З.И. БОРЕВИЧ

ОПРЕДЕЛИТЕЛИ И МАТРИЦЫ

$$A = \left\| \begin{matrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \end{matrix} \right.$$

З. И. БОРЕВИЧ

ОПРЕДЕЛИТЕЛИ И МАТРИЦЫ

ИЗДАНИЕ ВТОРОЕ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов высших учебных заведений

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1970

517. 1
Б 82
УДК 512. 83

АННОТАЦИЯ

Предлагаемая вниманию читателей книга является введением в линейную алгебру и составляет часть университетского курса по высшей алгебре. В ней изложены темы: теория определителей, теория систем линейных уравнений, действия над матрицами, алгебраическая теория квадратичных форм. Изложение этих тем, сопровождаемое большим количеством примеров, проводится на конкретной основе без использования понятия векторного пространства и имеет своей целью подготовить читателя к последующему естественному восприятию абстрактных понятий линейной алгебры.

В качестве приложения в книге изложена на матричном языке общая теория приведения уравнений кривых и поверхностей 2-го порядка к каноническому виду.

Книга рассчитана на студентов университетов, педагогических институтов и технических вузов, а также лиц, начинающих самостоятельное изучение высшей алгебры.

$\frac{2\text{-}2\text{-}3}{178\text{-}69}$

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Настоящая книга является введением в линейную алгебру. В ней изложены четыре темы: определители, системы линейных уравнений, матрицы и действия над ними, квадратичные формы.

Рассчитана книга главным образом на лиц, начинающих изучение высшей алгебры и еще не владеющих абстрактными алгебраическими понятиями. В силу этого изложение материала проведено на конкретной основе и имеет целью подготовить читателя к естественному восприятию абстрактных понятий при изучении им линейной алгебры в дальнейшем. В частности, теория систем линейных уравнений изложена без привлечения понятия многомерного векторного пространства.

В книге приведен лишь самый необходимый минимум упражнений, которых начинающему читателю, конечно, недостаточно. Имеется в виду, что читатель использует соответствующие разделы многократно изданных и легко доступных задачников: Д. К. Фаддеев и И. С. Соминский «Сборник задач по высшей алгебре» и И. В. Проскуряков «Сборник задач по линейной алгебре».

Автор выражает глубокую благодарность и признательность члену-корреспонденту Академии наук СССР профессору Дмитрию Константиновичу Фаддееву, в постоянном общении с которым писалась эта книга.

З. Боревич

ГЛАВА I

ТЕОРИЯ ОПРЕДЕЛИТЕЛЕЙ

§ 1. Определители второго и третьего порядков

Понятие определителя возникло в связи с проблемой отыскания формул для значений неизвестных в системе линейных уравнений.

Рассмотрим систему двух линейных уравнений с двумя неизвестными:

$$\left.\begin{aligned} a_1x + b_1y &= c_1, \\ a_2x + b_2y &= c_2. \end{aligned}\right\} \qquad (1)$$

Чтобы найти неизвестное x, умножим первое уравнение на b_2, а второе на $-b_1$. Складывая полученные левые и правые части, получим

$$(a_1b_2 - a_2b_1)\,x = c_1b_2 - c_2b_1.$$

Аналогично, умножая первое уравнение на $-a_2$, второе на a_1, найдем

$$(a_1b_2 - a_2b_1)\,y = a_1c_2 - a_2c_1.$$

Предполагая, что $a_1b_2 - a_2b_1 \neq 0$, получаем

$$x = \frac{c_1b_2 - c_2b_1}{a_1b_2 - a_2b_1}, \qquad y = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}. \qquad (2)$$

Непосредственной проверкой легко убедиться, что значения для x и y, даваемые формулами (2), действительно удовлетворяют системе (1). Нами доказано, таким образом, что если $a_1b_2 - a_2b_1 \neq 0$, то система (1) имеет единственное решение, определяемое формулами (2).

Рассмотрим таблицу чисел

$$\begin{Vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{Vmatrix}, \qquad (3)$$

составленную из коэффициентов при неизвестных x и y системы (1). Такую таблицу чисел называют (*квадратной*) *матрицей* 2-го порядка. Обозначим ее для краткости одной буквой A. Выражение $a_1b_2 - a_2b_1$, стоящее в знаменателе формул (2), определяется числами матрицы (3) по следующему правилу: надо взять произведение чисел, расположенных на главной диагонали (так мы называем диагональ, идущую из левого верхнего угла в правый нижний), и вычесть из него произведение чисел, расположенных на второй диагонали. Так полученное выражение $a_1b_2 - a_2b_1$ называется *определителем 2-го порядка*, составленным из элементов матрицы A (или короче: *определителем матрицы* A), и обозначается через

$$|A| = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1.$$

Например, если

$$A = \begin{Vmatrix} 2 & 3 \\ 5 & 8 \end{Vmatrix},$$

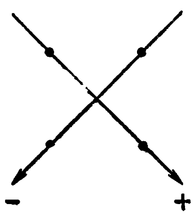
то

$$|A| = \begin{vmatrix} 2 & 3 \\ 5 & 8 \end{vmatrix} = 2\cdot 8 - 5\cdot 3 = 1.$$

Еще пример:

$$\begin{vmatrix} -1 & 2 \\ -3 & 4 \end{vmatrix} = (-1)\cdot 4 - (-3)\cdot 2 = 2.$$

Правило, по которому вычисляется определитель матрицы 2-го порядка, схематически можно изобразить следующим образом:



Приняв введенное определение определителя 2-го порядка, замечаем, что числители в формулах (2) могут быть представлены теперь в виде

$$c_1b_2-c_2b_1=\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}=|A_1|,\quad a_1c_2-a_2c_1=\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}=|A_2|,$$

где матрицы A_1 и A_2 получаются из A заменой первого, соответственно второго, столбца на свободные члены. Формулы (2) принимают теперь следующий вид:

$$x=\frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}=\frac{|A_1|}{|A|},\qquad y=\frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}=\frac{|A_2|}{|A|}. \tag{4}$$

Напомним еще раз, что эти формулы применимы лишь в случае, когда $|A|\neq 0$.

Пример. Пусть требуется решить систему уравнений

$$\left.\begin{aligned} 2x-3y&=-7,\\ 5x+4y&=12. \end{aligned}\right\}$$

Составляем матрицу из коэффициентов при неизвестных

$$A=\begin{Vmatrix} 2 & -3 \\ 5 & 4 \end{Vmatrix}$$

и вычисляем ее определитель

$$|A|=2\cdot 4-(-3)\cdot 5=23.$$

Так как $|A|\neq 0$, то формулы (4) применимы.

Вычисляем числители:

$$|A_1|=\begin{vmatrix} -7 & -3 \\ 12 & 4 \end{vmatrix}=-28+36=8,$$

$$|A_2|=\begin{vmatrix} 2 & -7 \\ 5 & 12 \end{vmatrix}=24+35=59.$$

Следовательно,

$$x=\frac{|A_1|}{|A|}=\frac{8}{23},\qquad y=\frac{|A_2|}{|A|}=\frac{59}{23}.$$

Рассмотрим теперь систему трех линейных уравнений с тремя неизвестными:

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1, \\ a_2x + b_2y + c_2z &= d_2, \\ a_3x + b_3y + c_3z &= d_3. \end{aligned}\right\} \tag{5}$$

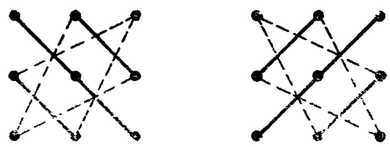
Чтобы найти x, умножим уравнения системы (5) соответственно на $b_2c_3 - b_3c_2$, $b_3c_1 - b_1c_3$, $b_1c_2 - b_2c_1$ и сложим полученные левые и правые части. После приведения подобных членов (относительно x, y и z) окажется, что коэффициенты при y и z равны нулю. Предполагая, что коэффициент при x отличен от нуля, получим

$$x = \frac{d_1b_2c_3 + d_2b_3c_1 + d_3b_1c_2 - d_1b_3c_2 - d_2b_1c_3 - d_3b_2c_1}{a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1}. \tag{6}$$

Составим, как и выше, таблицу чисел из коэффициентов при неизвестных:

$$A = \begin{Vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{Vmatrix}.$$

Эта таблица чисел называется *квадратной матрицей 3-го порядка* (у нее три строчки и три столбца). Выражение, стоящее в знаменателе формулы (6), составлено из чисел матрицы A по следующему правилу. Произведение чисел, расположенных на главной диагонали, и два произведения чисел, расположенных в вершинах двух равнобедренных треугольников с основанием, параллельным главной диагонали, и с вершиной в противоположном углу, берутся со знаком плюс. Три произведения, которые строятся по такому же правилу, но относительно второй диагонали, берутся со знаком минус. Схематически это правило может быть изображено следующим образом:



+ −

Так составленная сумма шести слагаемых (из которых три взяты со знаком плюс, а три — со знаком минус) называется *определителем* (*3-го порядка*) *матрицы* A и обозначается через

$$|A| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Например, если

$$A = \begin{Vmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ -1 & 2 & 5 \end{Vmatrix},$$

то

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ -1 & 2 & 5 \end{vmatrix} =$$

$$= 1\cdot 1\cdot 5 + 2\cdot 2\cdot 3 + (-1)\cdot 2\cdot 4 - (-1)\cdot 1\cdot 3 - \\ - 2\cdot 2\cdot 5 - 1\cdot 2\cdot 4 = -16.$$

Итак, знаменатель в формуле (6) представляется в виде определителя $|A|$. Что касается числителя, то, поскольку он получается из знаменателя заменой a_1, a_2, a_3 соответственно на d_1, d_2, d_3, его можно представить в виде определителя

$$\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix} = |A_1|.$$

Аналогичным образом, если уравнения системы (5) умножим последовательно на $a_3c_2 - a_2c_3$, $a_1c_3 - a_3c_1$, $a_2c_1 - a_1c_2$ и результаты сложим, найдем формулу для y. Наконец, умножая уравнения (5) последовательно на $a_2b_3 - a_3b_2$, $a_3b_1 - a_1b_3$, $a_1b_2 - a_2b_1$, найдем формулу для z. Окончательно будем иметь

$$x = \frac{|A_1|}{|A|}, \qquad y = \frac{|A_2|}{|A|}, \qquad z = \frac{|A_3|}{|A|}, \tag{7}$$

где матрицы A_1, A_2, A_3 получаются из A заменой соответствующего столбца на свободные члены.

Заметим, что формулы (7) нами выведены в предположении, что $|A| \neq 0$.

В качестве примера решим систему

$$\left.\begin{aligned} 2x - 3y + 5z &= 1, \\ x + 2y - 3z &= 7, \\ 3x - \ \ y - 2z &= 5. \end{aligned}\right\}$$

Так как определитель

$$|A| = \begin{vmatrix} 2 & -3 & 5 \\ 1 & 2 & -3 \\ 3 & -1 & -2 \end{vmatrix} = -8 - 5 + 27 - 30 - 6 - 6 = -28$$

отличен от нуля, то формулы (7) применимы. Имеем

$$|A_1| = \begin{vmatrix} 1 & -3 & 5 \\ 7 & 2 & -3 \\ 5 & -1 & -2 \end{vmatrix} = -89,$$

$$|A_2| = \begin{vmatrix} 2 & 1 & 5 \\ 1 & 7 & -3 \\ 3 & 5 & -2 \end{vmatrix} = -85,$$

$$|A_3| = \begin{vmatrix} 2 & -3 & 1 \\ 1 & 2 & 7 \\ 3 & -1 & 5 \end{vmatrix} = -21,$$

поэтому

$$x = \frac{89}{28}, \qquad y = \frac{85}{28}, \qquad z = \frac{3}{4}.$$

Чтобы получить формулы, аналогичные формулам (4) и (7), для систем n линейных уравнений с n неизвестными при любом n, надо дать определение определителя квадратной матрицы n-го порядка. Сделать это таким же способом, каким мы ввели определители 2-го и 3-го порядков, затруднительно, так как число слагаемых, из которых составляется определитель, очень быстро растет с увеличением порядка (определитель n-го порядка содержит $n!$ слагаемых, что уже при $n = 4$ равно 24). Мы вынуждены поэтому избрать иной путь, основанный на использовании некоторых сведений из теории перестановок. Обобщение формул (4) и (7) на общий случай будет дано нами в § 8.

Упражнения

1. Вычислить определители:

a) $\begin{vmatrix} 7 & -2 \\ 3 & -5 \end{vmatrix}$; b) $\begin{vmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{vmatrix}$;

c) $\begin{vmatrix} 1 & 2 & 3 \\ -2 & 3 & 5 \\ 5 & 3 & 4 \end{vmatrix}$; d) $\begin{vmatrix} 0 & a & a \\ a & 0 & a \\ a & a & 0 \end{vmatrix}$.

Отв. a) -29; b) 1; c) 0; d) $2a^3$.

2. Решить систему

$$\left.\begin{aligned} 3x+7y&=-11, \\ 5x-2y&=9. \end{aligned}\right\}$$

Отв. $x=1$, $y=-2$.

3. Решить систему

$$\left.\begin{aligned} x+2y+3z&=-7, \\ 2x+y+2z&=-2, \\ 3x+2y+z&=3. \end{aligned}\right\}$$

Отв. $x=2$, $y=0$, $z=-3$.

4. Проверить, что

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2-x_1)(x_3-x_1)(x_3-x_2).$$

§ 2. Определение определителя *n*-го порядка

Квадратная таблица n^2 чисел, расположенных в n строчках и n столбцах, будет называться нами *матрицей n-го порядка*. Чтобы подметить общее правило составления определителей матриц n-го порядка, присмотримся внимательнее к определителям 2-го и 3-го порядков. Отвлекаясь пока от знака, поставим вопрос: чем характеризуются произведения, входящие в состав определителя? Прежде всего мы замечаем, что число сомножителей в каждом произведении равно порядку матрицы. Далее, если мы обратимся к какому-нибудь конкретному произведению, то увидим, что в нем сомножители взяты по одному из каждой строчки и по одному из каждого столбца. Так обстоит дело для определителей 2-го и 3-го порядков, и в этом мы убеждаемся, анализируя определения первого параграфа. Естественно предположить, что мы получим разумное определение определителя матриц n-го порядка, если в его основу положим эту подмеченную закономерность.

Условимся сначала об одном весьма удобном способе обозначения элементов матрицы. Место каждого элемента в матрице вполне определяется указанием номера строчки и номера столбца, в которых находится наш элемент. Чтобы в обозначении элемента отразить его местонахождение в матрице, уславливаются все элементы матрицы обозначать одной буквой, но снабжать ее двумя индексами, из которых один обозначает номер строчки, а второй — номер столбца. Обычно эти индексы пишут справа внизу, причем сначала ставят номер строчки, а рядом — номер столбца. Например, $a_{2,5}$ (или a_{25}, читается: a — два-пять) есть элемент матрицы, расположенный во второй строчке и в пятом столбце; элемент a_{ik} расположен в i-й строчке и k-м столбце. (Конечно, вместо a можно взять любую другую букву.) Воспользовавшись введенной системой обозначений, матрицу n-го порядка запишем теперь в виде

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

Выясним, сколько можно составить различных произведений из элементов матрицы A, взятых по одному из каждой строчки и по одному из каждого столбца. Заметим, что под различными мы понимаем произведения, составленные разными способами, так что разные в нашем смысле произведения могут случайно оказаться равными по своим значениям. Пусть P — некоторое такое произведение. Множитель из первой строчки, входящий в P, имеет вид $a_{1\alpha}$, где α — номер столбца, т. е. $1 \leqslant \alpha \leqslant n$. Множитель из второй строчки имеет вид $a_{2\beta}$, $1 \leqslant \beta \leqslant n$, и т. д. Наконец, множитель из n-й строчки имеет вид $a_{n\omega}$, $1 \leqslant \omega \leqslant n$. Таким образом,

$$P = a_{1\alpha} a_{2\beta} \dots a_{n\omega}. \tag{2}$$

С другой стороны, все сомножители в этом произведении взяты по одному из каждого столбца, а значит, все вторые индексы $\alpha, \beta, \ldots, \omega$ различны (если бы, например, $\alpha=\beta$, то это означало бы, что $a_{1\alpha}$ и $a_{2\beta}$ принадлежат одному столбцу). Но в таком случае они образуют перестановку $(\alpha\beta \ldots \omega)$ из чисел $1, 2, \ldots, n$. С произведением P связывается, таким образом, некоторая вполне определенная перестановка: $(\alpha\beta \ldots \omega)$. Наоборот, если взять какую-нибудь произвольную перестановку $(\alpha\beta \ldots \omega)$ из чисел $1, 2, \ldots, n$ и по ней составить произведение (2), то это произведение, очевидно, будет иметь множители по одному из каждой строчки и по одному из каждого столбца. Из этих рассуждений следует, что число интересующих нас произведений, из которых будет строиться определитель матрицы (1), равно числу всех перестановок из чисел $1, 2, \ldots, n$, т. е. равно $n!$.

Дадим теперь следующее предварительное определение.

Определение. *Определителем матрицы A n-го порядка называется сумма всех $n!$ произведений элементов этой матрицы, взятых по одному из каждой строчки и по одному из каждого столбца; при этом каждое произведение снабжено знаком плюс или минус по некоторому правилу.*

Предполагая, что A есть матрица 4-го порядка, рассмотрим несколько примеров.

1. Произведения $a_{11}a_{23}a_{34}$ и $a_{12}a_{23}a_{34}a_{41}a_{32}$ не входят в определитель матрицы A, так как число сомножителей в них не равно порядку матрицы.

2. Произведение $a_{12}a_{31}a_{22}a_{44}$ также не входит в определитель, так как два сомножителя, a_{12} и a_{22}, принадлежат одному и тому же (второму) столбцу.

3. Произведения $a_{11}a_{22}a_{33}a_{44}$ и $a_{21}a_{43}a_{14}a_{32}$ содержат множители по одному из каждой строчки и по одному из каждого столбца — следовательно, они входят в состав определителя.

Чтобы приведенное выше определение сделать полным, надо еще сформулировать правило для знака, с которым берется то или иное произведение.

Рассмотрим опять определители 2-го и 3-го порядков, применив новую систему обозначений элементов с помощью двойных индексов:

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

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

$$- a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Во всех произведениях мы выписали сомножители в порядке следования строчек (первые индексы расположены в натуральном порядке). Присмотримся к перестановкам, которые образуют номера столбцов в этих произведениях. В определителе 2-го порядка в первом произведении (которое берется со знаком плюс) номера столбцов образуют натуральное расположение, во втором же произведении (которое берется со знаком минус) они образуют перестановку (21), в которой расположение номеров 1 и 2 противоположно их натуральному расположению.

В определителе 3-го порядка номера столбцов в произведениях, которые берутся со знаком плюс, образуют перестановки

$$\left.\begin{array}{l} (1\ \ 2\ \ 3), \\ (2\ \ 3\ \ 1), \\ (3\ \ 1\ \ 2), \end{array}\right\} \qquad (3)$$

а в трех остальных произведениях — перестановки

$$\left.\begin{array}{l} (1\ \ 3\ \ 2), \\ (2\ \ 1\ \ 3), \\ (3\ \ 2\ \ 1). \end{array}\right\} \qquad (4)$$

В пяти перестановках (3) и (4) (кроме первой) взаимное расположение некоторых пар номеров противоположно их расположению в перестановке (1 2 3). Такое явление называют *инверсией* (или беспорядком).

Определение. *Пусть имеется перестановка* ($\alpha\ \ \beta \ldots \omega$) *чисел* 1, 2, ..., *n. Мы говорим, что два числа, входящих в эту перестановку, образуют инверсию, если большее число из нашей пары предшествует* *меньшему. Число пар, образующих инверсию, называется числом инверсий перестановки.*

В натуральном расположении ($1\ \ 2\ \dots\ n$) нет пар, образующих инверсию. С другой стороны, наибольшее возможное число инверсий содержится в перестановке ($n\ \ n-1 \dots 2\ \ 1$); здесь каждая из $C_n^2 = \frac{n(n-1)}{2}$ пар образует инверсию. В перестановке ($3\ \ 1\ \ 5\ \ 2\ \ 4$) имеются четыре пары, образующие инверсию: 3 и 1, 3 и 2, 5 и 2, 5 и 4.

