. Тогда в E существует ортонормированный базис, в котором матрица $\mathcal{Q}$ имеет блочно-диагональный вид, причем блоки имеют размер не выше 2. В блоках размера 1 стоят ± 1, а блоки размера 2 имеют вид*

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{8.10}$$

Доказательство проводится индукцией по $\dim E$ с использованием теоремы 8.22, предложений 8.19–8.21.

Теорема. *Пусть $\mathcal{U}$ — унитарный оператор в эрмитовом пространстве E. Тогда в E существует собственный ортонормированный базис.*

Доказательство. Оператор $\mathcal{U}$ в комплексном пространстве всегда имеет собственный вектор e_1. Можно считать, что $\|e_1\| = 1$. Тогда $E = \langle e_1 \rangle \oplus \langle e_1 \rangle^{\perp}$, причем $\langle e_1 \rangle^{\perp}$ инвариантно относительно $\mathcal{U}$ по предложению 8.20. По индукции в $\langle e_1 \rangle^{\perp}$ существует собственный ортонормированный базис $e_2, \ldots, e_n$. Тогда $e_1, e_2, \ldots, e_n$ — искомый базис в E. □

Следствие. *Пусть U — унитарная матрица. Тогда существует такая унитарная матрица Q, что*

$$Q^{-1}UQ = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

где $\lambda_1, \dots, \lambda_n$ — комплексные числа, по модулю равные 1.

Теорема (Теорема о полярном разложения). *Пусть $\mathcal{A}$ — линейный оператор в евклидовом (эрмитовом) пространстве. Тогда $\mathcal{A} = \mathcal{Q}\mathcal{B}$, где $\mathcal{Q}$ — ортогональный (унитарный) оператор, $\mathcal{B}$ — симметричный оператор с неотрицательными собственными значениями. При этом $\mathcal{B}^2 = \mathcal{A}^*\mathcal{A}$.*

Доказательство изложено в [3, гл. 6, § 3, теорема 4] и в [6, гл. 3, § 3, теорема 15].

ГЛАВА 9

КРИВЫЕ И ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА

В этой главе классифицируются поверхности второго порядка, или квадрики, в евклидовых пространствах. В качестве приложения излагается полная классификация кривых второго порядка в двумерной плоскости и классификация поверхностей второго порядка в трехмерном евклидовом пространстве.

§ 1. Движения евклидова пространства

Расстоянием между векторами x и y в евклидовом пространстве E называется длина вектора $x - y$. *Движением* или *изометрией* евклидова пространства E называется отображение $\Phi : E \to E$, сохраняющее расстояние между любыми двумя векторами, т. е. $\|\Phi(x) - \Phi(y)\| = \|x - y\|$ для всех $x, y \in E$.

Примерами движений являются *сдвиги* на фиксированный вектор a, т. е. отображения $x \mapsto x + a$.

Предложение. *Композиция движений является движением.*

Доказательство. Пусть Φ, Ψ — движения. Тогда для любых $x, y \in E$ получаем

$$\|\Phi\Psi(x) - \Phi\Psi(y)\| = \|\Phi(\Psi(x)) - \Phi(\Psi(y))\| = \\ = \|\Psi(x) - \Psi(y)\| = \|x - y\|. \quad \square$$

Теорема. *Отображение $\Phi : E \to E$ является движением тогда и только тогда, когда существуют такой линейной ортогональный оператор φ и вектор b, что $\Phi(x) = \varphi(x)+b$ для любого $x \in E$. Другими словами, движение Φ является композицией ортогонального линейного оператора φ и сдвига на вектор b.*

Доказательство теоремы можно найти в [7, гл. 4, § 3].

§ 2. Квадрики

Квадрикой или *поверхностью второго порядка* в евклидовом пространстве E с базисом $\mathbf{e}$ из (6.5) называется множество всех таких векторов $x = e_1x_1 + \ldots + e_nx_n$, что

$$q(x) + 2(b_1x_1 + \ldots + b_nx_n) + c = 0, \tag{9.1}$$

где $q(x)$ — квадратичная функция, $b_1, \ldots, b_n$, c — вещественные числа.

Если $A = (a_{ij})$ симметричная матрица функции $q(x)$, то в силу (7.3) уравнение (9.1) имеет вид

$$\sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1\leqslant i<j\leqslant n} a_{ij}x_ix_j + 2(b_1x_1 + \ldots + b_nx_n) + c = 0. \tag{9.2}$$

Матрицей квадрики (9.1) назовем матрицу A квадратичной части $q(x)$. *Расширенной матрицей* квадрики (9.1) назовем матрицу

$$\widetilde{A} = \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} & b_n \\ b_1 & \ldots & b_n & c \end{pmatrix}.$$

Тогда уравнение (9.1) принимает вид

$$0 = {}^tXAX + 2\,bX + c = (x_1, \ldots, x_n, 1)\widetilde{A}\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}, \tag{9.3}$$

где

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad b = (b_1, \ldots, b_n).$$

Предложение. *Пусть в евклидовом пространстве E заданы два ортонормированных базиса $\mathbf{e}', \mathbf{e}$ с матрицей перехода C, т. е. $\mathbf{e}' = \mathbf{e}C$. Если $A, \widetilde{A}$ — матрицы квадрики в базисе $\mathbf{e}$, то в новом базисе $\mathbf{e}'$ они будут, соответственно, иметь вид*

$$A' = {}^tCAC, \quad \widetilde{A}' = {}^t\left(\begin{array}{c|c} C & 0 \\ \hline 0 & 1 \end{array}\right) \widetilde{A} \left(\begin{array}{c|c} C & 0 \\ \hline 0 & 1 \end{array}\right).$$

Если мы совершаем преобразование переноса начала координат

$$x_1 = x_1' + d_1, \ \ldots, \ x_n = x_n' + d_n,$$

то матрица A не меняется, а матрица $\widetilde{A}$ заменяется на матрицу

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ d_1 & d_2 & d_3 & \ldots & d_{n-1} & d_n & 1 \end{pmatrix} \widetilde{A} \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & d_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & d_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & d_n \\ 0 & 0 & 0 & \ldots & 0 & 0 & 1 \end{pmatrix}.$$

Доказательство. По предложению 6.5 столбец X' координат в новом базисе $\mathbf{e}'$ связан с X соотношением $X = CX'$. Поэтому, подставляя в (9.3), получаем требуемое утверждение. □

Теорема 9.1. *Функции $\det A$, $\operatorname{tr} A$, $\det \widetilde{A}$ и ранги $\operatorname{rk} A$, $\operatorname{rk} \widetilde{A}$ являются инвариантами квадрик при движении пространства, т. е. значения этих функций не меняются при замене переменных, связанных с движением пространства.*

Теорема 9.2. *Применяя движение евклидова пространства, уравнение квадрики в евклидовом пространстве можно привести к одному из следующих видов:*

1) $\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + \mu = 0, \quad \mu \neq 0$;

2) $\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 = 0$;

3) $\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + 2\mu x_{r+1} = 0$.

Доказательство. В силу теоремы 8.18 можно считать, что уравнение квадрики имеет вид

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + 2(\mu_1 x_1 + \ldots + \mu_n x_n) + c = 0,$$

где $\lambda_1, \ldots, \lambda_r \neq 0$. В этом случае, если $i = 1, \ldots, r$, то $\lambda_i x_i^2 + 2\mu_i x_i = \lambda_i \left(x_i + \frac{\mu_i}{\lambda_i}\right)^2 - \frac{\mu_i^2}{\lambda_i}$. Таким образом, совершая сдвиг, можно считать, что $\mu_1 = \ldots = \mu_r = 0$.

Если $\mu_{r+1} = \ldots = \mu_n = 0$, то уравнение имеет вид (1) или (2) из условий теоремы.

Пусть вектор $f = \mu_{r+1} e_{r+1} + \ldots + \mu_n e_n \neq 0$. Тогда вектор $g = \frac{f}{\|f\|}$ имеет длину 1, и он перпендикулярен векторам $e_1, \ldots, e_r$. Следовательно, систему $e_1, \ldots, e_r, g$ можно дополнить до ортонормированного базиса пространства E. В этом базисе уравнение имеет вид $\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + 2\mu x_{r+1} + c' = 0$, где $\mu \neq 0$. При этом $2\mu x_{r+1} + c' = 2\mu(x_{r+1} + \frac{c'}{2\mu})$. Совершая новый перенос, получаем уравнение вида (3). □

§ 3. Кривые второго порядка

В этом разделе, опираясь на теорему 9.2 мы приведем классификацию кривых второго порядка, т. е. квадрик в двумерном евклидовом пространстве.