Легко видеть, что для определения числа инверсий в некоторой перестановке ($\alpha\ \beta \dots \omega$) следует подсчитать, сколько для каждого числа имеется следующих за ним меньших чисел, а затем все найденные значения сложить. Например, число инверсий в перестановке ($4\ \ 7\ \ 1\ \ 5\ \ 3\ \ 6\ \ 2$) равно

$$3+5+0+2+1+1+0=12.$$

Определение. *Перестановка называется четной, если она содержит четное число инверсий, и называется нечетной в противном случае.*

Натуральное расположение ($1\ \ 2\ \dots\ n$) содержит 0 инверсий — следовательно, это четная перестановка. Перестановка ($2\ \ 6\ \ 1\ \ 5\ \ 3\ \ 4$) нечетная, так как у нее 7 инверсий.

Возвращаясь к перестановкам (3) и (4), легко замечаем, что три первые перестановки четные, а три остальные — нечетные. Это наблюдение мы и кладем в основу нашего следующего определения.

Правило для знака. *Пусть P — фиксированное произведение, входящее в состав определителя матрицы A n-го порядка. Выпишем сомножители в порядке следования строчек:*

$$P = a_{1\alpha} a_{2\beta} \dots a_{n\omega}.$$

Тогда номера столбцов дадут перестановку ($\alpha\ \beta \dots \omega$). Произведение P берется со знаком плюс, если эта перестановка четная, и со знаком минус, если она нечетная.

Обозначим через s число инверсий в перестановке ($\alpha\ \beta \dots \omega$). Тогда, объединяя вместе случаи четной и нечетной перестановок, можно сказать, что произведение P входит в определитель со знаком $(-1)^s$ (последнее означает, что знак определяется множителем $(-1)^s$).

Пример 1. Произведение $a_{11}a_{22}\ldots a_{nn}$ элементов, стоящих на главной диагонали (из левого верхнего угла в правый нижний), всегда берется со знаком плюс, так как натуральное расположение $(1\ 2\ \ldots n)$ является четной перестановкой.

Пример 2. Произведение $a_{1n}a_{2n-1}\ldots a_{n-12}a_{n1}$ элементов второй диагонали входит в определитель со знаком $(-1)^{\frac{n(n-1)}{2}}$, так как перестановка $(n\ n-1\ldots 2\ 1)$ имеет $\frac{n(n-1)}{2}$ инверсий.

Пример 3. Рассмотрим произведения $a_{34}a_{13}a_{42}a_{21}$ и $a_{24}a_{33}a_{12}a_{41}$, входящие в состав определителя матрицы 4-го порядка. Чтобы определить знаки, с которыми эти произведения входят в определитель, надо их переписать в виде $a_{13}a_{21}a_{34}a_{42}$ и $a_{12}a_{24}a_{33}a_{41}$. Выписываем теперь перестановки из номеров столбцов: $(3\ 1\ 4\ 2)$ и $(2\ 4\ 3\ 1)$. В первой перестановке мы имеем три инверсии, во второй — четыре. Следовательно, первое произведение входит в определитель со знаком минус, второе — со знаком плюс.

Данное нами определение определителя n-го порядка основывается на аналогии с определителями 2-го и 3-го порядков. Последние же были введены с целью получить компактные формулы для решений систем линейных уравнений (с двумя и тремя неизвестными). Разумность определения определителя в общем случае будет подтверждена в § 8, где мы получим формулы для значений неизвестных системы n линейных уравнений с n неизвестными, аналогичные формулам (4) и (7) из § 1 (опять-таки при условии, что определитель, составленный из коэффициентов при неизвестных, отличен от нуля).

Для определителя матрицы (1) принято обозначение

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

Если для обозначения суммы мы воспользуемся знаком $\sum$, то определитель матрицы (1) можно будет записать в виде формулы

$$|A|=\sum_{P}\pm P=\sum_{(\alpha\,\beta\,\ldots\,\omega)}\pm a_{1\alpha}a_{2\beta}\ldots a_{n\omega}.$$

В первой сумме P пробегает все $n!$ произведений элементов матрицы A, взятых по одному из каждой строчки и по одному из каждого столбца. Во второй сумме суммирование ведется по всем $n!$ перестановкам $(\alpha\ \beta\ \ldots\omega)$ из чисел $1, 2, \ldots, n$. В обоих случаях знак перед каждым произведением определяется согласно сформулированному выше правилу.

Вычисление определителя матрицы n-го порядка, основанное на одном лишь определении, представляет собой весьма трудоемкий процесс, так как число слагаемых, из которых составляется определитель, очень быстро растет с увеличением n (для определителя 4-го порядка имеем 24 слагаемых, для 5-го порядка — 120 слагаемых и т. д.). В дальнейшем нами будут указаны быстро приводящие к цели методы для вычисления определителей, основанные на свойствах последних. Однако в некоторых простейших случаях можно все же найти значение определителя, пользуясь только определением. Пусть, например, какая-нибудь строчка матрицы (или столбец) состоит сплошь из нулей. Так как в любое произведение, входящее в состав определителя, входит сомножитель из каждой строчки, в том числе и из строчки, состоящей из нулей, то все эти произведения будут равны нулю, а значит, их сумма также равна нулю. Таким образом, если все элементы какой-нибудь строчки (или столбца) матрицы n-го порядка равны нулю, то ее определитель равен нулю.

Отметим еще один важный частный случай. Пусть в матрице A (см. (1)) все элементы, расположенные выше главной диагонали, равны нулю, т. е. пусть $a_{i\alpha}=0$, если только $\alpha>i$. В этом случае те произведения (входящие в состав определителя $|A|$), которые содержат в качестве сомножителей хоть один элемент, расположенный выше главной диагонали, будут равны нулю. Выкинем все эти заведомо равные нулю произведения, содержащие множители $a_{i\alpha}$ с $\alpha>i$. Оставшиеся произведения будут иметь вид

$$a_{1\alpha_1}a_{2\alpha_2}\ldots a_{n\alpha_n},$$

где

$$\alpha_1 \leqslant 1 \quad \alpha_2 \leqslant 2, \ \ldots, \ \alpha_n \leqslant n.$$

Но из этих неравенств последовательно получаем

$$\alpha_1 = 1, \quad \alpha_2 = 2, \ \ldots, \ \alpha_n = n.$$

(Напомним, что натуральные числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ попарно различны.) Таким образом, только одно произведение, а именно $a_{11}a_{22}\ldots a_{nn}$, не содержит элементов, лежащих выше главной диагонали. Так как это произведение входит в состав определителя со знаком плюс, то окончательно получаем

$$\begin{vmatrix} a_{11} & 0 & 0 & \ldots 0 \\ a_{21} & a_{22} & 0 & \ldots 0 \\ a_{31} & a_{32} & a_{33} & \ldots 0 \\ \cdot & \cdot & \cdot & \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} & a_{n3} & \ldots a_{nn} \end{vmatrix} = a_{11}a_{22} \ldots a_{nn}.$$

Аналогичная формула может быть получена и для случая, когда равны нулю все элементы, расположенные ниже главной диагонали. Итак, *если все элементы матрицы n-го порядка, расположенные выше (или ниже) главной диагонали, равны нулю, то определитель такой матрицы равен произведению элементов, лежащих на главной диагонали.*

Квадратные матрицы, у которых все элементы выше или ниже главной диагонали равны нулю, иногда называются *треугольными матрицами.*

Замечание. Говоря об определителе матрицы, мы до сих пор предполагали, что порядок матрицы $\geqslant 2$. Ясно, что вполне допустимо рассматривать таблицу, состоящую лишь из одного числа, т. е. квадратную матрицу 1-го порядка. Применив определение определителя к этому случаю, мы легко увидим, что если $A = \|a\|$, то $|A| = a$, т. е. определитель матрицы 1-го порядка равен (единственному) элементу этой матрицы.

Упражнения

1. Какие из произведений

1) $a_{31}a_{43}a_{14}a_{23}$,
2) $a_{14}a_{21}a_{33}a_{42}$,
3) $a_{23}a_{42}a_{31}a_{14}$

входят в состав определителя матрицы 4-го порядка и с каким знаком?

Отв. 1) не входит; 2) входит со знаком плюс; 3) входит со знаком минус.

2. Пользуясь определением, вычислить определители:

$$1)\ \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{vmatrix};\quad 2)\ \begin{vmatrix} 2 & 1 & 0 & 3 \\ -1 & 2 & 0 & 4 \\ -2 & 3 & 0 & 5 \\ -3 & 4 & 0 & 6 \end{vmatrix};\quad 3)\ \begin{vmatrix} 1 & 0 & 0 & 2 \\ 3 & 0 & 0 & 4 \\ 0 & 5 & 6 & 0 \\ 0 & 7 & 8 & 0 \end{vmatrix}.$$

Отв. 1) 24; 2) 0; 3) 4.

3. Показать, что произведение элементов второй диагонали матрицы n-го порядка входит в ее определитель со знаком плюс при $n=4k$ и $n=4k+1$ и со знаком минус при $n=4k+2$ и $n=4k+3$.

§ 3. Транспозиции

В формулировке правила для знака, с которым то или иное произведение входит в состав определителя n-го порядка, существенную роль играет понятие перестановки. Естественно поэтому, что при доказательстве свойств определителей нам придется использовать некоторые сведения из теории перестановок. Изложению этих сведений мы и посвятим настоящий параграф.

Напомним, что *перестановкой* из n предметов (не обязательно чисел) называется расположение этих предметов в некотором определенном порядке. Если мы переставляемые предметы занумеруем натуральными числами $1, 2, \ldots, n$ (т. е. к каждому предмету привяжем номер), то произвольная перестановка предметов будет однозначно определяться перестановкой их номеров. (Например, для перестановки $(X_3X_1X_4X_2)$ предметов X_1, X_2, X_3, X_4 соответствующей перестановкой номеров будет $(3\ 1\ 4\ 2)$.) В силу этого мы можем ограничиться рассмотрением перестановок чисел $1, 2, \ldots, n$.

Пусть дана произвольная перестановка $(\alpha\ \beta \ldots \omega)$ чисел $1, 2, \ldots, n$. Если мы в этой перестановке поменяем местами два числа, то в результате получим новую перестановку тех же чисел. Например, после перемены местами чисел 4 и 5 в перестановке $(2\ 5\ 3\ 4\ 1)$ мы получим перестановку $(2\ 4\ 3\ 5\ 1)$.

Определение. *Операция перехода от одной перестановки к другой, при которой два числа меняются* *местами, а остальные остаются на своих местах, называется транспозицией.*

Записывать транспозиции мы будем так:

$$(2\ 6\ 1\ 5\ 4\ 3) \to (5\ 6\ 1\ 2\ 4\ 3).$$

(Здесь 2 и 5 поменялись местами.)

Теорема 1. *От произвольной перестановки можно перейти к любой другой при помощи нескольких последовательно выполненных транспозиций.*

Утверждение этой теоремы довольно очевидно. Например, переход от (5 1 7 4 2 3 6) к перестановке (3 4 1 5 7 6 2) достигается следующими транспозициями:

$$(5\ 1\ 7\ 4\ 2\ 3\ 6) \to (3\ 1\ 7\ 4\ 2\ 5\ 6) \to (3\ 4\ 7\ 1\ 2\ 5\ 6) \to$$
$$\to (3\ 4\ 1\ 7\ 2\ 5\ 6) \to (3\ 4\ 1\ 5\ 2\ 7\ 6) \to (3\ 4\ 1\ 5\ 7\ 2\ 6) \to$$
$$\to (3\ 4\ 1\ 5\ 7\ 6\ 2).$$

Доказательство. Доказательство теоремы проведем индукцией по числу n чисел, входящих в рассматриваемые перестановки. При $n = 2$ имеется только две перестановки (1 2) и (2 1), и от любой из них мы переходим ко второй при одной транспозиции. Таким образом, при $n = 2$ теорема справедлива.

Пусть теперь $n \geqslant 3$, и пусть теорема уже доказана для перестановок $n - 1$ чисел. Рассмотрим две произвольные перестановки n чисел:

$$(\alpha\ \beta\ \gamma \ldots \omega), \qquad (1)$$
$$(\alpha'\ \beta'\ \gamma' \ldots \omega'). \qquad (2)$$

Предположим сначала, что $\alpha = \alpha'$. В этом случае $(\beta\ \gamma \ldots \omega)$ и $(\beta'\ \gamma' \ldots \omega')$ будут являться перестановками одних и тех же $n - 1$ чисел. Следовательно, по индукционному предположению от первой из них можно перейти ко второй при помощи нескольких последовательно выполненных транспозиций. Если эти же транспозиции (и в том же порядке) выполнить в перестановке (1), то получим, очевидно, перестановку $(\alpha\ \beta'\ \gamma' \ldots \omega')$, т. е. перестановку (2) (поскольку $\alpha = \alpha'$). Таким образом, в случае $\alpha = \alpha'$ теорема доказана.

Пусть теперь $\alpha \neq \alpha'$. Исходная перестановка (1) имеет в этом случае вид

$$(\alpha\ \beta\ \gamma \ldots \alpha' \ldots \omega) \qquad (1a)$$

(α' совпадает с одним из чисел $\beta, \gamma, \ldots, \omega$). Поменяем в (1а) числа α и α' местами (т. е. сделаем одну транспозицию). Получим перестановку

$$(\alpha' \beta \gamma \ldots \alpha \ldots \omega). \tag{3}$$

Сравнивая (3) и (2), видим, что у этих перестановок числа, стоящие на первом месте, совпадают. Следовательно, по только что доказанному при помощи нескольких транспозиций мы можем перейти от перестановки (3) к перестановке (2). Таким образом, и в этом случае перестановка (1) может быть переведена при помощи транспозиций в перестановку (2). Теорема 1 доказана.

Теорема 2. *При выполнении одной транспозиции перестановка переходит в перестановку противоположного наименования, т. е. четная перестановка переходит в нечетную, а нечетная — в четную.*

Доказательство. Пусть мы имеем перестановку

$$(\alpha\ \beta \ldots \lambda\ \gamma_1 \gamma_2 \ldots \gamma_m\ \mu \ldots \omega) \tag{I}$$

из n чисел $1, 2, \ldots, n$. Поменяв здесь местами λ и μ, мы получим перестановку

$$(\alpha\ \beta \ldots \mu\ \gamma_1\ \gamma_2 \ldots \gamma_m\ \lambda \ldots \omega). \tag{II}$$

Нам надо показать, что перестановки (I) и (II) имеют противоположные наименования, т. е. если (I) — четная перестановка, то (II) — нечетная, и наоборот.

Можно, очевидно, предполагать, что $\lambda < \mu$, так как если это не так, то мы вправе перестановки (I) и (II) поменять местами.

Всего мы имеем $C_n^2 = \frac{n(n-1)}{2}$ пар из n чисел 1, 2, ..., n. Разобьем все эти пары на четыре группы.

1. Группа, состоящая только из одной пары: λ и μ.

2. Пары, образованные числом λ со всеми промежуточными числами $\gamma_1, \gamma_2, \ldots, \gamma_m$. Число пар в этой группе равно m.

3. Пары, образованные числом μ со всеми промежуточными числами $\gamma_1, \gamma_2, \ldots, \gamma_m$.

4. Все прочие пары.

В каждой из перечисленных групп обозначим соответственно через s_1, s_2, s_3, s_4 число тех пар, которые образуют инверсию в перестановке (I). Тогда общее число инверсий в перестановке (I) будет

$$s = s_1 + s_2 + s_3 + s_4.$$

Аналогично, если через t_1, t_2, t_3, t_4 обозначить число пар в каждой из групп, образующих инверсию в перестановке (II), то число всех инверсий в перестановке (II) будет

$$t = t_1 + t_2 + t_3 + t_4.$$

Для доказательства теоремы 2 мы должны установить, что числа s и t имеют противоположную друг другу четность, т. е. что разность $s - t$ есть число нечетное.

Так как по нашему предположению $\lambda < \mu$, то единственная пара из первой группы будет образовывать инверсию в перестановке (II) и не будет ее образовывать в перестановке (I). Другими словами,

$$s_1 = 0, \qquad t_1 = 1. \tag{4}$$

Обозначим через p число тех из промежуточных чисел $\gamma_1, \gamma_2, \ldots, \gamma_m$, которые меньше λ. Ясно, что эти p чисел вместе с λ дадут нам все те пары из второй группы, которые образуют инверсию в перестановке (I). Все прочие промежуточные числа (их число равно $m - p$) будут больше λ, поэтому во второй группе мы будем иметь ровно $m - p$ пар, образующих инверсию в перестановке (II). Этим доказано, что

$$s_2 = p, \quad t_2 = m - p. \tag{5}$$

Аналогично, обозначив через q число промежуточных чисел $\gamma_1, \gamma_2, \ldots, \gamma_m$, меньших μ, мы установим, что

$$s_3 = m - q, \quad t_3 = q. \tag{6}$$

Обратимся теперь к парам последней группы. Пусть ρ и σ — произвольная пара из четвертой группы. Если ρ и σ отличны от λ и μ, то числа ρ и σ в перестановках (I) и (II) стоят на одинаковых местах, и поэтому они одновременно либо образуют инверсию в обеих перестановках, либо нет. Рассмотрим далее случай, когда одно, и только одно, из чисел нашей пары совпадает с λ или μ. Пусть, например, $\rho = \lambda$. Тогда σ отлично от μ, γ_1, γ_2, ..., γ_m (в противном случае мы имели бы пару из первой или второй группы). Если σ в перестановке (I) расположено слева от λ, то в перестановке (II) оно также будет слева от λ. Если же в (I) число σ находится справа от λ, то (поскольку σ отлично от промежуточных чисел γ_1, ..., γ_m и от μ) оно будет находиться справа от λ и в (II). Рассмотрев аналогичным образом случай $\rho = \mu$, мы убедимся в том, что взаимное расположение всех пар из четвертой группы в перестановках (I) и (II) одинаково. Следовательно, каждая пара из этой группы, образующая инверсию в (I), будет образовывать ее и в (II), и наоборот, а значит,

$$s_4 = t_4. \tag{7}$$

В силу (4) — (7) мы имеем теперь

$$s - t = s_1 + s_2 + s_3 + s_4 - t_1 - t_2 - t_3 - t_4 =$$
$$= p + m - q - 1 - m + p - q = 2(p - q) - 1.$$

Таким образом, разность $s - t$ есть число нечетное. Но в таком случае s и t — числа разной четности, а значит, перестановки (I) и (II) имеют противоположные наименования. Теорема 2 доказана.

Замечание. Может случиться, что переставляемые числа λ и μ стоят рядом (т. е. $m = 0$). В этом случае вторая и третья группы пар отсутствуют (в частности, $p = 0$, $q = 0$) и доказательство несколько упрощается.

Следствие. *При четном числе транспозиций перестановка переходит в перестановку того же наименования, а при нечетном — в перестановку противоположного наименования.*

Действительно, по теореме 2 при одной транспозиции наименование перестановки меняется. После выполнения второй транспозиции в полученной перестановке наименование опять изменится, — значит, после двух транспозиций мы получим перестановку с исходным наименованием. После трех транспозиций мы получим перестановку противоположного наименования по сравнению с исходной и т. д.

Теорема 3. *Число всех четных перестановок из n чисел равно числу всех нечетных перестановок и равно, следовательно,* $\frac{n!}{2}$ $(n \geqslant 2)$.

Доказательство. Обозначим через a число всех четных и через b число всех нечетных перестановок из n чисел. Так как всего мы имеем $n!$ перестановок из n чисел, то $a+b=n!$. Нам надо доказать, что $a=b$.

Выпишем отдельно в виде таблицы все четные перестановки (таблица содержит, следовательно, a перестановок). Поменяем затем в каждой из выписанных четных перестановок первые два числа местами (перестановка (4 1 5 3 2) перейдет, например, в перестановку (1 4 5 3 2)). Мы получим в результате вторую таблицу, состоящую опять-таки из a перестановок. Все перестановки новой таблицы, очевидно, попарно различны, так как если бы в ней нашлись две одинаковые перестановки, то после вторичной перемены местами первых двух чисел в этих перестановках мы пришли бы к двум совпадающим перестановкам в первой таблице, чего не может быть (ведь все a перестановок первой таблицы различны). Далее, по теореме 2 все a перестановок второй таблицы будут нечетными. Число всех нечетных перестановок из n чисел мы обозначили через b. Но поскольку у нас нет уверенности, что во второй таблице (содержащей a различных нечетных перестановок) будут присутствовать все нечетные перестановки, то мы можем написать пока только неравенство

$$a \leqslant b.$$

Аналогичным образом, поменяв местами роли четных и нечетных перестановок, мы докажем, что

$$b \leqslant a.$$

Из полученных неравенств следует теперь, что

$$a=b,$$

и теорема 3 доказана.