Теорема. *Уравнение квадрики в двумерном евклидовом пространстве с помощью движения и, возможно, умножения уравнения на ненулевой коэффициент, приводится к одному из следующих видов*:

1) *эллипс* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ (рис. 9.1);

2) *мнимый эллипс* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$;

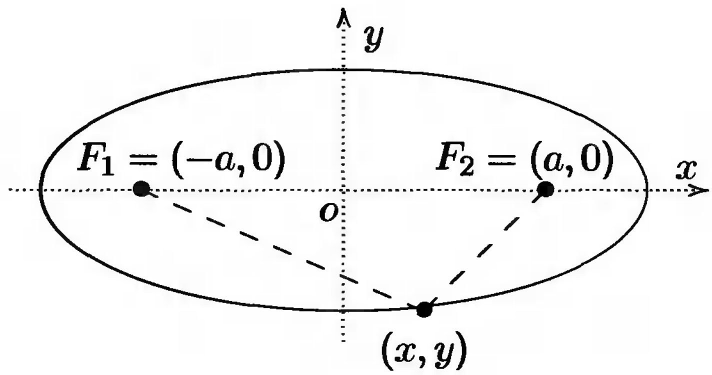


Рис. 9.1. Эллипс.

3) *гипербола* $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ (рис. 9.2);

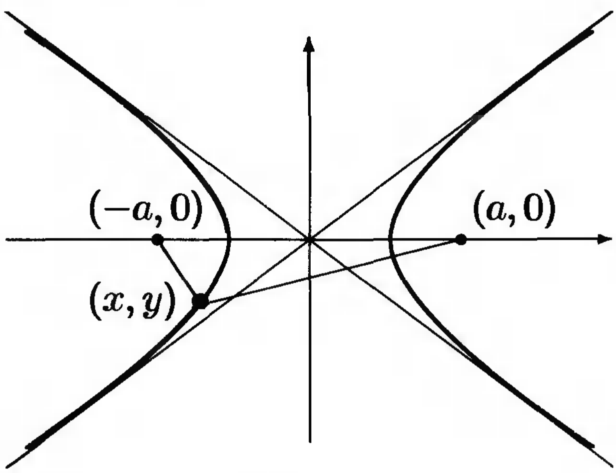


Рис. 9.2. Гипербола.

4) *пара пересекающихся прямых* $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$ (рис. 9.3);

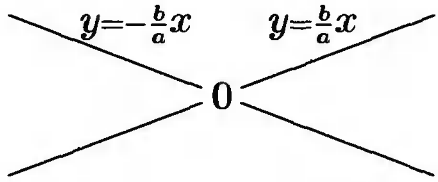


Рис. 9.3. Пара пересекающихся прямых.

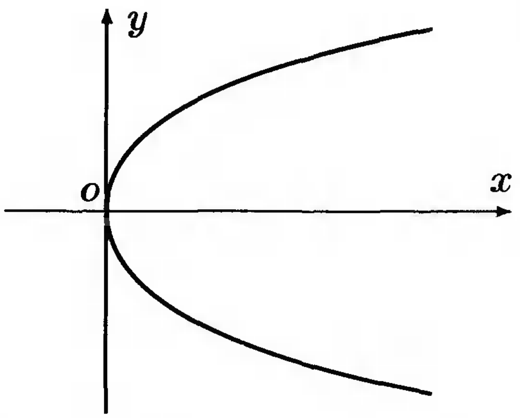


Рис. 9.4. Парабола при $p > 0$.

5) *пара мнимых пересекающихся прямых* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$;

6) *парабола* $y^2 = 2px$ (рис. 9.4);

7) *пара параллельных прямых* $y^2 = a^2$ (рис. 9.5);

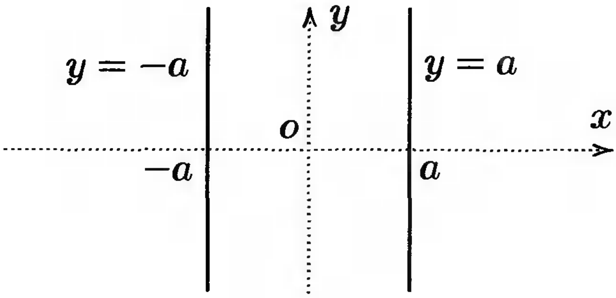


Рис. 9.5. Пара параллельных прямых.

8) *пара мнимых параллельных прямых* $y^2 = -a^2$;

9) *пара совпадающих прямых* $y^2 = 0$.

Приведём примеры, показывающие применение теоремы 9.1, позволяющее по уравнению квадрики определить вид кривой.

По теореме 9.2 с помощью движения уравнение кривой второго порядка приводится к одному из следующих канонических видов:

1) $\lambda_1 x^2 + \lambda_2 y^2 + \mu = 0, \quad \lambda_1, \lambda_2, \mu \neq 0;$

2) $\lambda_1 x^2 + \lambda_2 y^2 = 0, \quad \lambda_1, \lambda_2 \neq 0;$

3) $\lambda_1 x^2 + 2\mu y = 0, \quad \lambda_1, \mu \neq 0;$

4) $\lambda_1 x^2 + \mu = 0, \quad \lambda_1, \mu \neq 0;$

5) $\lambda_1 x^2 = 0, \quad \lambda_1 \neq 0.$

Для каждого из этих канонических видов матрицы квадратичных функций имеют, соответственно, вид

$$A_1 = A_2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad A_3 = A_4 = A_5 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Следовательно,

$$\det A_1 = \det A_2 = \lambda_1 \lambda_2, \quad \operatorname{tr} A_1 = \operatorname{tr} A_2 = \lambda_1 + \lambda_2,$$
$$\det A_3 = \det A_4 = \det A_5 = 0, \quad \operatorname{tr} A_3 = \operatorname{tr} A_4 = \operatorname{tr} A_5 = \lambda_1.$$

Расширенные матрицы имеют, соответственно, вид

$$\widetilde{A_1} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \mu \end{pmatrix}, \ \widetilde{A_2} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \widetilde{A_3} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & \mu \\ 0 & \mu & 0 \end{pmatrix},$$
$$\widetilde{A_4} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mu \end{pmatrix}, \quad \widetilde{A_5} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Таким образом,

$$\det \widetilde{A_1} = \lambda_1 \lambda_2 \mu = \mu \det A_1, \quad \operatorname{rk} \widetilde{A_1} = 3;$$
$$\det \widetilde{A_2} = 0, \quad \operatorname{rk} \widetilde{A_2} = 2; \qquad \det \widetilde{A_3} = -\lambda_2 \mu^2, \quad \operatorname{rk} \widetilde{A_3} = 3;$$
$$\det \widetilde{A_4} = 0, \quad \operatorname{rk} \widetilde{A_4} = 2; \qquad \det \widetilde{A_5} = 0, \quad \operatorname{rk} \widetilde{A_4} = 1.$$

Предположим теперь, что кривая в двумерном пространстве задана уравнением

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0. \tag{9.4}$$

Матрица квадратичной части A и расширенная матрица $\widetilde{A}$ в этом случае имеют вид

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{12} & a_{22} & b_2 \\ b_1 & b_2 & c \end{pmatrix}.$$

Следовательно, в силу теоремы 9.1 справедлива

Теорема 9.3. *Рассмотрим кривую, заданную уравнением* (9.4). *Справедливы следующие утверждения.*

1) *Кривая является эллипсом, если* $\det A > 0$ *и* $\operatorname{tr} A$, $\det \widetilde{A}$ *отличны от нуля и имеют разные знаки.*

2) *Кривая является мнимым эллипсом, если* $\det A > 0$, *причём* $\operatorname{tr} A$, $\det \widetilde{A}$ *отличны от нуля и имеют одинаковые знаки.*

3) *Кривая является гиперболой, если* $\det A < 0$, $\det \widetilde{A} \neq 0$.

4) *Кривая является парой пересекающихся прямых, если выполнены условия* $\det A < 0$, $\det \widetilde{A} = 0$.

5) *Кривая является парой мнимых пересекающихся прямых, если* $\det A > 0$, $\det \widetilde{A} = 0$.

6) *Кривая является параболой, если* $\det A = 0$, $\det \widetilde{A} \neq 0$.

7) *Кривая является парой параллельных прямых, если* $\operatorname{rk} A=1$, $\operatorname{rk} \widetilde{A} = 2$, *причем на кривой имеется хотя бы одна точка с вещественными координатами.*

8) *Кривая является парой мнимых параллельных прямых, если* $\operatorname{rk} A=1$, $\operatorname{rk} \widetilde{A} = 2$, *причем на кривой нет ни одной вещественной точки.*

9) *Кривая является парой совпадающих прямых, если* $\operatorname{rk} A = \operatorname{rk} \widetilde{A} = 1$.