Замечание. Теорема 3 показывает, что в определителе произвольного порядка число слагаемых, которые берутся со знаком плюс, равно числу слагаемых, которые берутся со знаком минус.

Упражнения

1. Перейти от перестановки (7 5 3 9 1 6 2 8 4) к натуральному расположению при помощи транспозиций.

Отв. Минимальное число требуемых транспозиций равно 4.

2. Чему равны числа s_1, s_2, s_3, s_4, t_1, t_2, t_3, t_4 (см. доказательство теоремы 2) в случае транспозиции

$$(7\ 5\ 3\ 9\ 1\ 6\ 2\ 8\ 4) \to (7\ 5\ 8\ 9\ 1\ 6\ 2\ 3\ 4)?$$

3. Пусть перестановка $(\alpha_1\ \alpha_2\ \dots\ \alpha_n)$ из чисел $1, 2, \dots, n$ имеет s инверсий. Сколько имеется инверсий в перестановке $(\alpha_n\ \dots\ \alpha_2\ \alpha_1)$?

Отв. $\frac{n(n-1)}{2} - s$.

4. Пусть P — произвольное произведение, входящее в состав определителя матрицы A n-го порядка (см. (1) из § 2). Запишем это произведение в порядке следования номеров строчек:

$$P = a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n},$$

и в порядке следования номеров столбцов:

$$P = a_{\beta_1 1} a_{\beta_2 2} \dots a_{\beta_n n}.$$

Доказать, что перестановки $(\alpha_1\ \alpha_2\ \dots\ \alpha_n)$ и $(\beta_1\ \beta_2\ \dots\ \beta_n)$ имеют одно и то же число инверсий.

5. Пусть

$$P = a_{\alpha_1\beta_1} a_{\alpha_2\beta_2} \dots a_{\alpha_n\beta_n}$$

есть произведение, входящее в определитель матрицы A n-го порядка, в котором сомножители записаны в произвольном порядке. Обозначим через s и t число инверсий в перестановках $(\alpha_1\ \alpha_2\ \dots\ \alpha_n)$ и $(\beta_1\ \beta_2\ \dots\ \beta_n)$ соответственно. Доказать, что произведение входит в определитель $|A|$ со знаком $(-1)^{s+t}$.

§ 4. Свойства определителей

В § 2 уже отмечалось, что вычисление определителей произвольного порядка значительно облегчается, если пользоваться их свойствами. В настоящем параграфе будут сформулированы и доказаны основные свойства определителей.

Наряду с матрицей

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

рассмотрим матрицу

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

которая получается из A заменой строчек на столбцы. Строчки матрицы A' являются, стало быть, столбцами для A. Можно также сказать, что A' получается из A переворачиванием вокруг главной диагонали. Переход от матрицы A к матрице A' называется *транспонированием*, а сама матрица A' называется *транспонированной с* A.

Непосредственной проверкой легко установить, что при $n=2$ и $n=3$ определитель транспонированной матрицы равен определителю исходной матрицы. Оказывается, что это свойство сохраняется и для любого n.

Свойство 1. *При транспонировании определитель матрицы не меняется. Другими словами, определитель транспонированной матрицы равен определителю исходной.*

Доказательство. Возьмем произвольное произведение

$$P = a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n},$$

входящее в состав определителя $|A|$ (вторые индексы $\alpha_1, \alpha_2, \dots, \alpha_n$ различны). Все сомножители в этом произведении являются также и элементами матрицы A', взятыми по одному из каждой строчки и по одному из каждого столбца, а значит, произведение P входит и в состав определителя $|A'|$. Аналогичным образом убеждаемся, что всякое произведение, входящее в $|A'|$, входит и в $|A|$. Таким образом, определители $|A|$ и $|A'|$ составляются из одних и тех же произведений. Для доказательства свойства 1 остается теперь проверить, что каждое произведение P входит в наши определители $|A|$ и $|A'|$ с одним и тем же знаком.

Если через s мы обозначим число инверсий в перестановке $(\alpha_1\ \alpha_2 \dots \alpha_n)$, то, согласно правилу для знака, произведение P будет входить в $|A|$ со знаком $(-1)^s$.

Займемся теперь определением знака, с которым P входит в определитель $|A'|$. Раньше, в частности при формулировке правила для знака, элементы матрицы мы обозначали двумя индексами, из которых первый указывал на номер строчки, а второй — на номер столбца. В матрице A' в результате операции транспонирования обозначения оказались другими: здесь первый индекс обозначает номер столбца, а второй — номер строчки. Чтобы в A' систему обозначений сделать прежней, введем для ее элементов новые обозначения, полагая $b_{ij}=a_{ji}$. Тогда матрица A' примет вид

$$A'=\begin{Vmatrix} b_{11} & b_{12} \dots b_{1n} \\ b_{21} & a_{22} \dots b_{2n} \\ \dots & \dots \\ b_{n1} & b_{n2} \dots b_{nn} \end{Vmatrix}.$$

Теперь у каждого элемента первый индекс обозначает номер строчки, а второй — номер столбца.

Произведение P в новых обозначениях имеет вид

$$P=b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_n n}. \tag{1}$$

Перепишем это произведение в порядке следования строчек матрицы A', т. е. так, чтобы первые индексы расположились в натуральном порядке:

$$P=b_{1\beta_1} b_{2\beta_2} \dots b_{n\beta_n}. \tag{2}$$

Если через t мы обозначим число инверсий в перестановке $(\beta_1\ \beta_2 \dots \beta_n)$ номеров столбцов, то, согласно правилу для знака, произведение P будет входить в определитель $|A'|$ матрицы A' со знаком $(-1)^t$.

Мы должны показать, что $(-1)^s=(-1)^t$. Для этого достаточно, очевидно, установить, что числа s и t имеют одинаковую четность.

В силу теоремы 1 из § 3 от произведения (1) мы можем перейти к произведению (2) при помощи транспозиций (под транспозицией понимается здесь перемена местами двух сомножителей). Пусть для этого понадобилось r транспозиций. При транспозиции сомножителей одновременно претерпевают транспозицию прикрепленные к ним первые и вторые индексы. При этом в результате выполнения наших r транспозиций перестановка $(\alpha_1\ \alpha_2 \dots \alpha_n)$ первых индексов переходит в натуральное расположение, а натуральное расположение вторых индексов переходит в перестановку

$(\beta_1\ \beta_2\ \dots\ \beta_n)$. Запишем последние два факта в виде

$$(\alpha_1\ \alpha_2\ \dots\ \alpha_n) \to \dots \to (1\ 2\ \dots\ n),$$
$$(1\ 2\ \dots\ n) \underbrace{\to \dots \to}_{r \text{ транспозиций}} (\beta_1\ \beta_2\ \dots\ \beta_n).$$

Предположим, что число s четное. Тогда перестановки $(\alpha_1\ \alpha_2\ \dots\ \alpha_n)$ и $(1\ 2\ \dots\ n)$ имеют одно и то же наименование (обе четные), а значит, по следствию теоремы 2 из § 3 число транспозиций r должно быть четным (при нечетном r наименование изменилось бы). Но при четном r перестановка $(\beta_1\ \beta_2\ \dots\ \beta_n)$ также четная, так как ее наименование должно совпадать с наименованием натурального расположения. Этим доказано, что при четном s число t также четное.

Аналогичным образом устанавливаем, что при нечетном s число t также нечетное (в этом случае число транспозиций r будет нечетным).

Таким образом, числа s и t всегда имеют одинаковую четность, и, следовательно, $(-1)^s = (-1)^t$.

Итак, определители $|A|$ и $|A'|$ составлены из одних и тех же $n!$ произведений, при этом произвольно взятое произведение P входит в оба определителя с одним и тем же знаком. Следовательно,

$$|A| = \sum_{(\alpha_1\alpha_2\dots\alpha_n)} (-1)^s a_{1\alpha_1}a_{2\alpha_2}\dots a_{n\alpha_n} = \sum_P (-1)^s P =$$
$$= \sum_P (-1)^t P = \sum_{(\beta_1\beta_2\dots\beta_n)} (-1)^t b_{1\beta_1}b_{2\beta_2}\dots b_{n\beta_n} = |A'|.$$

Свойство 1 доказано.

Замечания. 1) При доказательстве свойства 1 нами установлено, что числа s и t имеют одинаковую четность. Можно было бы показать (см. упражнение 4 из § 3), что в действительности $s = t$.

2) Доказательство свойства 1 было бы намного проще, если бы мы воспользовались утверждением упражнения 5 из § 3.

Доказанное свойство 1 означает, что с точки зрения теории определителей строчки и столбцы матрицы занимают равноправное положение. Если поэтому нам будет известно некоторое свойство определителей, относящееся к строчкам, то можно будет сказать, что такое же свойство имеет место и по отношению к столбцам.

Свойство 2. *Если в квадратной матрице поменять местами две строчки (или два столбца), оставив остальные на своих местах, то определитель полученной матрицы будет равен определителю исходной с противоположным знаком. Короче: при перемене местами двух строчек (или столбцов) определитель меняет знак.*

Доказательство. Согласно свойству 1 достаточно ограничиться рассмотрением случая, когда переставляются строчки.

Предположим, что в матрице

$$A=\begin{Vmatrix} a_{11} & a_{12} \dots a_{1n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{i1} & a_{i2} \dots a_{in} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{j1} & a_{j2} \dots a_{jn} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} \dots a_{nn} \end{Vmatrix}$$

n-го порядка меняются местами i-я и j-я строчки. В результате этой транспозиции строчек мы получаем матрицу

$$B=\begin{array}{ll} \begin{Vmatrix} a_{11} & a_{12} \dots a_{1n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{j1} & a_{j2} \dots a_{jn} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{i1} & a_{i2} \dots a_{in} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} & a_{n2} \dots a_{nn} \end{Vmatrix} & \begin{array}{l} \\ \\ i\text{-я строчка} \\ \\ j\text{-я строчка} \\ \\ \\ \end{array} \end{array}$$

Нам надо доказать, что $|B|=-|A|$.

Возьмем произвольное произведение

$$P=a_{1\alpha_1}a_{2\alpha_2}\dots a_{i\alpha_i}\dots a_{j\alpha_j}\dots a_{n\alpha_n}, \tag{3}$$

входящее в состав определителя $|A|$. Все сомножители в этом произведении являются также и элементами матрицы B, причем и в матрице B они принадлежат разным строчкам и разным столбцам. Но это значит, что P входит и в состав определителя $|B|$. Ясно, что всякое произведение, входящее в $|B|$, будет входить также и в $|A|$. Следовательно, состав всех $n!$ произведений для обоих определителей один и тот же. Для доказательства свойства 2 нам достаточно теперь установить, что произведение P входит в определители $|A|$ и $|B|$ с противоположными знаками.

В произведении (3) сомножители расположены в порядке следования строчек матрицы A. Если поэтому через s мы обозначим число инверсий в перестановке

$$(\alpha_1 \, \alpha_2 \ldots \alpha_i \ldots \alpha_j \ldots \alpha_n) \tag{4}$$

номеров столбцов, то знак произведения P, с которым оно входит в $|A|$, будет равен $(-1)^s$.

Для определения знака, с которым P входит в $|B|$, введем для элементов матрицы B новые обозначения, полагая

$$a_{i\alpha} = b_{j\alpha},$$
$$a_{j\alpha} = b_{i\alpha},$$
$$a_{k\alpha} = b_{k\alpha} \quad \text{при } k \neq i \text{ и } k \neq j.$$

В этих обозначениях матрица B принимает вид

$$B = \begin{array}{c} \left\| \begin{array}{llll} b_{11} & b_{12} & \ldots & b_{1n} \\ \cdot & \cdot & \cdot\;\cdot & \cdot \\ b_{i1} & b_{i2} & \ldots & b_{in} \\ \cdot & \cdot & \cdot\;\cdot & \cdot \\ b_{j1} & b_{j2} & \ldots & b_{jn} \\ \cdot & \cdot & \cdot\;\cdot & \cdot \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{array} \right\| \end{array} \begin{array}{l} \\ \\ i\text{-я строчка} \\ \\ j\text{-я строчка} \\ \\ \\ \end{array}$$

(индексы соответствуют теперь номерам строчек и столбцов). Перепишем произведение (3), используя введенные обозначения:

$$P = b_{1\alpha_1} b_{2\alpha_2} \ldots b_{j\alpha_i} \ldots b_{i\alpha_j} \ldots b_{n\alpha_n}.$$

Чтобы применить правило для определения знака, мы должны сомножители здесь расположить в порядке следования строчек матрицы B. Для этого нам надо лишь поменять местами множители $b_{i\alpha_j}$ и $b_{j\alpha_i}$ (все остальные занимают требуемые места):

$$P = b_{1\alpha_1} b_{2\alpha_2} \ldots b_{i\alpha_j} \ldots b_{j\alpha_i} \ldots b_{n\alpha_n}.$$

Если через t мы обозначим число инверсий в перестановке

$$(\alpha_1\ \alpha_2\ \ldots\ \alpha_j\ \ldots\ \alpha_i\ \ldots\ \alpha_n), \tag{5}$$

то знак, с которым P входит в $|B|$, будет равен $(-1)^t$.

Перестановка (5) получается из (4) одной транспозицией (α_i и α_j поменялись местами). Следовательно, по теореме 2 из § 3 эти перестановки имеют противоположные наименования (числа s и t — разной четности), а значит,

$$(-1)^t = -(-1)^s.$$

Мы показали, таким образом, что всякое произведение P, входящее в $|A|$ со знаком $(-1)^s$, входит также и в $|B|$, но с противоположным знаком $-(-1)^s$. Следовательно,

$$|B| = \sum_P (-1)^t P = \sum_P -(-1)^s P =$$
$$= -\left(\sum_P (-1)^s P\right) = -|A|$$

(P пробегает все $n!$ произведений). Свойство 2 доказано.

Свойство 3. *Если квадратная матрица имеет две одинаковые строчки (или два одинаковых столбца), то ее определитель равен нулю.*

Доказательство. В силу свойства 1 опять достаточно ограничиться рассмотрением случая равных строчек. Пусть i-я и j-я строчки матрицы A одинаковы, т. е.

$$a_{i1} = a_{j1},\ a_{i2} = a_{j2},\ \ldots,\ a_{in} = a_{jn}.$$

Поменяв в матрице A эти строчки местами, мы получим матрицу B, которая ничем не будет отличаться от матрицы A. Определители одинаковых матриц, очевидно, равны, поэтому $|B| = |A|$. С другой стороны, по свойству 2 имеем $|B| = -|A|$. Таким образом, $|A| = -|A|$, откуда $2|A| = 0$, а значит, $|A| = 0$. Свойство 3 доказано.

Чтобы сформулировать следующее свойство определителей, нам необходимо предварительно ввести два новых понятия и доказать относящиеся к ним некоторые вспомогательные утверждения.

Определение. *Пусть дана матрица A n-го порядка. Вычеркнем в ней i-ю строчку и k-й столбец и сдвинем, не нарушая порядка, оставшиеся элементы. Определитель полученной матрицы $(n-1)$-го порядка называется минором элемента i-й строчки и k-го столбца матрицы A и обозначается нами через Δ_{ik}.*

Например, если

$$A=\begin{Vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{Vmatrix}, \tag{6}$$

то минорами элементов второй строчки третьего столбца и четвертой строчки четвертого столбца будут соответственно определители

$$\Delta_{23}=\begin{vmatrix} a_1 & b_1 & d_1 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix}, \qquad \Delta_{44}=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Так как определитель всякой матрицы есть число, то миноры элементов матрицы также являются числами. Для матрицы n-го порядка имеем, очевидно, n^2 различным образом составленных миноров (для каждого элемента свой минор).

Лемма 1. *Если все элементы первой строчки матрицы A, кроме a_{11}, равны нулю, то ее определитель равен произведению элемента a_{11} на его минор:*

$$|A|=\begin{vmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}=$$

$$=a_{11}\Delta_{11}=a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

Доказательство. Определитель $|A|$ состоит из $n!$ слагаемых. Однако многие из них равны нулю.

Это будут те произведения, которые содержат нулевой множитель из первой строчки. Выбросим все эти нулевые слагаемые. Тогда можно будет сказать, что определитель $|A|$ составлен из произведений вида

$$P = a_{11} a_{2\beta} a_{3\gamma} \ldots a_{n\omega}, \tag{7}$$

где индексы β, γ, ..., ω образуют, очевидно, перестановку чисел 2, 3, ..., n. Определитель $|A|$ мы можем, следовательно, записать в виде суммы

$$|A| = \sum_{(\beta\gamma \ldots \omega)} \pm a_{11} a_{2\beta} a_{3\gamma} \ldots a_{n\omega},$$

в которой суммирование ведется по всем $(n-1)!$ перестановкам $(\beta\ \gamma \ldots \omega)$ из чисел 2, 3, ..., n (перед каждым произведением знак, разумеется, поставлен в соответствии с правилом для знака). Все слагаемые этой суммы содержат общий множитель a_{11}. После вынесения его за скобку мы получим

$$|A| = a_{11} \left(\sum_{(\beta\gamma \ldots \omega)} \pm a_{2\beta} a_{3\gamma} \ldots a_{n\omega} \right). \tag{8}$$

Докажем, что выражение, стоящее в скобках правой части равенства (8), равно определителю $|A^*|$ матрицы $(n-1)$-го порядка

$$A^* = \left\| \begin{matrix} a_{22} & a_{23} & \ldots & a_{2n} \\ a_{32} & a_{33} & \ldots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n2} & a_{n3} & \ldots & a_{nn} \end{matrix} \right\|,$$

которая получается из A удалением первой строчки и первого столбца.

Прежде всего замечаем, что слагаемыми рассматриваемой суммы являются все $(n-1)!$ произведений элементов матрицы A^*, взятых по одному из каждой строчки и по одному из каждого столбца и снабженных некоторым знаком. Для доказательства нашего утверждения надо, следовательно, лишь убедиться, что перед каждым произведением стоит именно тот знак, какой требуется согласно правилу для знака. Другими словами, надо убедиться, что произведение (7) и произведение

$$P^* = a_{2\beta} a_{3\gamma} \ldots a_{n\omega}$$

входят соответственно в определители $|A|$ и $|A^*|$ с одним и тем же знаком.

Чтобы для элементов матрицы A^* индексы соответствовали номерам строчек и столбцов, мы должны все индексы уменьшить на единицу. После такого изменения обозначений вторые индексы в произведении P^* будут образовывать перестановку $(\beta - 1, \gamma - 1, \ldots, \omega - 1)$. Число инверсий в ней, очевидно, такое же, как и в перестановке $(\beta\ \gamma\ \ldots\ \omega)$ чисел $2, 3, \ldots, n$. Обозначим это общее число инверсий через s. Тогда знак, с которым P^* входит в $|A^*|$, будет равен $(-1)^s$. С другой стороны, число инверсий в перестановке $(1\ \beta\ \gamma\ \ldots\ \omega)$ также равно s, так как стоящая на первом месте единица не образует инверсий с последующими числами. Следовательно, произведение (7) входит в $|A|$ с тем же знаком $(-1)^s$.

Таким образом, доказано, что P и P^* входят в $|A|$ и $|A^*|$ с одинаковыми знаками, а значит, равенство (8) можно переписать в виде $|A| = a_{11}|A^*|$. Но определитель $|A^*|$ равен минору Δ_{11}, поэтому $|A| = a_{11}\Delta_{11}$, и лемма 1 доказана.

Замечание. Лемма 1 справедлива также и для случая, когда нулевые элементы расположены в первом столбце (достаточно применить свойство 1).

Рассмотрим теперь матрицу

$$A = \begin{Vmatrix} a_{11} & \ldots & a_{1,k-1} & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i-1,1} & \ldots & a_{i-1,k-1} & a_{i-1,k} & a_{i-1,k+1} & \ldots & a_{i-1,n} \\ 0 & \ldots & 0 & a_{ik} & 0 & \ldots & 0 \\ a_{i+1,1} & \ldots & a_{i+1,k-1} & a_{i+1,k} & a_{i+1,k+1} & \ldots & a_{i+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{n,k-1} & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{Vmatrix},$$

у которой равны нулю элементы произвольно взятой i-й строчки (кроме одного элемента a_{ik}), и попытаемся распространить утверждение леммы 1 на этот более общий случай. Совершая перестановки строчек и столбцов, преобразуем нашу матрицу так, чтобы к ней можно было применить лемму 1. Будем передвигать i-ю строчку вверх, переставляя ее последовательно с соседними строчками (чтобы не нарушить взаимного расположения других строчек). Так как нам придется сделать $i-1$ перестановок строчек, а при каждой перестановке определитель меняет знак (свойство 2), то, переходя к определителям, получим

$$|A| = (-1)^{i-1} \begin{vmatrix} 0 & \dots 0 & a_{ik} & 0 & \dots 0 \\ a_{11} & \dots a_{1,k-1} & a_{1k} & a_{1,k+1} & \dots a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i-1,1} & \dots a_{i-1,k-1} & a_{i-1,k} & a_{i-1,k+1} & \dots a_{i-1,n} \\ a_{i+1,1} & \dots a_{i+1,k-1} & a_{i+1,k} & a_{i+1,k+1} & \dots a_{i+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots a_{n,k-1} & a_{nk} & a_{n,k+1} & \dots a_{nn} \end{vmatrix}.$$

Строчка с нулями находится теперь наверху. Чтобы элемент a_{ik} оказался в левом верхнем углу, надо k-й столбец поставить на место первого. Переставляя его $k-1$ раз с соседними слева столбцами, мы придем, в силу свойства 2 для столбцов, к равенству

$$|A| = (-1)^{i-1}(-1)^{k-1} \begin{vmatrix} a_{ik} & 0 & \dots 0 & 0 & \dots 0 \\ a_{1k} & a_{11} & \dots a_{1,k-1} & a_{1,k+1} & \dots a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i-1,k} & a_{i-1,1} & \dots a_{i-1,k-1} & a_{i-1,k+1} & \dots a_{i-1,n} \\ a_{i+1,k} & a_{i+1,1} & \dots a_{i+1,k-1} & a_{i+1,k+1} & \dots a_{i+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{nk} & a_{n1} & \dots a_{n,k-1} & a_{n,k+1} & \dots a_{nn} \end{vmatrix}.$$

К матрице, стоящей под знаком определителя в правой части полученного равенства, можно применить теперь лемму 1. Замечая, что числа $i+k-2$ и $i+k$ имеют одинаковую четность, мы последнее равенство, согласно лемме 1, можем переписать так:

$$|A| = a_{ik}(-1)^{i+k} \begin{vmatrix} a_{11} & \dots a_{1,k-1} & a_{1,k+1} & \dots a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i-1,1} & \dots a_{i-1,k-1} & a_{i-1,k+1} & \dots a_{i-1,n} \\ a_{i+1,1} & \dots a_{i+1,k-1} & a_{i+1,k+1} & \dots a_{i+1,n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots a_{n,k-1} & a_{n,k+1} & \dots a_{nn} \end{vmatrix}$$

или

$$|A| = a_{ik}(-1)^{i+k}\Delta_{ik}, \tag{9}$$

где Δ_{ik} — минор элемента a_{ik}.

Полученный результат приводит нас к выводу о целесообразности следующего определения.

Определение. *Пусть A — матрица n-го порядка. Минор Δ_{ik} элемента i-й строчки и k-го столбца матрицы A, взятый со знаком $(-1)^{i+k}$, называется алгебраическим дополнением этого элемента.*

Если алгебраическое дополнение элемента a_{ik} обозначено через A_{ik}, то, согласно определению, имеем

$$A_{ik} = (-1)^{i+k}\Delta_{ik}. \tag{10}$$

Например, для элементов c_2 и d_4 матрицы (6) алгебраические дополнения соответственно равны

$$A_{23} = (-1)^{2+3}\Delta_{23} = -\begin{vmatrix} a_1 & b_1 & d_1 \\ a_3 & b_3 & d_3 \\ a_4 & b_4 & d_4 \end{vmatrix},$$

$$A_{44} = (-1)^{4+4}\Delta_{44} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Для элементов, расположенных на главной диагонали, алгебраическое дополнение совпадает с минором: $A_{ii} = (-1)^{2i}\Delta_{ii} = \Delta_{ii}$. Миноры двух соседних элементов строчки (или столбца) всегда приобретают противоположные знаки, так как при изменении i или k на одну единицу четность числа $i+k$ меняется. Распределение знаков у миноров (для получения алгебраических дополнений) можно изобразить следующим образом:

+	−	+	−
−	+	−	+
+	−	+	−
−	+	−	+

Приняв введенное определение алгебраического дополнения, мы получаем, в силу равенства (9), следующую лемму.

Лемма 2. *Если все элементы i-й строчки матрицы A, кроме элемента a_{ik}, равны нулю, то ее определитель равен произведению элемента a_{ik} на его алгебраическое дополнение A_{ik}:*

$$|A| = a_{ik}A_{ik}.$$

В дальнейшем мы часто будем пользоваться следующим почти очевидным фактом.

Лемма 3. *Если в матрицах A и B n-го порядка все строчки, за исключением i-й, одинаковы, то алгебраические дополнения соответствующих элементов i-й строчки этих матриц равны между собой.*

Действительно, вычеркивая в матрицах A и B i-ю строчку и k-й столбец, мы получим одинаковые матрицы $n-1$-го порядка. Следовательно, для элементов i-й строчки и k-го столбца матриц A и B миноры равны. Так как оба минора снабжаются одним и тем же знаком $(-1)^{i+k}$, то алгебраические дополнения также получаются равными.

Очевидно, что леммы 2 и 3 справедливы также и по отношению к столбцам.

Сохраняя для матрицы A n-го порядка обозначение, введенное в начале параграфа, переходим к формулировке следующего свойства определителей.

Свойство 4. *Определитель матрицы A n-го порядка равен сумме произведений всех элементов какой-нибудь одной фиксированной строчки на их алгебраические дополнения, т. е. для любого $i = 1, 2, \ldots, n$ имеет место равенство*

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}, \tag{11}$$

называемое разложением определителя $|A|$ по элементам i-й строчки.

Аналогично для любого $k = 1, 2, \ldots, n$ имеет место разложение определителя $|A|$ по элементам k-го столбца:

$$|A| = a_{1k}A_{1k} + a_{2k}A_{2k} + \ldots + a_{nk}A_{nk}. \tag{12}$$

Доказательство. Наряду с матрицей A рассмотрим матрицу

$$B=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n} \\ x_1 & x_2 & \dots & x_n \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}, \tag{13}$$

получающуюся из A заменой элементов ее i-й строчки значениями $x_1, x_2, \dots, x_n$, которые будем считать переменными. Согласно лемме 3 алгебраические дополнения $A_{i1}, A_{i2}, \dots, A_{in}$ элементов i-й строчки матрицы A будут также алгебраическими дополнениями соответствующих элементов i-й строчки матрицы B при любых значениях $x_1, x_2, \dots, x_n$.

Каждое слагаемое, входящее в определитель $|B|$, содержит один, и только один, сомножитель из i-й строчки, т. е. содержит одну, и только одну, из переменных $x_1, x_2, \dots, x_n$. Обозначим через S_1 сумму тех слагаемых из $|B|$, которые содержат x_1, через S_2 — сумму тех слагаемых, которые содержат x_2, и т. д., наконец, через S_n — сумму тех слагаемых, которые содержат x_n. Ясно, что тогда

$$|B|=S_1+S_2+\dots+S_n.$$

Так как все слагаемые суммы S_k содержат множитель x_k, то

$$S_k=x_kT_k,$$

где T_k — некоторое число, не зависящее от переменных $x_1, x_2, \dots, x_n$. Имеем, следовательно,

$$|B|=x_1T_1+\dots+x_kT_k+\dots+x_nT_n. \tag{14}$$

Выберем теперь для наших переменных следующие значения:

$$x_1=0, \dots, x_{k-1}=0, x_k=1, x_{k+1}=0, \dots, x_n=0. \tag{15}$$

При этих значениях правая часть равенства (14) будет равна T_k. Левая же часть обратится в определитель

$$\begin{vmatrix} a_{11} & \ldots & a_{1k} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i-1,1} & \ldots & a_{i-1,k} & \ldots & a_{i-1,n} \\ 0 & \ldots & 1 & \ldots & 0 \\ a_{i+1,1} & \ldots & a_{i+1,k} & \ldots & a_{i+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nk} & \ldots & a_{nn} \end{vmatrix}.$$

Так как здесь все элементы i-й строчки, кроме одного, равны нулю, то, по лемме 2, последний определитель равен $1 \cdot A_{ik} = A_{ik}$. Таким образом, при значениях (15) равенство (14) принимает вид

$$A_{ik} = T_k \qquad (k = 1, 2, \ldots, n).$$

Подставляя найденные значения для $T_1, T_2, \ldots, T_n$ в (14), получаем тождество

$$|B| = x_1 A_{i1} + x_2 A_{i2} + \ldots + x_n A_{in}, \tag{16}$$

справедливое при любых значениях переменных $x_1, x_2, \ldots, x_n$. Положим теперь в равенстве (16)

$$x_1 = a_{i1},\ x_2 = a_{i2},\ \ldots,\ x_n = a_{in}.$$

Так как при этих значениях переменных матрица B превращается в A, то получаем разложение (11).

Что касается разложения (12), то оно может быть доказано точно таким же образом (надо будет использовать леммы 2 и 3 для столбцов). Можно также для доказательства (12) применить свойство 1.

Доказанное свойство 4 позволяет вычисление определителя n-го порядка свести к вычислению n определителей $(n-1)$-го порядка.

Пример 1. Вычислить определитель

$$\Delta = \begin{vmatrix} 2 & -1 & 1 & 1 \\ a & b & c & d \\ 1 & 1 & 2 & -1 \\ -1 & 1 & 1 & 2 \end{vmatrix},$$

разложив его по элементам второй строчки.

Получим

$$\Delta = aA_{21} + bA_{22} + cA_{23} + dA_{24} =$$

$$= -a\Delta_{21} + b\Delta_{22} - c\Delta_{23} + d\Delta_{24} =$$

$$= -a\begin{vmatrix} -1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{vmatrix} + b\begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ -1 & 1 & 2 \end{vmatrix} -$$

$$-c\begin{vmatrix} 2 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 2 \end{vmatrix} + d\begin{vmatrix} 2 & -1 & 1 \\ 1 & 1 & 2 \\ -1 & 1 & 1 \end{vmatrix} =$$

$$= 9a + 12b - 9c + 3d.$$

Пример 2. Вычислить определитель

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

В данном случае для разложения целесообразно выбрать третий столбец, так как наличие нулевых элементов дает возможность не вычислять соответствующих алгебраических дополнений (произведение нуля на любое число равно нулю). Получим

$$\Delta = 0\cdot A_{13} + 0\cdot A_{23} + 0\cdot A_{33} + 3\cdot A_{43} = -3\Delta_{43} =$$

$$= -3\begin{vmatrix} -x & 1 & -1 \\ -1 & -x & 1 \\ 1 & -1 & -x \end{vmatrix} = 3x^3 + 9x.$$

Пример 3. Вычислить определитель n-го порядка

$$\Delta = \begin{vmatrix} a & -b & 0 & \dots & 0 & 0 \\ 0 & a & -b & \dots & 0 & 0 \\ 0 & 0 & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a & -b \\ -b & 0 & 0 & \dots & 0 & a \end{vmatrix}. \qquad (17)$$

Разложив наш определитель по элементам первого столбца, получим

$$\Delta = a\begin{vmatrix} a & -b & \dots & 0 & 0 \\ 0 & a & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & a & -b \\ 0 & 0 & \dots & 0 & a \end{vmatrix}_{n-1} +$$

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

Так как в первом определителе нули — под главной диагональю, а во втором — над главной диагональю, то оба они равны произведению элементов, расположенных на главной диагонали. Таким образом,

$$\Delta = a\cdot a^{n-1} + (-b)(-1)^{n+1}(-b)^{n-1} = a^n - b^n.$$

Свойство 5. *Сумма произведений элементов одной строчки матрицы A n-го порядка на алгебраические дополнения соответствующих элементов другой строчки равна нулю, т. е.*

$$a_{j1}A_{i1} + a_{j2}A_{i2} + \ldots + a_{jn}A_{in} = 0 \quad \textit{при} \quad j \neq i. \qquad (18)$$

Аналогично для столбцов

$$a_{1l}A_{1k} + a_{2l}A_{2k} + \ldots + a_{nl}A_{nk} = 0 \quad \textit{при} \quad l \neq k. \qquad (19)$$

Доказательство. Обратимся к равенству (16) из доказательства свойства 4. Если мы положим

$$x_1 = a_{j1},\ x_2 = a_{j2},\ \ldots,\ x_n = a_{jn},$$

то правая часть равенства (16) превратится в сумму $a_{j1}A_{i1} + a_{j2}A_{i2} + \ldots + a_{jn}A_{in}$. Что касается матрицы B, то она превратится в матрицу, у которой строчка из элементов a_{j1}, a_{j2}, ..., a_{jn} встречается дважды (на месте i-й и j-й строчек). Так как по свойству 3 определитель такой матрицы равен нулю, то получаем, следовательно, равенство (18). Равенство (19) можно доказать либо таким же способом, либо с помощью свойства 1.

Свойство 6. *Если все элементы какой-нибудь строчки (или столбца) матрицы n-го порядка умножить на число c, то ее определитель также умножится на c.*

Доказательство. Умножив все элементы i-й строчки матрицы

$$A=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}$$

на число c, мы получим матрицу

$$A_1=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ ca_{i1} & ca_{i2} & \dots & ca_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}.$$

По лемме 3 алгебраические дополнения $A_{i1}, A_{i2}, \dots$ $\dots, A_{in}$ элементов i-й строчки матрицы A равны алгебраическим дополнениям соответствующих элементов i-й строчки матрицы A_1. Разложив определитель $|A_1|$ по элементам i-й строчки, мы получим

$$|A_1| = ca_{i1}A_{i1} + ca_{i2}A_{i2} + \dots + ca_{in}A_{in} =$$
$$= c(a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}).$$

Так как выражение в скобках равно $|A|$ (свойство 4), то $|A_1| = c|A|$. Этим свойство 6 для строчек доказано. В силу свойства 1 оно справедливо и для столбцов.

Свойство 6 может быть сформулировано также следующим образом: общий множитель для элементов какой-нибудь строчки (или столбца) матрицы n-го порядка можно выносить за знак определителя.

Свойство 7. *Если матрица n-го порядка имеет две пропорциональные строчки (или два пропорциональных столбца), то ее определитель равен нулю.*

Доказательство. Пусть i-я и j-я строчки матрицы n-го порядка пропорциональны, т. е.

$$a_{j1} = ca_{i1},\ a_{j2} = ca_{i2},\ \dots,\ a_{jn} = ca_{in}.$$

Тогда

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \ldots & a_{jn} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = \begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ ca_{i1} & ca_{i2} & \ldots & ca_{in} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} =$$

$$= c \begin{vmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = c \cdot 0 = 0.$$

(Здесь мы применили свойство 6 о вынесении общего множителя из строчки за знак определителя и свойство 3 о равенстве нулю определителя матрицы с равными строчками.) Свойство 7 для столбцов следует из свойства 1.

Свойство 8. *Если все элементы i-й строчки матрицы n-го порядка представлены в виде суммы двух слагаемых*

$$a_{ik} = a'_{ik} + a''_{ik} \qquad (k = 1, 2, \ldots, n), \tag{20}$$

то ее определитель можно представить в виде суммы определителей двух матриц, у которых элементами i-й строчки являются соответственно первые и вторые слагаемые разложений (20), а все остальные строчки — такие же, как у исходной матрицы:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{i1} + a''_{i1} & a'_{i2} + a''_{i2} & \ldots & a'_{in} + a''_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{i1} & a'_{i2} & \ldots & a'_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a''_{i1} & a''_{i2} & \ldots & a''_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}. \tag{21}$$

Аналогичное утверждение, разумеется, имеет место и для столбцов.

Доказательство. По лемме 3 алгебраические дополнения элементов i-й строчки для всех трех матриц, стоящих под знаком определителя в (21), одинаковы. Обозначим их через $A_{i1}, A_{i2}, \ldots, A_{in}$. По свойству 4 определитель слева равен

$$(a'_{i1}+a''_{i1})A_{i1}+(a'_{i2}+a''_{i2})A_{i2}+\ldots+(a'_{in}+a''_{in})A_{in}. \quad (22)$$

Аналогично, применив то же свойство к определителям справа, получим, что их сумма равна

$$(a'_{i1}A_{i1}+a'_{i2}A_{i2}+\ldots+a'_{in}A_{in})+ \\ +(a''_{i1}A_{i1}+a''_{i2}A_{i2}+\ldots+a''_{in}A_{in}). \quad (23)$$

Сравнивая выражения (22) и (23), видим, что они равны между собой, т. е. левая часть равенства (21) действительно равна правой. Свойство 8 доказано.

Ясно, что если все элементы какой-нибудь строчки представлены в виде суммы k слагаемых, то определитель можно представить в виде суммы k определителей.

Пример 1. Найдем значение определителя

$$\Delta=\begin{vmatrix} 1+2a & 1 & a & x \\ 1+2b & 2 & b & x \\ 1+2c & 3 & c & x \\ 1+2d & 4 & d & x \end{vmatrix}.$$

Элементы первого столбца являются здесь суммами двух слагаемых, поэтому согласно свойству 8 имеем

$$\Delta=\begin{vmatrix} 1 & 1 & a & x \\ 1 & 2 & b & x \\ 1 & 3 & c & x \\ 1 & 4 & d & x \end{vmatrix}+\begin{vmatrix} 2a & 1 & a & x \\ 2b & 2 & b & x \\ 2c & 3 & c & x \\ 2d & 4 & d & x \end{vmatrix}.$$

В первом определителе первый столбец пропорционален последнему, во втором же первый столбец пропорционален третьему. Следовательно, по свойству 7 оба они равны нулю, а значит, $\Delta=0$.

Пример 2. Вычислить определитель n-го порядка

$$\Delta=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 2 & 2 & \dots & 2 & 2 \\ 0 & 0 & 3 & \dots & 3 & 3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 & n-1 \\ x & x & x & \dots & x & n+x \end{vmatrix}.$$

Элементы последней строчки представим в виде сумм

$$0+x,\ 0+x,\ \dots,\ 0+x,\ n+x.$$

Тогда

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

$$+\begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 2 & 2 & \dots & 2 & 2 \\ 0 & 0 & 3 & \dots & 3 & 3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 & n-1 \\ x & x & x & \dots & x & x \end{vmatrix}.$$

В первом определителе под главной диагональю везде нули, поэтому он равен произведению элементов главной диагонали, т. е. $1\cdot 2\cdot 3\ \dots\ (n-1)\cdot n=n!$. Второй определитель равен нулю, так как у него первая и последняя строчки пропорциональны. Таким образом $\Delta=n!+0=n!$.

Свойство 9. *Определитель матрицы n-го порядка не изменится, если к элементам одной ее строчки прибавить соответствующие элементы другой строчки, умноженные на одно и то же произвольное число. То же самое справедливо и для столбцов.*

Доказательство. Для разнообразия доказательство этого свойства проведем для столбцов. Пусть к элементам k-го столбца матрицы

$$A=\begin{Vmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1l} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & \dots & a_{2l} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nl} & \dots & a_{nn} \end{Vmatrix}$$

прибавлены соответствующие элементы l-го столбца, умноженные на число c. Для определителя полученной матрицы A_1 имеем

$$|A_1|=\begin{vmatrix} a_{11} & \dots & a_{1k}+ca_{1l} & \dots & a_{1l} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k}+ca_{2l} & \dots & a_{2l} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nk}+ca_{nl} & \dots & a_{nl} & \dots & a_{nn} \end{vmatrix}=$$

$$=\begin{vmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1l} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & \dots & a_{2l} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nk} & \dots & a_{nl} & \dots & a_{nn} \end{vmatrix}+$$

$$+\begin{vmatrix} a_{11} & \dots & ca_{1l} & \dots & a_{1l} & \dots & a_{1n} \\ a_{21} & \dots & ca_{2l} & \dots & a_{2l} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & ca_{nl} & \dots & a_{nl} & \dots & a_{nn} \end{vmatrix}=|A|+0=|A|$$

(при выкладках мы использовали свойства 8 и 7 для столбцов). Таким образом, $|A_1|=|A|$, а это и требовалось доказать.