Например, пусть задана кривая

$$x^2 - 2xy + 3y^2 - 2x + 4y - 5 = 0.$$

Тогда

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}, \quad \det A = 3,\ \operatorname{tr} A = 5,$$

$$\widetilde{A} = A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 4 & 2 \\ -1 & 2 & -5 \end{pmatrix}, \quad \det \widetilde{A} = -19 < 0.$$

Следовательно, эта кривая является эллипсом.

§ 4. Эллипс и гипербола

При классификации кривых второго порядка возникли следующие важные кривые — эллипс и гипербола. В этом разделе мы дадим геометрическую характеризацию этих кривых.

Эллипсом называется множество всех точек плоскости, сумма расстояний до которых от двух заданных точек, называемых *фокусами*, постоянно. *Гиперболой* называется множество точек плоскости, модуль разности расстояний до которых от двух заданных точек, называемых *фокусами*, постоянно.

Пусть фокусы эллипса из рис. 9.1 (гиперболы из рис. 9.2) расположены на оси OX и имеют координаты $(-a, 0)$, $(a, 0)$, где $a > 0$. Точка (x, y) лежит на эллипсе из рис. 9.1 (гиперболе из рис. 9.2) тогда и только тогда, когда

$$\left|\sqrt{(x+a)^2 + y^2} \pm \sqrt{(x-a)^2 + y^2}\right| = r.$$

Рассмотрим треугольник с вершинами в точках

$$(-a, 0), \quad (x, y), \quad (a, 0).$$

Сумма длин любых двух его сторон больше длины третьей. Поэтому в эллипсе $r > 2a$, а в гиперболе $0 < r < 2a$.

Раскрывая модуль, перенося второй радикал в правую часть и возводя в квадрат, получаем

$$\pm 2r\sqrt{(x-a)^2+y^2} = r^2 - 4ax.$$

Снова возводя в квадрат, получаем

$$x^2\frac{4}{r^2} + y^2\frac{4}{r^2-4a^2} = 1.$$

Остается воспользоваться условиями $r > 2a$ или $0 < r < 2a$.

§ 5. Поверхности второго порядка

В этом разделе, опираясь на теорему 9.2, мы приведем классификацию поверхностей второго порядка (квадрик) в трехмерном евклидовом пространстве.

Теорема 9.4. *После применения движения и, возможно, умножения уравнения на ненулевой коэффициент, уравнение квадрики в трехмерном пространстве приводится к одному из следующих видов*:

1) *эллипсоид* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ (рис. 9.6);

2) *мнимый эллипсоид* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = -1$;

3) *однополостный гиперболоид* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$ (рис. 9.6);

4) *двуполостный гиперболоид* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1$ (рис. 9.6);

5) *конус* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0$ (рис. 9.6);

6) *мнимый конус* $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=0;$

7) *эллиптический параболоид* $\dfrac{x^2}{p}+\dfrac{y^2}{q}=2z,\ p,q>0$ (рис. 9.6);

8) *гиперболический параболоид* $\dfrac{x^2}{p}-\dfrac{y^2}{q}=2z, p,q>0$ (рис. 9.6);

9) *эллиптический цилиндр* $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ (рис. 9.7);

10) *мнимый эллиптический цилиндр* $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=-1;$

11) *гиперболический цилиндр* $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ (рис. 9.7);

12) *параболический цилиндр* $y^2=2px$ (рис. 9.7);

13) *пара пересекающихся плоскостей* $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=0;$

14) *пара мнимых пересекающихся плоскостей* $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=0;$