Пример 1. Вычислить определитель

$$\Delta=\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 5 & 9 \\ 0 & 0 & 3 & 7 \\ -2 & -4 & -6 & 1 \end{vmatrix}.$$

Здесь целесообразно первую строчку, умноженную на 2, прибавить к четвертой. Так как при таком преобразовании определитель не меняется, то

$$\Delta=\begin{vmatrix}1&2&3&4\\0&2&5&9\\0&0&3&7\\0&0&0&9\end{vmatrix}=1\cdot2\cdot3\cdot9=54.$$

Пример 2. В определителе

$$\Delta=\begin{vmatrix}1&2&3&4\\1&3&3&4\\1&-1&7&4\\1&-2&5&9\end{vmatrix}$$

легко можно получить нули над главной диагональю. Для этого первый столбец, умноженный на —2, прибавим ко второму. Затем в полученном определителе первый столбец, умноженный на —3, прибавим к третьему. Во вновь полученном определителе опять первый столбец, умноженный на —4, прибавим к последнему. Так как, по свойству 9, при наших преобразованиях матрицы определитель не меняется, то в результате этих трех последовательно выполненных операций мы получим

$$\Delta=\begin{vmatrix}1&0&0&0\\1&1&0&0\\1&-3&4&0\\1&-4&2&5\end{vmatrix}=1\cdot1\cdot4\cdot5=20.$$

Пример 3. Вычислить определитель n-го порядка

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

Вычтем здесь из первого столбца второй (т. е. к первому столбцу прибавим второй, умноженный на -1), затем из второго столбца вычтем третий, из третьего — четвертый и т. д., наконец, из предпоследнего столбца вычтем последний. В результате мы получим

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

В полученном определителе первую строчку (умноженную на 1) прибавим последовательно ко всем остальным:

$$\Delta=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 2 & 2 & \dots & 2 & 3 \\ 0 & 0 & 2 & \dots & 2 & 4 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 2 & n \\ 0 & 0 & 0 & \dots & 0 & n+1 \end{vmatrix}=$$

$$=1\cdot 2\cdot 2\cdots 2\cdot(n+1)=2^{n-2}(n+1).$$

Упражнения

1. Как изменится определитель матрицы n-го порядка, если ее столбцы записать в обратном порядке?

Отв. Приобретет множитель $(-1)^{\frac{n(n-1)}{2}}$.

2. Вычислить определитель (17), пользуясь лишь определением определителя.

3. Доказать свойство 3, основываясь только на определении определителя (без использования свойства 2).

4. Вычислить определитель, разложив его по элементам последнего столбца,

$$\begin{vmatrix} 1 & 1 & 3 & x \\ 1 & 3 & 1 & y \\ 3 & 1 & 1 & z \\ 1 & 1 & 1 & t \end{vmatrix}.$$

Отв. $4x+4y+4z-20t$.

5. Вычислить определитель

$$\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & 1+x_1y_3 \\ 1+x_2y_1 & 1+x_2y_2 & 1+x_2y_3 \\ 1+x_3y_1 & 1+x_3y_2 & 1+x_3y_3 \end{vmatrix}.$$

Отв. 0.

6. Вычислить определители n-го порядка

$$\text{a)} \begin{vmatrix} 1 & 1 & 1\ldots & 1 & 1 \\ -x & a & a\ldots & a & a \\ -x & -x & a\ldots & a & a \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -x & -x & -x\ldots & a & a \\ -x & -x & -x\ldots & -x & a \end{vmatrix};$$

$$\text{b)} \begin{vmatrix} 1 & 1 & 1\ldots 1 & 1 \\ -x & x & 0\ldots 0 & 0 \\ -x & 0 & x\ldots 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ -x & 0 & 0\ldots x & 0 \\ -x & 0 & 0\ldots 0 & x \end{vmatrix}.$$

Отв. a) $(a+x)^{n-1}$; b) nx^{n-1}.

§ 5. Примеры вычисления определителей

1. Вычислить определитель

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

Если к этому определителю мы непосредственно применим свойство 4, то получим четыре определителя 3-го порядка. Но такой путь вряд ли целесообразен. Поставим перед собой цель, пользуясь свойством 9, получить, например, в первом столбце три нулевых элемента. Предварительно сделаем вспомогательные преобразования, уменьшающие числа первого столбца, а именно: к последней строчке прибавим третью, затем из третьей вычтем вторую и, наконец, из второй вычтем первую. Тогда будем иметь

$$\Delta = \begin{vmatrix} 2 & 3 & -4 & 5 \\ 1 & -8 & 6 & -1 \\ 2 & 9 & 1 & -6 \\ 1 & 6 & 8 & 1 \end{vmatrix}.$$

Если теперь четвертую строчку вычесть из второй, а затем ее же, но умноженную на 2 вычесть из первой и третьей, то получим

$$\Delta = \begin{vmatrix} 0 & -9 & -20 & 3 \\ 0 & -14 & -2 & -2 \\ 0 & -3 & -15 & -8 \\ 1 & 6 & 8 & 1 \end{vmatrix} = -\begin{vmatrix} -9 & -20 & 3 \\ -14 & -2 & -2 \\ -3 & -15 & -8 \end{vmatrix} =$$

(выносим из строчек множители -1, -2, -1 соответственно)

$$= 2\begin{vmatrix} 9 & 20 & -3 \\ 7 & 1 & 1 \\ 3 & 15 & 8 \end{vmatrix} =$$

(третий столбец вычитаем из второго и, умноженный на 7, — из первого)

$$= 2\begin{vmatrix} 30 & 23 & -3 \\ 0 & 0 & 1 \\ -53 & 7 & 8 \end{vmatrix} = -2\begin{vmatrix} 30 & 23 \\ -53 & 7 \end{vmatrix} =$$

$$= -2(210 + 1219) = -2858.$$

2. Вычислить определитель n-го порядка

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

Вычтем из первой строчки вторую, из второй — третью и т. д., наконец, из предпоследней — последнюю (последняя остается без изменения). Получим

$$\Delta=\begin{vmatrix} n-1 & -1 & -1 & \dots & -1 & -1 \\ 0 & n-1 & -1 & \dots & -1 & -1 \\ 0 & 0 & n-1 & \dots & -1 & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 & -1 \\ 1 & 1 & 1 & \dots & 1 & n \end{vmatrix}_n .$$

Элементы последней строчки представим в виде сумм двух слагаемых:

$$0+1,\ 0+1,\ \dots,\ 0+1,\ (n-1)+1.$$

Применив свойство 8, будем иметь

$$\Delta=\begin{vmatrix} n-1 & -1 & -1 & \dots & -1 & -1 \\ 0 & n-1 & -1 & \dots & -1 & -1 \\ 0 & 0 & n-1 & \dots & -1 & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 & -1 \\ 0 & 0 & 0 & \dots & 0 & n-1 \end{vmatrix}_n +$$

$$+\begin{vmatrix} n-1 & -1 & -1 & \dots & -1 & -1 \\ 0 & n-1 & -1 & \dots & -1 & -1 \\ 0 & 0 & n-1 & \dots & -1 & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 & -1 \\ 1 & 1 & 1 & \dots & 1 & 1 \end{vmatrix}_n .$$

В первом определителе нули под главной диагональю, во втором мы можем последнюю строчку прибавить ко всем предшествующим и получить нули над главной диагональю. Таким образом,

$$\Delta=(n-1)^n+\begin{vmatrix} n & 0 & 0 & \dots & 0 & 0 \\ 1 & n & 0 & \dots & 0 & 0 \\ 1 & 1 & n & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & n & 0 \\ 1 & 1 & 1 & \dots & 1 & 1 \end{vmatrix}=(n-1)^n+n^{n-1}.$$

3. Определитель вида

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

называется *определителем Вандермонда порядка n* (*или степенным определителем*).

Вычтем в определителе Δ_n из каждого столбца, начиная с последнего, предыдущий столбец, умноженный на x_n. В результате этих операций последняя строчка примет вид

$$1,\ 0,\ 0,\ \dots,\ 0.$$

Что касается произвольной i-й строчки при $1 \leqslant i \leqslant n-1$, то она будет выглядеть следующим образом:

$$1,\ x_i - x_n,\ x_i(x_i - x_n),\ x_i^2(x_i - x_n),\ \dots,\ x_i^{n-2}(x_i - x_n).$$

Разлагая полученный определитель по элементам последней строчки, получим

$$\Delta_n = (-1)^{n+1} \begin{vmatrix} x_1 - x_n & x_1(x_1 - x_n) & \dots & x_1^{n-2}(x_1 - x_n) \\ x_2 - x_n & x_2(x_2 - x_n) & \dots & x_2^{n-2}(x_2 - x_n) \\ \cdot & \cdot & \cdot & \cdot \\ x_{n-1} - x_n & x_{n-1}(x_{n-1} - x_n) & \dots & x_{n-1}^{n-2}(x_{n-1} - x_n) \end{vmatrix}.$$

Вынесем теперь из строчек множители $x_1 - x_n$, $x_2 - x_n, \dots, x_{n-1} - x_n$:

$$\Delta_n = (-1)^{n-1}(x_1 - x_n)(x_2 - x_n)\dots(x_{n-1} - x_n) \times$$

$$\times \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^{n-2} \end{vmatrix}.$$

Если каждую из $n-1$ разностей в полученном равенстве мы умножим на -1, то это равенство примет вид

$$\Delta_n = (x_n - x_1)(x_n - x_2) \dots (x_n - x_{n-1}) \Delta_{n-1},$$

где Δ_{n-1} — определитель Вандермонда порядка $n-1$. С определителем Δ_{n-1} мы можем поступить точно так же, т. е. выразить его через Δ_{n-2}. В результате для

Δ_n получим выражение

$$\Delta_n=(x_n-x_1)(x_n-x_2)\ldots(x_n-x_{n-1})\times$$
$$\times(x_{n-1}-x_1)(x_{n-1}-x_2)\ldots(x_{n-1}-x_{n-2})\cdot\Delta_{n-2}.$$

Продолжая этот процесс дальше, мы на $(n-2)$-м шаге найдем выражение Δ_n через определитель Δ_2, который равен

$$\begin{vmatrix}1 & x_1\\ 1 & x_2\end{vmatrix}=(x_2-x_1).$$

В итоге получим

$$\begin{aligned}\Delta_n=(x_n-x_1)(x_n-x_2)\quad&\ldots\quad(x_n-x_{n-1})\times\\ \times(x_{n-1}-x_1)(x_{n-1}-x_2)\;&\ldots\;(x_{n-1}-x_{n-2})\times\\ \cdots\cdots\cdots&\cdots\cdots\cdots\\ \times(x_3-x_1)(x_3-x_2)&\times\\ \times(x_2-x_1).&\end{aligned}$$

Если воспользоваться знаком $\prod$, применяемым для обозначения произведений, то полученное значение для Δ_n можно записать в виде

$$\Delta_n=\prod_{n\geqslant i>j\geqslant 1}(x_i-x_j)$$

(произведение распространяется на все те пары индексов i и j среди значений $1, 2, \ldots, n$, для которых $i>j$).

Из формулы для Δ_n следует, что если значения $x_1, x_2, \ldots, x_n$ попарно различны, то определитель Вандермонда $\Delta_n=\Delta_n(x_1, x_2, \ldots, x_n)$ отличен от нуля. Действительно, все сомножители в формуле для Δ_n отличны от нуля, поэтому и их произведение отлично от нуля.

Приведем конкретный пример определителя Вандермонда:

$$\begin{vmatrix}1 & 1 & 1 & 1\\ 1 & 2 & 4 & 8\\ 1 & 3 & 9 & 27\\ 1 & 4 & 16 & 64\end{vmatrix}=$$
$$=(4-1)(4-2)(4-3)(3-1)(3-2)(2-1)=12.$$

4. Пусть нам дан определитель n-го порядка ($\alpha \neq \beta$)

$$\Delta_n = \begin{vmatrix} \alpha+\beta & \alpha\beta & 0 & \dots & 0 & 0 \\ 1 & \alpha+\beta & \alpha\beta & \dots & 0 & 0 \\ 0 & 1 & \alpha+\beta & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \alpha+\beta & \alpha\beta \\ 0 & 0 & 0 & \dots & 1 & \alpha+\beta \end{vmatrix}.$$

Представим элементы первого столбца в виде сумм двух слагаемых:

$$\alpha+\beta,\ 1+0,\ 0+0,\ \dots,\ 0+0.$$

Тогда по свойству 8 из § 4 определитель Δ_n может быть представлен в виде суммы двух определителей:

$$\Delta_n = \begin{vmatrix} \alpha & \alpha\beta & 0 & \dots & 0 & 0 \\ 1 & \alpha+\beta & \alpha\beta & \dots & 0 & 0 \\ 0 & 1 & \alpha+\beta & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \alpha+\beta & \alpha\beta \\ 0 & 0 & 0 & \dots & 1 & \alpha+\beta \end{vmatrix} + \begin{vmatrix} \beta & \alpha\beta & 0 & \dots & 0 & 0 \\ 0 & \alpha+\beta & \alpha\beta & \dots & 0 & 0 \\ 0 & 1 & \alpha+\beta & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \alpha+\beta & \alpha\beta \\ 0 & 0 & 0 & \dots & 1 & \alpha+\beta \end{vmatrix}.$$

Второй из этих определителей после разложения по первому столбцу дает нам $\beta\Delta_{n-1}$, где Δ_{n-1} — определитель такого же вида, как и исходный, но на единицу меньшего порядка.

С первым определителем поступим следующим образом. Первый столбец, умноженный на β, вычтем из второго, затем полученный второй столбец, умноженный на β, вычтем из третьего и т. д. В конечном итоге, как легко видеть, мы придем к определителю

$$\begin{vmatrix} \alpha & 0 & 0 & \dots & 0 & 0 \\ 1 & \alpha & 0 & \dots & 0 & 0 \\ 0 & 1 & \alpha & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \alpha & 0 \\ 0 & 0 & 0 & \dots & 1 & \alpha \end{vmatrix} = \alpha^n.$$

Нами получено, таким образом, соотношение

$$\Delta_n = \alpha^n + \beta\Delta_{n-1},$$

связывающее заданный определитель порядка n с определителем такой же структуры порядка $n-1$. Соотношения такого типа носят название *рекуррентных*.

Обращаясь к определителям типа Δ_n при малых n, легко находим, что

$$\Delta_1 = \alpha + \beta,$$

$$\Delta_2 = \begin{vmatrix} \alpha+\beta & \alpha\beta \\ 1 & \alpha+\beta \end{vmatrix} = \alpha^2 + \alpha\beta + \beta^2,$$

$$\Delta_3 = \begin{vmatrix} \alpha+\beta & \alpha\beta & 0 \\ 1 & \alpha+\beta & \alpha\beta \\ 0 & 1 & \alpha+\beta \end{vmatrix} = \alpha^3 + \alpha^2\beta + \alpha\beta^2 + \beta^3.$$

Полученные результаты мы можем записать также в виде

$$\Delta_1 = \frac{\alpha^2-\beta^2}{\alpha-\beta}, \qquad \Delta_2 = \frac{\alpha^3-\beta^3}{\alpha-\beta}, \qquad \Delta_3 = \frac{\alpha^4-\beta^4}{\alpha-\beta}.$$

Найденные значения подсказывают нам, что, по-видимому, для любого порядка k справедлива формула

$$\Delta_k = \frac{\alpha^{k+1}-\beta^{k+1}}{\alpha-\beta}.$$

Это пока что только догадка. Но мы можем попытаться доказать справедливость написанной формулы для всех k методом математической индукции, используя найденное рекуррентное соотношение. Предположим, что справедливость формулы уже установлена для $k = n-1$, т. е. что

$$\Delta_{n-1} = \frac{\alpha^n-\beta^n}{\alpha-\beta}.$$

Тогда

$$\Delta_n = \alpha^n + \beta\Delta_{n-1} = \alpha^n + \beta\frac{\alpha^n - \beta^n}{\alpha - \beta} = \frac{\alpha^{n+1} - \beta^{n+1}}{\alpha - \beta}.$$

Таким образом, если формула справедлива при $k = n - 1$, то она справедлива и при $k = n$, а значит, она справедлива при всех k. Выражение для определителя Δ_n найдено.

Для определителя Δ_n можно легко получить другое рекуррентное соотношение. Разложим определитель Δ_n по элементам первого столбца. Один из получающихся миноров будет равен Δ_{n-1}, второй же разложением по первой строчке сводится к Δ_{n-2}. В результате мы получим рекуррентное соотношение

$$\Delta_n = (\alpha + \beta)\,\Delta_{n-1} - \alpha\beta\Delta_{n-2},$$

связывающее определитель Δ_n с определителями такого же типа двух предшествующих порядков. Пользуясь этим новым соотношением, мы могли бы, конечно, также доказать справедливость подмеченной формулы для Δ_k.

Укажем еще на один прием вычисления рассмотренного определителя Δ_n. Вернемся для этого к выводу нашего первого рекуррентного соотношения. Поменяв роли α и β, мы тем же способом получим еще одно соотношение:

$$\Delta_n = \beta^n + \alpha\Delta_{n-1}.$$

Нахождение значения для Δ_n сводится теперь к решению системы двух линейных уравнений с двумя неизвестными Δ_n и Δ_{n-1}:

$$\left.\begin{aligned} \Delta_n &= \alpha^n + \beta\Delta_{n-1}, \\ \Delta_n &= \beta^n + \alpha\Delta_{n-1}. \end{aligned}\right\}$$

(Конечно, нас интересует только неизвестное Δ_n.)

5. Вычислить определитель

$$\Delta = \begin{vmatrix} a_1 & x & x & \dots & x \\ y & a_2 & x & \dots & x \\ y & y & a_3 & \dots & x \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y & y & y & \dots & a_n \end{vmatrix}.$$

Представим элементы первой строчки в виде сумм двух слагаемых:

$$(a_1 - x) + x,\ 0 + x,\ \ldots,\ 0 + x,$$

и разложим определитель Δ на сумму двух определителей:

$$\Delta = \begin{vmatrix} a_1 - x & 0 & 0 & \ldots & 0 \\ y & a_2 & x & \ldots & x \\ y & y & a_3 & \ldots & x \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y & y & y & \ldots & a_n \end{vmatrix} + \begin{vmatrix} x & x & x & \ldots & x \\ y & a_2 & x & \ldots & x \\ y & y & a_3 & \ldots & x \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y & y & y & \ldots & a_n \end{vmatrix}.$$

Первый определитель после разложения по элементам первой строчки записывается в виде $(a_1 - x)D$, где D — минор элемента $a_1 - x$. Во втором определителе мы вынесем из первой строчки x за знак определителя, а затем первую строчку, умноженную на y, вычтем из всех прочих строчек. В результате под главной диагональю окажутся нули, а на главной диагонали будут стоять элементы $1,\ a_2 - y,\ a_3 - y,\ \ldots,\ a_n - y$. Для определителя Δ мы получаем представление

$$\Delta = (a_1 - x) D + x (a_2 - y)(a_3 - y) \ldots (a_n - y).$$

Рассмотрим полученное равенство как тождество относительно переменных x и y. Если в этом тождестве мы поменяем местами переменные x и y, то получим некоторое новое равенство. Но определители Δ и D при перемене местами x и y не меняются, так как такая перемена сводится к транспонированию соответствующих матриц (см. свойство 1 из § 4). Поэтому мы имеем

$$\Delta = (a_1 - y) D + y (a_2 - x)(a_3 - x) \ldots (a_n - x).$$

Определитель Δ может быть найден теперь из системы двух уравнений с двумя неизвестными. Исключив D, получим окончательно

$$\Delta = \frac{x(a_1 - y)(a_2 - y)\ldots(a_n - y) - y(a_1 - x)(a_2 - x)\ldots(a_n - x)}{x - y}.$$

Рассмотренные нами примеры показывают, что вычисление определителей n-го порядка требует умелого использования свойств определителей с учетом специфических особенностей того или иного определителя.

Встречавшиеся в примерах приемы позволяют выделить следующие методы вычисления определителей:

1) для численных определителей — получение нулей в какой-нибудь строчке (или столбце) и сведение к одному определителю на единицу меньшего порядка;

2) преобразование матрицы определителя к треугольному виду (с нулями выше или ниже главной диагонали);

3) получение рекуррентного соотношения с последующим применением метода математической индукции;

4) получение двух (или нескольких) рекуррентных соотношений и нахождение определителя из системы линейных уравнений.

В конкретных примерах перечисленные основные методы используются, как правило, в сочетаниях друг с другом.

Упражнения

1. Вычислить определители

$$\text{a)}\ \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{vmatrix}; \qquad \text{b)}\ \begin{vmatrix} 7 & 8 & 5 & 5 & 3 \\ 10 & 11 & 6 & 7 & 5 \\ 5 & 3 & 6 & 2 & 5 \\ 6 & 7 & 5 & 4 & 2 \\ 7 & 10 & 7 & 5 & 0 \end{vmatrix}.$$

Отв. a) -20; b) 5.

2. Вычислить определитель

$$\begin{vmatrix} 0 & x & y & z \\ x & 0 & z & y \\ y & z & 0 & x \\ z & y & x & 0 \end{vmatrix}.$$

Отв. $x^4+y^4+z^4-2x^2y^2-2x^2z^2-2y^2z^2=$
$=(x+y+z)(x-y-z)(x+y-z)(x-y+z)$.

3. Вычислить определитель (порядка $n+1$)

$$\begin{vmatrix} x & 0 & 0 & \dots & 0 & a_0 \\ -1 & x & 0 & \dots & 0 & a_1 \\ 0 & -1 & x & \dots & 0 & a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x & a_{n-1} \\ 0 & 0 & 0 & \dots & -1 & a_n \end{vmatrix}.$$

Отв. $a_0+a_1x+a_2x^2+\dots+a_nx^n$.

Указание. К первой строчке прибавить вторую, умноженную на x, третью, умноженную на x^2, и т. д., наконец, последнюю, умноженную на x^n. Затем разложить полученный определитель по первой строчке.

4. Вычислить определитель n-го порядка

$$\begin{vmatrix} \alpha+\beta & \alpha\beta & 0 & \dots & 0 & 0 \\ 2 & \alpha+\beta & \alpha\beta & \dots & 0 & 0 \\ 0 & 1 & \alpha+\beta & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha+\beta & \alpha\beta \\ 0 & 0 & 0 & \dots & 1 & \alpha+\beta \end{vmatrix}_n .$$

Отв. $\alpha^n+\beta^n$.

Указание. Представить в виде суммы двух определителей по первому столбцу или получить рекуррентное соотношение разложением по последнему столбцу.

5. Вычислить определитель n-го порядка

$$\begin{vmatrix} 1 & a & 0 & \dots & 0 & 0 \\ 1 & 1+a & a & \dots & 0 & 0 \\ 0 & 1 & 1+a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1+a & a \\ 0 & 0 & 0 & \dots & 1 & 1+a \end{vmatrix} .$$

Отв. 1.

6. Вычислить определитель

$$\begin{vmatrix} 1 & -a & -a & -a & \dots & -a \\ 2 & 1 & -a & -a & \dots & -a \\ 3 & 2 & 1 & -a & \dots & -a \\ 4 & 3 & 2 & 1 & \dots & -a \\ \dots & \dots & \dots & \dots & \dots & \dots \\ n & n-1 & n-2 & n-3 & \dots & 1 \end{vmatrix} .$$

Отв. $(1+a)^n-a^n$.

Указание. Из каждого столбца, начиная с первого, вычесть последующий; затем представить в виде суммы двух определителей по последнему столбцу.

7. Вычислить определитель n-го порядка

$$\begin{vmatrix} x & -y & 0 & \dots & 0 & 0 \\ 0 & x & -y & \dots & 0 & 0 \\ 0 & 0 & x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x & -y \\ 1 & 1 & 1 & \dots & 1 & 1 \end{vmatrix}_n .$$

Отв. $\dfrac{x^n-y^n}{x-y}$.

Указание. Разложение определителя по элементам первого столбца и (второй раз) по элементам последнего столбца дает два различных рекуррентных соотношения.

8. Вычислить определитель n-го порядка

$$\begin{vmatrix} x+1 & 1 & 1 & \dots & 1 & 1 \\ -1 & x & 0 & \dots & 0 & 0 \\ 0 & -1 & x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x & 0 \\ 0 & 0 & 0 & \dots & -1 & x \end{vmatrix}_n .$$

Отв. $\dfrac{x^{n+1}-1}{x-1}$.

9. Вычислить определитель n-го порядка

$$\begin{vmatrix} x & a & a & \dots & a \\ b & x & 0 & \dots & 0 \\ b & 0 & x & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b & 0 & 0 & \dots & x \end{vmatrix}_n .$$

Отв. $x^n-(n-1)\,abx^{n-2}$.

10. Вычислить определитель n-го порядка

$$\begin{vmatrix} a+1 & x & x & \dots & x & x \\ 1 & a & x & \dots & x & x \\ 1 & 0 & a & \dots & x & x \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & \dots & a & x \\ 1 & 0 & 0 & \dots & 0 & a \end{vmatrix}_n .$$

Отв. $a^n+(a-x)^{n-1}$.

11. Вычислить определитель

$$\begin{vmatrix} 1+x_1 & 1 & 1 & \dots & 1 \\ 2 & 2+x_2 & 2 & \dots & 2 \\ 3 & 3 & 3+x_3 & \dots & 3 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ n & n & n & \dots & n+x_n \end{vmatrix} .$$

Отв. $x_1x_2\dots x_n\left(1+\dfrac{1}{x_1}+\dfrac{2}{x_2}+\dots+\dfrac{n}{x_n}\right)$.

§ 6. Определитель ступенчатой матрицы

Рассмотрим матрицу

$$A=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1m} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2m} & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mm} & 0 & \dots & 0 \\ a_{m+1,1} & a_{m+1,2} & \dots & a_{m+1,m} & a_{m+1,m+1} & \dots & a_{m+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nm} & a_{n,m+1} & \dots & a_{nn} \end{Vmatrix}$$

n-го порядка, у которой все элементы a_{ik} при $1\leqslant i\leqslant m$ и $m+1\leqslant k\leqslant n$ равны нулю. В дальнейшем такую матрицу будем называть *ступенчатой*. Обозначим через A_1 матрицу m-го порядка, расположенную в левом верхнем углу матрицы A, и через A_2 — матрицу порядка $n-m$, расположенную в правом нижнем углу. Матрицы A_1 и A_2 называются *диагональными клетками* ступенчатой матрицы A. Символически матрицу A можно записать также в виде

$$A=\begin{Vmatrix} A_1 & O \\ B & A_2 \end{Vmatrix},$$

где B — прямоугольная матрица, имеющая $n-m$ строк и m столбцов, а O — прямоугольная матрица, имеющая m строк и $n-m$ столбцов и состоящая сплошь из нулей.

Теорема 1. *Определитель ступенчатой матрицы равен произведению определителей ее диагональных клеток, т. е.*

$$|A|=\begin{vmatrix} A_1 & O \\ B & A_2 \end{vmatrix}=|A_1|\cdot|A_2|.$$

Доказательство. Доказательство теоремы мы проведем индукцией по порядку m верхней диагональной клетки A_1. Если $m=1$, т. е. $A_1=\|a_{11}\|$ — матрица 1-го порядка, то утверждение теоремы справедливо, так как оно в этом случае превращается в утверждение леммы 1 из § 4 (см. замечание в конце § 2).

Предположим теперь, что $m\geqslant 2$ и что для случая, когда порядок верхней диагональной клетки равен $m-1$, теорема уже доказана.

Разложим определитель $|A|$ по элементам первой строчки:

$$|A|=a_{11}A_{11}+a_{12}A_{12}+\ldots+a_{1m}A_{1m}. \qquad (1)$$

Здесь A_{1k} — алгебраическое дополнение элемента a_{1k} в матрице A $(k=1,\ 2,\ \ldots,\ m)$. Согласно формуле (10) из § 4 мы имеем

$$A_{1k}=(-1)^{1+k}\Delta_{1k},$$

где Δ_{1k} — минор элемента a_{1k} в матрице A, т. е определитель матрицы порядка $n-1$, полученной вычеркиванием в A первой строчки и k-го столбца. Но после такого вычеркивания мы получаем, очевидно, ступенчатую матрицу (обозначим ее через $A^{(k)}$), у которой верхняя клетка имеет порядок $m-1$. К матрице $A^{(k)}$ применимо поэтому индукционное предположение. Определитель верхней диагональной клетки матрицы $A^{(k)}$ совпадает с минором Δ'_{1k} элемента a_{1k} в матрице A_1. Что касается нижней диагональной клетки в $A^{(k)}$, то она совпадает с A_2. Следовательно, на основе индукционного предположения мы имеем

$$\Delta_{1k} = \Delta'_{1k} \cdot |A_2|. \qquad (2)$$

Обозначим через A'_{1k} алгебраическое дополнение элемента a_{1k} в матрице A_1. Так как $A'_{1k} = (-1)^{1+k} \Delta'_{1k}$, то из (2) вытекает равенство

$$A_{1k} = A'_{1k} \cdot |A_2|.$$

Возвращаясь к разложению (1), получаем теперь

$$|A| = (a_{11}A'_{11} + a_{12}A'_{12} + \ldots + a_{1m}A'_{1m}) \cdot |A_2| = |A_1| \cdot |A_2|,$$

что и требовалось доказать.

Введенное нами понятие ступенчатой матрицы легко может быть обобщено. Именно будем под ступенчатой матрицей понимать матрицу вида

$$A = \begin{Vmatrix} \boxed{A_1} & & & \text{нули} \\ & \boxed{A_2} & & \\ & & \ddots & \\ & * & & \boxed{A_k} \end{Vmatrix}, \qquad (3)$$

где $A_1, A_2, \ldots, A_k$ — квадратные матрицы, называемые *диагональными клетками* (выше клеток находятся нули, а ниже — произвольные числа). Индукцией по числу диагональных клеток k легко установить, что и в этом случае определитель ступенчатой матрицы равен произведению определителей ее диагональных клеток, т. е.

$$|A| = |A_1| \cdot |A_2| \ldots |A_k|.$$

Ввиду свойства 1 из § 4 как теорема 1, так и ее обобщение справедливы и для «перевернутых» ступенчатых матриц, т. е. для случая, когда нули находятся ниже диагональных клеток.

Доказанная нами теорема 1 является частным случаем одной более общей теоремы. Эту теорему, называемую теоремой Лапласа, мы сформулируем здесь без доказательства.

Пусть нам дана квадратная матрица n-го порядка

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

Зафиксировав некоторое натуральное число $k<n$, выберем в A какие-нибудь k строчек и k столбцов. Пусть это будут строчки с номерами $i_1<i_2<\ldots<i_k$ и столбцы с номерами $j_1<j_2<\ldots<j_k$. Из элементов на пересечениях выбранных строчек и столбцов можно образовать определитель k-го порядка

$$\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \ldots & a_{i_1j_k} \\ a_{i_2j_1} & a_{i_2j_2} & \ldots & a_{i_2j_k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i_kj_1} & a_{i_kj_2} & \ldots & a_{i_kj_k} \end{vmatrix}.$$

Этот определитель называется *минором k-го порядка* матрицы A. Вычеркнем теперь в A выбранные строчки и столбцы. Определитель полученной таким образом матрицы порядка $n-k$ называется *дополнительным минором* для первого минора. Дополнительный минор, взятый со знаком

$$(-1)^{i_1+i_2+\ldots+i_k+j_1+\ldots+j_k},$$

называется *алгебраическим дополнением* исходного минора.

При $k=1$, когда выбирается одна строчка и один столбец (с номерами, скажем, i и j), введенные понятия минора, дополнительного минора и алгебраического дополнения для данного минора превращаются соответственно в хорошо известные нам понятия элемента a_{ij} (определителя 1-го порядка), минора Δ_{ij} и алгебраического дополнения A_{ij}.

Приведем еще пример для $k=2$, $n=5$. Если в матрице A мы выберем 2-ю и 5-ю строчки и 3-й и 4-й столбцы, то соответствующим минором будет

$$\begin{vmatrix} a_{23} & a_{24} \\ a_{53} & a_{54} \end{vmatrix}.$$

Алгебраическое дополнение для этого минора равно

$$(-1)^{2+5+3+4}\begin{vmatrix} a_{11}a_{12}a_{15} \\ a_{31}a_{32}a_{35} \\ a_{41}a_{42}a_{45} \end{vmatrix}.$$

Теорема 2 (теорема Лапласа). *Выделим в матрице A n-го порядка какие-нибудь k строчек. Тогда ее определитель $|A|$ может быть представлен как сумма произведений всевозможных миноров k-го порядка, составленных из выбранных k строчек, на их алгебраические дополнения. То же самое справедливо и для случая, когда выбраны k столбцов.*

Чтобы составить какой-нибудь минор из заданных k строчек, мы должны выбрать еще k столбцов. Так как это можно сделать C_n^k способами, то сумма, о которой говорится в теореме Лапласа, состоит из C_n^k слагаемых.

Для определителя 4-го порядка, в котором выбраны две первые строчки, теорема Лапласа дает следующее разложение:

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ x & y & z & t \\ \alpha & \beta & \gamma & \delta \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ a & b \end{vmatrix} \cdot (-1)^{1+2+1+2} \begin{vmatrix} z & t \\ \gamma & \delta \end{vmatrix} +$$

$$+ \begin{vmatrix} 1 & 1 \\ a & c \end{vmatrix} \cdot (-1)^{1+2+1+3} \begin{vmatrix} y & t \\ \beta & \delta \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ a & d \end{vmatrix} \cdot (-1)^{1+2+1+4} \begin{vmatrix} y & z \\ \beta & \gamma \end{vmatrix} +$$

$$+ \begin{vmatrix} 1 & 1 \\ b & c \end{vmatrix} \cdot (-1)^{1+2+2+3} \begin{vmatrix} x & t \\ \alpha & \delta \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ b & d \end{vmatrix} \cdot (-1)^{1+2+2+4} \begin{vmatrix} x & z \\ \alpha & \gamma \end{vmatrix} +$$

$$+ \begin{vmatrix} 1 & 1 \\ c & d \end{vmatrix} \cdot (-1)^{1+2+3+4} \begin{vmatrix} x & y \\ \alpha & \beta \end{vmatrix}.$$

Покажем теперь, как с помощью теоремы Лапласа может быть получена теорема 1. Выберем в ступенчатой матрице A, выписанной в начале параграфа, первые m строчек. Одним из миноров, построенных на этих строчках, будет определитель верхней диагональной клетки $|A_1|$ (если взять первые m столбцов). Всякий другой минор из выбранных строчек обязательно захватит столбец из нулей и поэтому будет равен нулю. Следовательно, разложение Лапласа в этом случае сводится к одному слагаемому (все остальные равны нулю). Дополнительный минор для $|A_1|$ равен $|A_2|$, поэтому

$$|A| = |A_1| \cdot (-1)^{1+2+\ldots+m+1+2+\ldots+m} |A_2| = |A_1| \cdot |A_2|.$$

Упражнения

1. Вычислить определители

a) $\begin{vmatrix} 2 & 3 & 0 & 0 & 0 \\ 4 & 5 & 0 & 0 & 0 \\ 1 & 7 & 1 & 1 & 1 \\ -1 & 3 & 2 & 3 & 4 \\ 9 & 8 & 4 & 9 & 16 \end{vmatrix}$; b) $\begin{vmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 & 0 \\ -2 & 3 & 2 & -1 & 0 & 0 \\ 7 & 2 & 3 & 2 & 0 & 0 \\ 5 & -1 & 3 & 5 & 7 & -5 \\ 2 & 3 & 7 & 2 & 2 & -1 \end{vmatrix}$.

Отв. a) -4; b) 42.

2. Вычислить определитель, применив теорему Лапласа ко второй и четвертой строчкам,

$$\begin{vmatrix} 1 & 2 & 3 & 1 & 5 \\ 0 & 1 & 0 & 5 & 1 \\ 2 & 1 & 2 & 3 & 2 \\ 0 & 3 & 0 & 1 & 3 \\ 3 & 2 & 1 & 3 & 4 \end{vmatrix}.$$

Отв. 168.

ГЛАВА II

СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 7. Основные понятия

Система m линейных уравнений с n неизвестными имеет вид

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{array}\right\} \qquad (1)$$

В общем случае число уравнений в системе не обязательно совпадает с числом неизвестных: m может быть меньше, равно или больше числа n.

Числа a_{ik} (вещественные или комплексные) называются *коэффициентами системы*; b_i — *свободными членами*; $x_1, x_2, \ldots, x_n$ — *неизвестными*.

Определение. *Совокупность n чисел $\alpha_1, \alpha_2, \ldots$ $\ldots, \alpha_n$ называется решением системы* (1), *если после замены неизвестных $x_1, x_2, \ldots, x_n$ числами $\alpha_1, \alpha_2, \ldots, \alpha_n$ соответственно каждое из уравнений системы превращается в верное равенство.*

Пример 1.

$$\left.\begin{array}{l} x + y + z = 1, \\ 2x + 2y + 2z = 3. \end{array}\right\}$$

Эта система двух уравнений с тремя неизвестными решений не имеет, так как любая тройка чисел, удовлетворяющая первому уравнению, не может удовлетворять второму.

Пример 2.

$$\left.\begin{array}{l} x + y = 1, \\ 2x + 7y = -3. \end{array}\right\}$$

Легко видеть, что эта система имеет единственное решение: $x=2$, $y=-1$.

Пример 3.

$$\left.\begin{aligned} x+y&=1,\\ 2x+2y&=2,\\ 3x+3y&=3. \end{aligned}\right\}$$

Пара чисел, 1, 0 есть одно из решений этой системы трех уравнений с двумя неизвестными; $-1, 2$ — другое решение. Эта система имеет бесконечно много решений: значения $x=\alpha$, $y=1-\alpha$ при любом α удовлетворяют системе.

Приведенные примеры систем показывают, что, вообще говоря, система может либо вовсе не иметь решений, либо иметь единственное решение, либо иметь их несколько (как мы увидим в дальнейшем, в последнем случае система всегда имеет бесконечно много решений).

Определение. *Система, не имеющая ни одного решения, называется несовместной. Система, обладающая хотя бы одним решением, называется совместной.*

Система примера 1 несовместна. Системы примеров 2 и 3 совместны.

Относительно каждой системы линейных уравнений могут быть поставлены следующие вопросы.

1. Совместна заданная система или нет?

2. В случае, если система совместна, как определить, сколько она имеет решений — одно или несколько?

3. Как найти все решения системы?

Дать ответ на все эти вопросы и должна теория линейных уравнений.

§ 8. Теорема Крамера

Ограничимся сначала рассмотрением систем, у которых число уравнений равно числу неизвестных.

Пусть мы имеем систему n линейных уравнений с n неизвестными:

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\ldots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n&=b_n. \end{aligned}\right\} \qquad (1)$$

Определитель

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

составленный из коэффициентов при неизвестных, называется *определителем системы* (1). Введем дополнительное ограничение, предположив, что определитель Δ этой системы не равен нулю. При этом предположении в § 1 нами были получены формулы для решений систем двух линейных уравнений с двумя неизвестными и трех линейных уравнений с тремя неизвестными. Сейчас мы получим аналогичные формулы для произвольного n.

Теорема Крамера. *Если определитель Δ системы n линейных уравнений с n неизвестными отличен от нуля, то эта система имеет единственное решение. Это решение может быть найдено по формулам*

$$x_1=\frac{\Delta_1}{\Delta},\quad x_2=\frac{\Delta_2}{\Delta},\ \dots,\ x_n=\frac{\Delta_n}{\Delta}, \tag{2}$$

где Δ_k — определитель, получающийся из Δ заменой k-го столбца свободными членами системы

$$\Delta_k=\begin{vmatrix} a_{11} \dots a_{1,k-1} & b_1 & a_{1,k+1} \dots a_{1n} \\ a_{21} \dots a_{2,k-1} & b_2 & a_{2,k+1} \dots a_{2n} \\ \cdot\ \cdot\ \cdot\ \cdot & \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1} \dots a_{n,k-1} & b_n & a_{n,k+1} \dots a_{nn} \end{vmatrix}. \tag{3}$$

Формулы (2) носят название *формул Крамера.*

Доказательство. 1°. Предположив сначала, что решение существует, докажем, что оно единственное.