15) *пара параллельных плоскостей* $y^2=b^2;$

16) *пара мнимых параллельных плоскостей* $y^2+b^2=0;$

17) *пара совпадающих плоскостей* $y^2=0.$

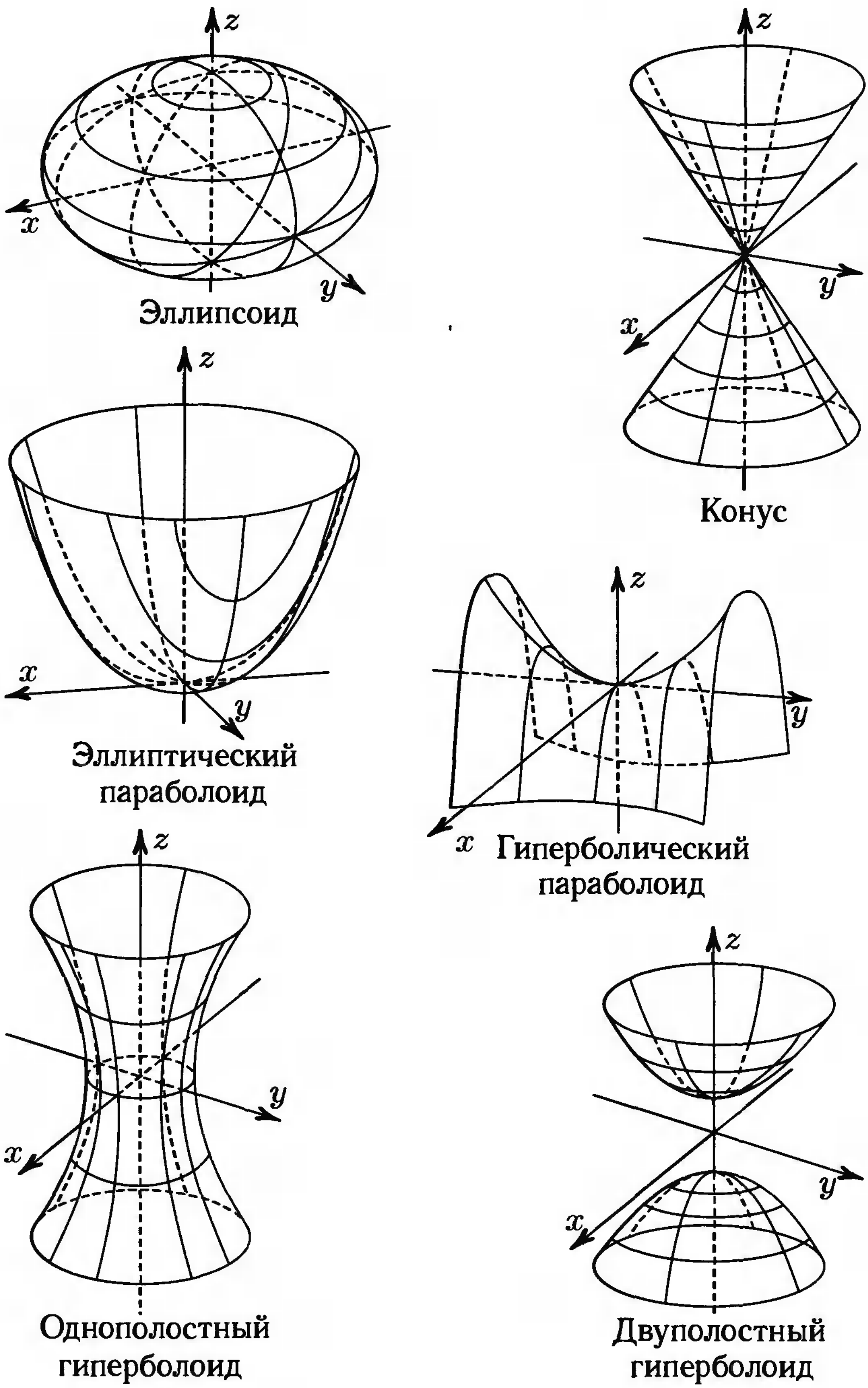


Рис. 9.6. Трехмерные квадрики.