Пусть числа $\alpha_1, \alpha_2, \dots, \alpha_n$ образуют решение системы (1), определитель которой Δ не равен нулю. Тогда мы имеем n верных равенств:

$$\left.\begin{aligned} a_{11}\alpha_1+a_{12}\alpha_2+\dots+a_{1n}\alpha_n&=b_1,\\ a_{21}\alpha_1+a_{22}\alpha_2+\dots+a_{2n}\alpha_n&=b_2,\\ \dots\dots\dots\dots\dots&\dots\\ a_{n1}\alpha_1+a_{n2}\alpha_2+\dots+a_{nn}\alpha_n&=b_n. \end{aligned}\right\} \tag{4}$$

Умножим левую и правую части каждого из равенств (4) последовательно на алгебраические дополнения A_{11}, $A_{21}, \ldots, A_{n1}$ элементов первого столбца определителя Δ и сложим все полученные левые и правые части. В результате получим

$$\begin{aligned}(a_{11}A_{11}+a_{21}A_{21}+\ldots+a_{n1}A_{n1})\,\alpha_1+\\ +(a_{12}A_{11}+a_{22}A_{21}+\ldots+a_{n2}A_{n1})\,\alpha_2+\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ +(a_{1n}A_{11}+a_{2n}A_{21}+\ldots+a_{nn}A_{n1})\,\alpha_n=\\ =b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1}.\end{aligned}$$

Коэффициент, стоящий здесь при α_1, равен Δ, как сумма произведений элементов первого столбца определителя Δ на их алгебраические дополнения (свойство 4 из § 4). Коэффициенты же при $\alpha_2, \alpha_3, \ldots, \alpha_n$ все равны нулю, как суммы произведений элементов 2-го, 3-го, ..., n-го столбцов определителя Δ на алгебраические дополнения того же первого столбца (свойство 5 из § 4 для столбцов). Следовательно, мы имеем

$$\Delta\alpha_1=b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1}.$$

Совершенно таким же образом (если равенства (4) умножим последовательно на алгебраические дополнения $A_{1k}, A_{2k}, \ldots, A_{nk}$ элементов k-го столбца определителя Δ и результаты сложим) получим

$$\Delta\alpha_k=b_1A_{1k}+b_2A_{2k}+\ldots+b_nA_{nk}. \qquad (5)$$

Разложим определитель (3) по элементам k-го столбца. Так как матрицы определителей Δ и Δ_k отличаются лишь элементами k-го столбца, то, по лемме 3 из § 4 (примененной для столбцов), алгебраические дополнения элементов $b_1, b_2, \ldots, b_n$ в определителе Δ_k соответственно равны $A_{1k}, A_{2k}, \ldots, A_{nk}$. Следовательно,

$$\Delta_k=b_1A_{1k}+b_2A_{2k}+\ldots+b_nA_{nk}.$$

Равенство (5) принимает теперь вид

$$\Delta\alpha_k=\Delta_k \qquad (k=1, 2, \ldots, n).$$

Вспоминая, что $\Delta\neq 0$, окончательно находим

$$\alpha_1=\frac{\Delta_1}{\Delta}, \quad \alpha_2=\frac{\Delta_2}{\Delta}, \ \ldots, \ \alpha_n=\frac{\Delta_n}{\Delta}. \qquad (6)$$

Если теперь допустить, что система (1) имеет еще решение, скажем $\beta_1, \beta_2, \dots, \beta_n$, то, повторив приведенное рассуждение для этого решения, мы получим для $\beta_1, \beta_2, \dots, \beta_n$ формулы, в правых частях которых будут стоять те же выражения, что и в формулах (6). Но тогда $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \dots, \alpha_n = \beta_n$.

Нами доказано, таким образом, что если (при $\Delta \neq 0$) система (1) имеет решение, то это решение единственно, и оно может быть найдено по формулам (2).

2°. Для доказательства существования решения нам достаточно проверить, что совокупность чисел (6) образует решение системы (1). Заметим при этом, что выражения (6) имеют смысл, так как по условию $\Delta \neq 0$.

Подставив числа (6) вместо неизвестных $x_1, x_2, \dots, x_n$ в левую часть i-го уравнения системы (1), мы получим

$$
\begin{aligned}
a_{i1}\frac{\Delta_1}{\Delta} + a_{i2}\frac{\Delta_2}{\Delta} + \dots + a_{in}\frac{\Delta_n}{\Delta} = \\
&= \frac{a_{i1}}{\Delta}(b_1A_{11} + \dots + b_iA_{i1} + \dots + b_nA_{n1}) + \\
&+ \frac{a_{i2}}{\Delta}(b_1A_{12} + \dots + b_iA_{i2} + \dots + b_nA_{n2}) + \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ \frac{a_{in}}{\Delta}(b_1A_{1n} + \dots + b_iA_{in} + \dots + b_nA_{nn}) = \\
&= \frac{b_1}{\Delta}(a_{i1}A_{11} + a_{i2}A_{12} + \dots + a_{in}A_{1n}) + \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ \frac{b_i}{\Delta}(a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}) + \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&+ \frac{b_n}{\Delta}(a_{i1}A_{n1} + a_{i2}A_{n2} + \dots + a_{in}A_{nn}) = \\
&\qquad = \frac{b_1}{\Delta}\cdot 0 + \dots + \frac{b_i}{\Delta}\cdot\Delta + \dots + \frac{b_n}{\Delta}\cdot 0 = b_i.
\end{aligned}
$$

При выкладках мы воспользовались тем, что

$$
a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = \begin{cases} \Delta & \text{при } i = j, \\ 0 & \text{при } i \neq j \end{cases}
$$

(свойства 4 и 5 из § 4 для строчек).

Таким образом, при замене неизвестных $x_1, x_2, \ldots, x_n$ числами (6) каждое уравнение системы (1) превращается в верное равенство (тождество), а это и значит, что эти числа образуют решение. Теорема Крамера доказана полностью.

Пример. Решить систему линейных уравнений

$$\left.\begin{aligned} x_1 - 2x_2 + 3x_3 - x_4 &= 6,\\ 2x_1 + 3x_2 - 4x_3 + 4x_4 &= -7,\\ 3x_1 + x_2 - 2x_3 - 2x_4 &= 9,\\ x_1 - 3x_2 + 7x_3 + 6x_4 &= -7. \end{aligned}\right\}$$

Вычисляем определитель системы:

$$\Delta = \begin{vmatrix} 1 & -2 & 3 & -1\\ 2 & 3 & -4 & 4\\ 3 & 1 & -2 & -2\\ 1 & -3 & 7 & 6 \end{vmatrix} = 35.$$

Так как $\Delta \neq 0$, то система имеет единственное решение, и для его нахождения мы можем применить формулы Крамера (2). Вычисляем определители:

$$\Delta_1 = \begin{vmatrix} 6 & -2 & 3 & -1\\ -7 & 3 & -4 & 4\\ 9 & 1 & -2 & -2\\ -7 & -3 & 7 & 6 \end{vmatrix} = 70,$$

$$\Delta_2 = \begin{vmatrix} 1 & 6 & 3 & -1\\ 2 & -7 & -4 & 4\\ 3 & 9 & -2 & -2\\ 1 & -7 & 7 & 6 \end{vmatrix} = -35,$$

$$\Delta_3 = \begin{vmatrix} 1 & -2 & 6 & -1\\ 2 & 3 & -7 & 4\\ 3 & 1 & 9 & -2\\ 1 & -3 & -7 & 6 \end{vmatrix} = 0,$$

$$\Delta_4 = \begin{vmatrix} 1 & -2 & 3 & 6\\ 2 & 3 & -4 & -7\\ 3 & 1 & -2 & 9\\ 1 & -3 & 7 & -7 \end{vmatrix} = -70.$$

Следовательно,

$$x_1=\frac{\Delta_1}{\Delta}=2, \qquad x_2=\frac{\Delta_2}{\Delta}=-1,$$

$$x_3=\frac{\Delta_3}{\Delta}=0, \qquad x_4=\frac{\Delta_4}{\Delta}=-2.$$

Замечание. В случае, когда число неизвестных n велико, практическое использование формул Крамера затруднено в связи с необходимостью большого числа вычислений. (Требуется вычислить $n+1$ определителей порядка n; в § 10 нами изложен другой, более удобный для практических целей способ решения систем линейных уравнений с числовыми коэффициентами.) Однако формулы Крамера имеют весьма важное значение в теоретическом плане, поскольку они дают возможность явно записать значения для неизвестных через коэффициенты системы.

Упражнения

1. Решить систему с помощью формул Крамера:

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

Отв. $x_1=2,\ x_2=1,\ x_3=-3,\ x_4=1.$

2. Решить систему линейных уравнений:

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

§ 9. Ранг матрицы. Элементарные преобразования

В предыдущем параграфе нами исследованы системы линейных уравнений, у которых число неизвестных совпадает с числом уравнений и определитель которых не равен нулю. Теперь мы должны перейти к исследованию и решению произвольных систем линейных уравнений. Однако предварительно нам надо будет познакомиться с некоторыми сведениями, относящимися к теории прямоугольных матриц. Это отступление обусловлено тем, что произвольная система линейных уравнений (см. (1) из § 7) с точностью до обозначения неизвестных вполне определяется таблицей коэффициентов при неизвестных и свободными членами, и поэтому свойства системы должны проявляться в свойствах соответствующей матрицы.

Итак, пусть нам дана прямоугольная матрица (т. е. таблица mn чисел, расположенных в m строчках и n столбцах)

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}. \qquad (1)$$

Конечно, в частном случае допускается равенство $m = n$, т. е. матрица A может быть квадратной.

Пусть k — какое-нибудь натуральное число, не превосходящее m и n. Выберем в A произвольным способом k строчек с номерами $i_1 < i_2 < \dots < i_k$ и k столбцов с номерами $j_1 < j_2 < \dots < j_k$. Из элементов матрицы A, лежащих на пересечении выбранных k строчек и k столбцов, можно образовать определитель

$$\begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i_k j_1} & a_{i_k j_2} & \dots & a_{i_k j_k} \end{vmatrix}.$$

Этот определитель называется *минором k-го порядка* матрицы A.

С понятием минора матрицы мы уже встречались в § 6, но там мы рассматривали лишь случай квадратной матрицы. Здесь же мы ввели понятие минора для произвольных прямоугольных матриц. При $k = 1$ под минором 1-го порядка мы понимаем, разумеется, элемент a_{ij} матрицы A.

Из m строчек мы можем C_m^k различными способами выбрать k строчек. Аналогично из n столбцов можно выбрать C_n^k различными способами k столбцов. Следовательно, в матрице с m строчками и n столбцами мы имеем $C_m^k \cdot C_n^k$ различным образом составленных миноров k-го порядка.

Пример. Для матрицы

$$A=\begin{Vmatrix} 1 & 2 & 3 & 4 \\ a & b & c & d \\ x & y & z & t \end{Vmatrix}$$

определители

$$\begin{vmatrix} b, & d \\ y, & t \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 4 \\ a & b & d \\ x & y & t \end{vmatrix}$$

являются минорами 2-го и 3-го порядка.

Предложение 1. *Если в матрице A все миноры k-го порядка равны нулю, то равны нулю и все миноры более высокого порядка (если таковые существуют).*

Доказательство. Возьмем в матрице A произвольный минор $(k+1)$-го порядка. Он равен сумме произведений элементов какой-нибудь строчки на их алгебраические дополнения. Но алгебраическое дополнение любого элемента нашего минора с точностью до знака (см. формулу (10) из § 4) равно некоторому минору k-го порядка матрицы A. А так как все миноры k-го порядка матрицы A по условию равны нулю, то равны нулю рассматриваемые алгебраические дополнения и, далее, равен нулю выбранный минор $(k+1)$-го порядка. Итак, все миноры порядка $k+1$ матрицы A равны нулю. Аналогичным образом докажем, что все миноры порядка $k+2$ матрицы A также равны нулю и т. д. Предложение 1 доказано.

В теории систем линейных уравнений существенную роль играет понятие ранга матрицы. Именно в терминах ранга в § 11 и будет сформулировано необходимое и достаточное условие для совместности произвольной системы линейных уравнений.

Определение. *Рангом матрицы A называется такое целое число r, что среди миноров r-го порядка матрицы A имеется хоть один, не равный нулю, а все миноры $(r+1)$-го порядка (если только их можно составить) сплошь равны нулю.*

Если все элементы матрицы равны нулю, то ранг такой матрицы считается равным нулю (здесь равны нулю миноры всех порядков).

Пусть ранг матрицы A равен r. Тогда, согласно предложению 1 (которое надо применить для $k=r+1$), в матрице A равны нулю все миноры порядка выше, чем r. Следовательно, ранг матрицы может быть определен как наивысший из порядков не равных нулю миноров этой матрицы. Если поэтому о некоторой матрице известно, что все ее миноры k-го порядка равны нулю, то ранг этой матрицы меньше k.

Примеры. Найдем ранги следующих матриц:

$$A_1=\begin{Vmatrix}1 & 1 & 1\\ 2 & 2 & 2\\ 3 & 3 & 3\end{Vmatrix},\quad B_1=\begin{Vmatrix}1 & 0\\ 0 & 0\\ 0 & 0\end{Vmatrix},$$

$$A_2=\begin{Vmatrix}1 & 2 & 3 & 4\\ 0 & 5 & 6 & 7\\ 0 & 0 & 0 & 0\end{Vmatrix},\quad B_2=\begin{Vmatrix}1 & 1 & 3 & 1\\ 2 & 1 & 4 & 3\\ 1 & 2 & 5 & 0\\ 5 & 4 & 13 & 6\end{Vmatrix}.$$

$$A_3=\begin{Vmatrix}1 & 1 & 1\\ 1 & 2 & 3\\ 1 & 4 & 9\end{Vmatrix}.$$

Ранг матриц A_1 и B_1 равен 1, так как все миноры 2-го порядка этих матриц равны нулю, а среди миноров 1-го порядка, т. е. среди элементов этих матриц, есть не равные нулю. Ранг матриц A_2 и B_2 равен 2. Действительно,

$$\begin{vmatrix}1 & 2\\ 0 & 5\end{vmatrix}\neq 0 \text{ и } \begin{vmatrix}1 & 1\\ 2 & 1\end{vmatrix}\neq 0$$

(миноры 2-го порядка матриц A_2 и B_2 соответственно), а все миноры 3-го порядка обеих матриц равны нулю (для A_2 — четыре минора, для B_2 — шестнадцать миноров 3-го порядка).

Ранг матрицы A_3 равен 3, так как единственный минор 3-го порядка этой матрицы не равен нулю.

На примере матрицы B_2 мы видим, что задача определения ранга матрицы (если исходить только из определения), вообще говоря, требует вычисления большого числа определителей. Существует, однако, прием, позволяющий свести объем вычислений к минимуму. Прием этот основан на использовании элементарных преобразований матриц.

Определение. *Под элементарными преобразованиями матрицы понимаются следующие операции:*

1°) *умножение какой-либо строчки матрицы на число, отличное от нуля;*

2°) *прибавление к элементам одной строчки матрицы соответствующих элементов другой строчки, умноженных на одно и то же число;*

3°) *перемена местами двух строчек;*

1*, 2*, 3* — *аналогичные операции над столбцами.*

Применяя к матрице A какое-либо элементарное преобразование, мы получаем новую матрицу B. Этот факт мы будем записывать так: $A \to B$.

Предложение 2. *Элементарные преобразования обратимы, т. е. если матрица B получается из A при помощи какого-либо элементарного преобразования, то и матрица A может быть получена из B также при помощи некоторого элементарного преобразования (называемого обратным к первому).*

Доказательство. Пусть матрица B получается из A умножением i-й строчки на число $c \neq 0$. Умножая i-ю строчку матрицы B на c^{-1} (т. е. применяя к B элементарное преобразование), мы получим исходную матрицу A. Пусть B получается из A прибавлением к i-й строчке элементов j-й строчки, умноженных на число α. Прибавляя к элементам i-й строчки матрицы B элементы ее j-й строчки, умноженные на $-\alpha$, мы возвращаемся к матрице A. Наконец, если B получается из A перестановкой i-й и j-й строчек, то, переставляя в B те же i-ю и j-ю строчки, мы снова получим исходную матрицу A. Совершенно аналогичным образом проверяется обратимость элементарных преобразований над столбцами.

Предложение 3. *Элементарное преобразование третьего типа (т. е. 3° или 3*) равносильно нескольким последовательно выполненным преобразованиям первых двух типов.*

Доказательство. Пусть мы имеем элементарное преобразование

$$A = \begin{Vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & b_1 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_2 & b_2 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix} \to \begin{Vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ a_2 & b_2 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & b_1 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix} = B$$

типа 3°, при котором отмеченные строчки поменялись местами (при сохранении всех остальных без изменения). Переход от A к B при помощи преобразований типа 1° и 2° может быть осуществлен следующим образом:

$$A=\begin{Vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & b_1 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_2 & b_2 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix} \xrightarrow{2^\circ} \begin{Vmatrix} \cdot & \cdot & \cdot \\ a_1+a_2 & b_1+b_2 & \ldots \\ \cdot & \cdot & \cdot \\ a_2 & b_2 & \\ \cdot & \cdot & \cdot \end{Vmatrix} \xrightarrow{2^\circ}$$

$$\xrightarrow{2^\circ} \begin{Vmatrix} \cdot & \cdot & \cdot \\ a_1+a_2 & b_1+b_2 & \ldots \\ \cdot & \cdot & \cdot \\ -a_1 & -b_1 & \ldots \\ \cdot & \cdot & \cdot \end{Vmatrix} \xrightarrow{2^\circ} \begin{Vmatrix} \cdot & \cdot & \cdot \\ a_2 & b_2 & \ldots \\ \cdot & \cdot & \cdot \\ -a_1 & -b_1 & \ldots \\ \cdot & \cdot & \cdot \end{Vmatrix} \xrightarrow{1^\circ}$$

$$\xrightarrow{1^\circ} \begin{Vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ a_2 & b_2 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1 & b_1 & \ldots & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{Vmatrix} = B.$$

Здесь мы к верхней из выписанных строчек прибавили нижнюю, полученную верхнюю строчку вычли из нижней, затем снова к верхней строчке прибавили нижнюю и, наконец, нижнюю строчку умножили на $c=-1$. В результате этих операций мы от A пришли к B, т. е. получили тот же самый эффект, что и при элементарном преобразовании 3°. Это и составляет содержание предложения 3.

В дальнейшем ранг матрицы A мы будем обозначать через r_A.

Теорема 1. *При элементарном преобразовании ранг матрицы не меняется. Другими словами, если* $A \to B$, *то* $r_A = r_B$.

Доказательство. 1) Предположим, что B получается из A умножением i-й строчки на $c \neq 0$. Рассмотрим в матрицах A и B одинаковым образом составленные миноры Δ и δ (некоторого порядка k). Матрицы A и B отличаются только элементами i-й строчки. Поэтому $\delta=\Delta$, если i-я строчка не входит в состав миноров, и $\delta=c\Delta$, если входит. Отсюда следует (поскольку $c\neq 0$), что δ и Δ либо одновременно отличны от нуля, либо одновременно равны нулю. Но в таком случае наибольший порядок не равного нулю минора для обеих матриц будет один и тот же, т. е. $r_A = r_B$.

2) Пусть B получается из A прибавлением к i-й строчке элементов j-й строчки, умноженных на α. Положим $r_A = r$ и докажем, что в матрице B все миноры порядка $r+1$ равны нулю.

Пусть δ — произвольный минор $(r+1)$-го порядка матрицы B. Если i-я строчка матрицы B не входит в δ, то δ является также минором $(r+1)$-го порядка матрицы A, а потому $\delta = 0$ (так как ранг матрицы A равен r, то все ее миноры порядка $r+1$ равны нулю).

Пусть в δ входят обе строчки — i-я и j-я:

$$\delta = \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{ik_1} + \alpha a_{jk_1} & a_{ik_2} + \alpha a_{jk_2} & \ldots \\ \cdots & \cdots & \cdots \\ a_{jk_1} & a_{jk_2} & \\ \cdots & \cdots & \cdots \end{vmatrix}.$$

В силу свойства 9 из § 4 определитель δ равен определителю

$$\begin{vmatrix} \cdots & \cdots & \cdots \\ a_{ik_1} & a_{ik_2} & \ldots \\ \cdots & \cdots & \cdots \\ a_{jk_1} & a_{jk_2} & \ldots \\ \cdots & \cdots & \cdots \end{vmatrix},$$

т. е. равен определителю порядка $r+1$ матрицы A. Так что и в этом случае $\delta = 0$.