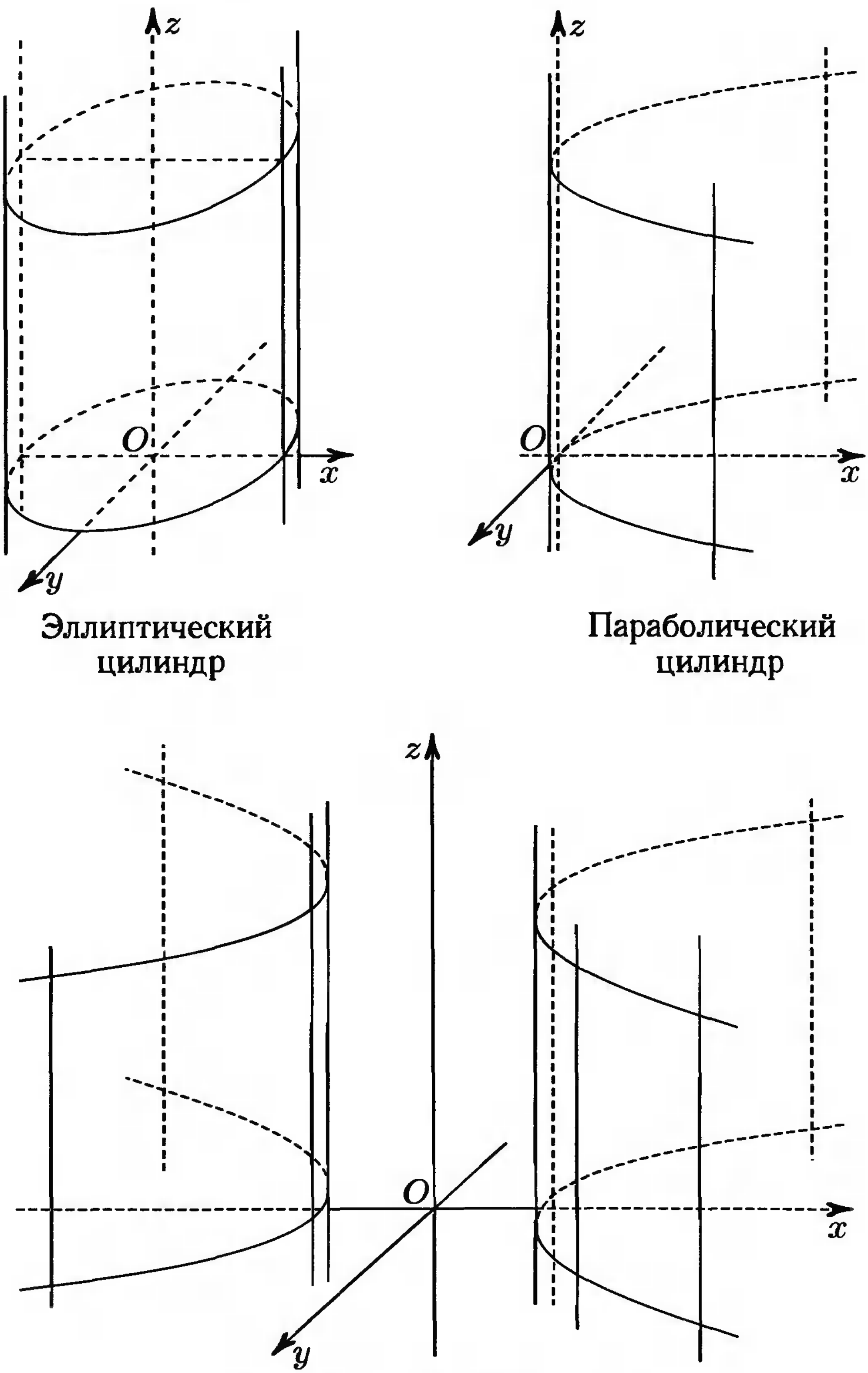


Рис. 9.7. Цилиндры.

СПИСОК ЛИТЕРАТУРЫ

[1] *Артамонов В. А.* Аналитическая геометрия и линейная алгебра. Курс лекций. — Изд. МГУ — 1999.

[2] *Артамонов В. А., Словохотов Ю. Л.* Группы и их приложения в физике, химии и кристаллографии. — М.: Издательский центр «Академия», 2005.

[3] *Винберг Э. Б.* Курс алгебры — М.: Факториал Пресс, 2002.

[4] *Ильин В. А., Ким Г. Д.* Линейная алгебра и аналитическая геометрия. — М.: Издательство МГУ, 1998.

[5] *Кострикин А. И.* Введение в алгебру. Часть I. — М.: МАИК Наука, 2000.

[6] *Кострикин А. И.* Введение в алгебру. Часть II. — М.: МАИК Наука, 2000.

[7] *Кострикин А. И.* Введение в алгебру. Часть III. — М.: МАИК Наука, 2000.

[8] *Курош А. Г.* Курс высшей алгебры. — М.: Наука, 1971.

[9] Сборник задач по алгебре. Под. ред. А. И. Кострикина. — М: МАИК Наука, 2001.

[10] Михалев А. В., Михалев А. А., Начала алгебры. Часть 1. — М.: ИнтУИТ, 2005.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

СПИСОК ОБОЗНАЧЕНИЙ