Осталось рассмотреть случай, когда i-я строчка входит в δ, а j-я не входит. В этом случае, применяя свойства 8 и 6 из § 4, мы имеем

$$\delta = \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{ik_1} + \alpha a_{jk_1} & a_{ik_2} + \alpha a_{jk_1} & \ldots \\ \cdots & \cdots & \cdots \end{vmatrix} =$$

$$= \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{ik_1} & a_{ik_2} & \ldots \\ \cdots & \cdots & \cdots \end{vmatrix} + \alpha \begin{vmatrix} \cdots & \cdots & \cdots \\ a_{jk_1} & a_{jk_2} & \ldots \\ \cdots & \cdots & \cdots \end{vmatrix} = \Delta + \alpha\Delta_1.$$

Здесь Δ есть минор порядка $r+1$ матрицы A, поэтому $\Delta = 0$. Что же касается определителя Δ_1, то он, строго говоря, не является минором матрицы A, так как в нем выписанная строчка может оказаться не на своем месте. Но при помощи транспозиций с соседними строчками мы можем эту строчку передвинуть в надлежащее место так, чтобы взаимное расположение строчек в Δ_1 соответствовало их расположению в матрице A. Этим мы преобразуем определитель Δ_1 в минор Δ' (порядка $r+1$) матрицы A. Из равенств $\Delta'=0$ и $\Delta_1=\pm\Delta'$ следует теперь, что $\Delta_1=0$, а значит, равен нулю и минор $\delta=\Delta+\alpha\Delta_1$.

Объединяя вместе разобранные случаи, мы можем констатировать, что если $A \xrightarrow{2^\circ} B$ и $r_A=r$, то все миноры $(r+1)$-го порядка матрицы B равны нулю и, следовательно, $r_B\leqslant r$, т. е. $r_B\leqslant r_A$. Теперь, в силу обратимости элементарных преобразований (предложение 2), в нашем рассуждении матрицы A и B можно поменять местами. В результате мы получим также неравенство $r_A\leqslant r_B$, которое вместе с предшествующим дает нам равенство $r_A=r_B$. Этим мы завершили доказательство теоремы 1 для элементарного преобразования типа 2°.

3) Справедливость теоремы 1 для элементарного преобразования 3° следует, ввиду доказанного, из предложения 3.

Для столбцов доказательство проводится, конечно, аналогично.

Доказанная теорема 1 может быть с пользой применена к вычислению ранга (для матриц с конкретными числовыми элементами), а именно, если при помощи нескольких последовательно выполненных элементарных преобразований мы перешли от матрицы A к некоторой другой матрице C, то, согласно теореме 1, $r_C=r_A$. Вычислив ранг r_C, мы тем самым будем знать и ранг r_A. Оказывается, что, отправляясь от любой матрицы A, всегда можно прийти к такой матрице C, вычисление ранга которой не представляет затруднений. Для этого следует добиться, чтобы в матрице C было достаточно много нулей.

Пример. Найти ранг матрицы

$$\begin{Vmatrix} 1 & 3 & 2 & 0 & 5 \\ 2 & 6 & 9 & 7 & 12 \\ -2 & -5 & 2 & 4 & 5 \\ 1 & 4 & 8 & 4 & 20 \end{Vmatrix}.$$

Подвергнем эту матрицу следующим преобразованиям. Ко второй строчке прибавим первую, умноженную на —2, затем к третьей строчке прибавим первую, умноженную на 2, наконец, из четвертой вычтем первую. После этих трех последовательно выполненных элементарных преобразований получим матрицу

$$\begin{Vmatrix} 1 & 3 & 2 & 0 & 5 \\ 0 & 0 & 5 & 7 & 2 \\ 0 & 1 & 6 & 4 & 15 \\ 0 & 1 & 6 & 4 & 15 \end{Vmatrix}.$$

Вычитая из четвертой строчки третью, а затем переставив местами вторую и третью строчки, получаем

$$\begin{Vmatrix} 1 & 3 & 2 & 0 & 5 \\ 0 & 1 & 6 & 4 & 15 \\ 0 & 0 & 5 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{Vmatrix}.$$

Ранг полученной матрицы равен 3, так как минор 3-го порядка, лежащий в левом верхнем углу, не равен нулю, а все миноры 4-го порядка равны нулю. Следовательно, ранг исходной матрицы также равен 3.

Если мы в последней матрице третью строчку разделим на 5 (т. е. выполним еще одно элементарное преобразование типа 1°), то получим матрицу

$$\begin{Vmatrix} 1 & 3 & 2 & 0 & 5 \\ 0 & 1 & 6 & 4 & 15 \\ 0 & 0 & 1 & \frac{7}{5} & \frac{2}{5} \\ 0 & 0 & 0 & 0 & 0 \end{Vmatrix}$$

с единицами вдоль «главной диагонали». Матрицу такого вида мы будем называть *трапециевидной*. Трапециевидная матрица в общем случае выглядит следующим образом:

$$\begin{Vmatrix} 1 & c_{12} & c_{13} & \dots & \dots & c_{1n} \\ 0 & 1 & c_{23} & \dots & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & c_{r,\,r+1} & \dots\, c_{rn} \\ 0 & 0 & \dots & 0 & 0 & \dots\, 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots\, 0 \end{Vmatrix}.$$

Частным случаем трапециевидной матрицы является треугольная матрица вида

$$\begin{Vmatrix} 1 & a & b & c \\ 0 & 1 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Другими примерами трапециевидных матриц являются

$$\begin{Vmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 1 & a & b & c \\ 0 & 1 & d & e \end{Vmatrix}.$$

В рассмотренном примере мы привели матрицу к трапециевидной форме при помощи элементарных преобразований только над строчками матрицы. В общем случае для перехода к трапециевидной матрице приходится помимо преобразований 1°, 2° и 3° применять еще и преобразование 3*.

Теорема 2. *Всякая матрица при помощи элементарных преобразований над строчками и при помощи перестановок столбцов может быть преобразована в трапециевидную матрицу.*

Доказательство. Пусть нам дана произвольная прямоугольная матрица (1) (см. стр. 73). Если все ее элементы равны нулю, то она сама является трапециевидной (число единиц на главной «диагонали» равно нулю). Допустим, что не все элементы матрицы (1) нули. Тогда, переставляя строчки и столбцы (т. е. применяя преобразования 3° и 3*), мы можем не равный нулю элемент поместить в левый верхний угол. Можно считать теперь, что $a_{11} \neq 0$. Разделив первую строчку на a_{11} (преобразование 1°), мы получим матрицу, у которой $a_{11} = 1$. Вычтем теперь первую строчку, умноженную соответственно на $a_{21}, a_{31}, \ldots, a_{m1}$, из всех последующих строчек (преобразования 2°). В результате мы придем к матрице

$$\left\| \begin{matrix} 1 & c_{12} & c_{13} & \ldots & c_{1n} \\ 0 & a'_{22} & a'_{23} & \ldots & a'_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & a'_{m2} & a'_{m3} & \ldots & a'_{mn} \end{matrix} \right\|.$$

Если окажется, что все элементы a'_{ij} равны нулю, то мы уже достигли цели: полученная матрица трапециевидная. Если же не все a'_{ij} нули, то при помощи преобразований 3° и 3* добиваемся того, чтобы $a'_{22} \neq 0$, а затем при помощи преобразований 1° и 2° приходим к матрице

$$\left\| \begin{matrix} 1 & c_{12} & c_{13} & \ldots & c_{1n} \\ 0 & 1 & c_{23} & \ldots & c_{2n} \\ 0 & 0 & a''_{33} & \ldots & a''_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & a''_{m3} & \ldots & a''_{mn} \end{matrix} \right\|.$$

Если здесь все a''_{ij} равны нулю, то все готово. Если же не все a''_{ij} нули, то продолжаем наш процесс дальше. В результате на некотором шаге мы либо исчерпаем все строчки (число единиц на главной «диагонали» будет равно m), либо придем к матрице, у которой несколько последних строчек состоит сплошь из нулей. В том и другом случае мы получим трапециевидную матрицу. Теорема 2 доказана.

В заключение параграфа сделаем следующие два замечания.

Замечание 1. Если к матрице A ранга r приписать строчку или столбец из нулей, то ранг полученной матрицы будет также равен r.

Действительно, не равный нулю минор r-го порядка матрицы A будет также минором r-го порядка и матрицы B, и, с другой стороны, всякий минор матрицы B порядка $r+1$ равен нулю, поскольку он либо является минором $(r+1)$-го порядка матрицы A, либо содержит строчку (или столбец) из нулей.

Замечание 2. Если к матрице A приписать какую-нибудь строчку (или столбец), то ранг полученной матрицы B может превосходить ранг исходной матрицы не более чем на одну единицу.

Действительно, если $r_A = r$, то, очевидно $r_B \geqslant r$. Возьмем произвольный минор δ порядка $r+2$ матрицы B. Если в δ приписанная строчка (или столбец) не входит, то δ будет также минором порядка $r+2$ матрицы A и поэтому $\delta = 0$. Если в δ приписанные элементы входят, то, разлагая δ по этим элементам, мы выразим наш минор через определители порядка $r+1$, которые будут уже минорами $(r+1)$-го порядка матрицы A. Так что и в этом случае $\delta = 0$. Таким образом, все миноры порядка $r+2$ матрицы B равны нулю, а значит $r_B \leqslant r+1$.

Упражнения

1. Вычислить ранги матриц

a) $\begin{Vmatrix} 3 & 4 & -2 & 1 & 5 \\ 4 & -3 & 5 & 2 & -3 \\ 5 & 15 & -11 & 1 & 18 \\ 7 & -24 & 26 & 5 & -27 \end{Vmatrix}$; b) $\begin{Vmatrix} 2 & 3 & -1 & 4 & 1 \\ 3 & 2 & 5 & -3 & 4 \\ 4 & -3 & 2 & 1 & 7 \\ 1 & 8 & 2 & 0 & -2 \end{Vmatrix}$.

Отв. a) 2; b) 3.

2. Показать, что всякая матрица ранга r при помощи элементарных преобразований может быть преобразована в матрицу, у которой элементы $a_{11}, a_{22}, \ldots, a_{rr}$ вдоль главной «диагонали» равны 1, а все прочие элементы равны 0.

§ 10. Метод исключения неизвестных

Мы изложим здесь способ решения систем линейных уравнений, который носит название метода последовательного исключения неизвестных или метода Гаусса.

Пусть нам дана произвольная система m линейных уравнений с n неизвестными:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots&\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned}\right\} \quad (1)$$

Матрица

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

составленная из коэффициентов при неизвестных, называется *матрицей системы* (1). Матрица

$$B = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{Vmatrix},$$

получающаяся из A добавлением столбца из свободных членов системы (1), называется *расширенной матрицей системы* (1). Матрица B, очевидно, вполне определяет систему (1) с точностью до обозначения неизвестных.

Определение. *Под элементарными преобразованиями системы линейных уравнений понимаются следующие операции*:

1°) *умножение какого-либо уравнения системы на число, отличное от нуля*;

2°) *прибавление к одному уравнению другого уравнения, умноженного на произвольное число*;

3°) *перемена местами двух уравнений в системе.*

Делая элементарное преобразование в системе, мы получаем новую систему. Очевидно, что каждому элементарному преобразованию системы соответствует аналогичное преобразование над строчками расширенной матрицы этой системы, и наоборот, каждому элементарному преобразованию над строчками расширенной матрицы соответствует некоторое элементарное преобразование в системе. Таким образом, элементарные преобразования в системе сводятся к соответствующим преобразованиям над строчками ее расширенной матрицы. Отсюда, в силу предложения 2 из § 9, следует, в частности, что *элементарные преобразования системы обратимы*, т. е. если мы, сделав элементарное преобразование, перешли от одной системы к другой, то мы можем вернуться от полученной новой системы к первоначальной, выполнив опять некоторое элементарное преобразование.

В некоторых случаях может возникнуть надобность записать неизвестные в другом порядке. Ясно, что в матрице A для этого потребуется по-другому расположить ее столбцы. Согласно теореме 1 из § 3 переход в матрице A к другому расположению столбцов может быть осуществлен посредством нескольких элементарных преобразований типа 3*.

Определение. *Две системы линейных уравнений от одних и тех же неизвестных называются равносильными, если каждое решение одной из них является решением другой, и наоборот (или если обе системы несовместны).*

Заметим, что число уравнений в системах может быть различным.

Если мы имеем две равносильные системы, то, определив решения одной из них, мы тем самым будем знать решения другой. Ясно, что решать мы будем ту систему, которая проще.

Уравнение вида

$$0\cdot x_1+0\cdot x_2+\ldots+0\cdot x_n=0$$

удовлетворяется, очевидно, при любых значениях неизвестных. Следовательно, если мы припишем такое уравнение к некоторой системе или, наоборот, вычеркнем его из системы, то новая система будет равносильна первоначальной.

Теорема. *При элементарных преобразованиях система переходит в равносильную систему.*

Доказательство. Пусть в системе (1) сделано одно из элементарных преобразований. Полученную систему мы обозначим через (1′).

При преобразованиях 1° и 2° одно из уравнений в системе, например первое, заменяется соответственно на уравнение вида

$$ca_{11}x_1+ca_{12}x_2+\ldots+ca_{1n}x_n=cb_1,$$

$$(a_{11}+\alpha a_{21})x_1+(a_{12}+\alpha a_{22})x_2+\ldots+(a_{1n}+\alpha a_{2n})x_n= \\ =b_1+\alpha b_2.$$

Ясно, что если числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ удовлетворяют всем уравнениям системы (1), то они удовлетворяют и этим двум уравнениям. Следовательно, всякое решение системы (1) является также решением и системы (1′) (в случае преобразования 3° это очевидно). Благодаря обратимости элементарных преобразований справедливо и обратное: всякое решение системы (1′) является решением и системы (1). Таким образом, если системы (1) и (1′) имеют решения, то эти решения для обеих систем одинаковы. Ясно также, что если одна из систем (1) или (1′) несовместна, то несовместной будет и другая система. Следовательно, системы (1) и (1′) равносильны, а это и требовалось доказать.

При помощи элементарных преобразований мы можем значительно упростить заданную систему. Решив полученную упрощенную систему, мы, согласно теореме, найдем тем самым и решения исходной системы. При этом упрощений можно достигать, конечно, разными способами.

Метод Гаусса решения систем линейных уравнений состоит в том, что при помощи элементарных преобразований систему приводят к такому виду, чтобы ее матрица из коэффициентов оказалась трапециевидной или близкой к трапециевидной; при этом иногда приходится изменять нумерацию неизвестных, т. е. применять преобразования 3*. После того как матрица системы приняла трапециевидную форму, уже не представляет труда разобраться в вопросе о совместности системы, определить число решений и найти сами решения.

Для иллюстрации рассмотрим несколько примеров.

П р и м е р 1. Решить систему

$$\left.\begin{aligned} x_1+2x_2-3x_3+5x_4&=1,\\ x_1+3x_2-13x_3+22x_4&=-1\\ 3x_1+5x_2+x_3-2x_4&=5,\\ 2x_1+3x_2+4x_3-7x_4&=4. \end{aligned}\right\}$$

Пользуясь элементарными преобразованиями, подвергнем систему упрощению. Для этого из второго, третьего и четвертого уравнений вычтем первое, умноженное соответственно на 1, 3, 2. Мы получим

$$\left.\begin{aligned} x_1+2x_2-3x_3+5x_4&=1,\\ x_2-10x_3+17x_4&=-2,\\ -x_2+10x_3-17x_4&=2,\\ -x_2+10x_3-17x_4&=2. \end{aligned}\right\}$$

Второе уравнение прибавим теперь к третьему и четвертому. В результате приходим к системе

$$\left.\begin{aligned} x_1+2x_2-\ 3x_3+\ 5x_4&=1,\\ x_2-10x_3+17x_4&=-2,\\ 0&=0,\\ 0&=0. \end{aligned}\right\}$$

Последние два уравнения превратились в уравнения вида $0\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4=0$. Эти уравнения удовлетворяются при любых значениях неизвестных, и их можно отбросить. Чтобы удовлетворить второму уравнению, мы можем для x_3 и x_4 выбрать произвольные значения $x_3=\alpha$, $x_4=\beta$, тогда значение для x_2 определится уже однозначно: $x_2=10\alpha-17\beta-2$. Из первого уравнения значение для x_1 также находится однозначно: $x_1=-17\alpha+29\beta+5$. Так как последняя система равносильна заданной, то формулы

$$\begin{aligned} x_1&=-17\alpha+29\beta+5,\\ x_2&=10\alpha-17\beta-2,\\ x_3&=\alpha,\\ x_4&=\beta \end{aligned}$$

при произвольных α и β дают нам все решения заданной системы.

Пример 2. Рассмотрим систему

$$\left.\begin{aligned} x_1-2x_2+3x_3-\ 4x_4&=2,\\ 3x_1+3x_2-5x_3+\ \ x_4&=-3,\\ -2x_1+\ x_2+2x_3-\ 3x_4&=5,\\ 3x_1\qquad\ +3x_3-10x_4&=8. \end{aligned}\right\}$$

Так как каждому элементарному преобразованию системы соответствует элементарное преобразование над строчками расширенной матрицы этой системы и наоборот, то можно вместо системы оперировать с расширенной матрицей этой системы (производя при этом элементарные преобразования только над строчками):

$$\left\|\begin{array}{rrrrr} 1 & -2 & 3 & -4 & 2 \\ 3 & 3 & -5 & 1 & -3 \\ -2 & 1 & 2 & -3 & 5 \\ 3 & 0 & 3 & -10 & 8 \end{array}\right\| \to$$

$$\to \left\|\begin{array}{rrrrr} 1 & -2 & 3 & -4 & 2 \\ 0 & 9 & -14 & 13 & -9 \\ 0 & -3 & 8 & -11 & 9 \\ 0 & 6 & -6 & 2 & 2 \end{array}\right\| \to$$

$$\to \left\|\begin{array}{rrrrr} 1 & -2 & 3 & -4 & 2 \\ 0 & -3 & 8 & -11 & 9 \\ 0 & 9 & -14 & 13 & -9 \\ 0 & 6 & -6 & 2 & 2 \end{array}\right\| \to$$

$$\to \left\|\begin{array}{rrrrr} 1 & -2 & 3 & -4 & 2 \\ 0 & -3 & 8 & -11 & 9 \\ 0 & 0 & 10 & -20 & 18 \\ 0 & 0 & 10 & -20 & 20 \end{array}\right\| \to \left\|\begin{array}{rrrrr} 1 & -2 & 3 & -4 & 2 \\ 0 & -3 & 8 & -11 & 9 \\ 0 & 0 & 10 & -20 & 18 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right\|.$$

Таким образом, заданная система равносильна следующей:

$$\left.\begin{array}{r} x_1 - 2x_2 + 3x_3 - 4x_4 = 2, \\ -3x_2 + 8x_3 - 11x_4 = 9, \\ 10x_3 - 20x_4 = 18, \\ 0 = 2. \end{array}\right\}$$

Полученная система несовместна, так как ее последнее уравнение $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 2$ не может быть удовлетворено никакими значениями неизвестных. Следовательно, заданная система также несовместна.

Пример 3. Решить систему

$$\left.\begin{array}{r} x_1 + 2x_2 + 3x_3 - 2x_4 = 1, \\ 2x_1 - x_2 - 2x_3 - 3x_4 = 2, \\ 3x_1 + 2x_2 - x_3 + 2x_4 = -5, \\ 2x_1 - 3x_2 + 2x_3 + x_4 = 11. \end{array}\right\}$$

При помощи элементарных преобразований над строчками упрощаем расширенную матрицу системы:

$$\begin{Vmatrix} 1 & 2 & 3 & -2 & 1 \\ 2 & -1 & -2 & -3 & 2 \\ 3 & 2 & -1 & 2 & -5 \\ 2 & -3 & 2 & 1 & 11 \end{Vmatrix} \to$$

$$\to \begin{Vmatrix} 1 & 2 & 3 & -2 & 1 \\ 0 & -5 & -8 & 1 & 0 \\ 0 & -4 & -10 & 8 & -8 \\ 0 & -7 & -4 & 5 & 9 \end{Vmatrix} \to$$

(из второй строчки вычтем третью и изменим знак у второй строчки)

$$\to \begin{Vmatrix} 1 & 2 & 3 & -2 & 1 \\ 0 & 1 & -2 & 7 & -8 \\ 0 & -4 & -10 & 8 & -8 \\ 0 & -7 & -4 & 5 & 9 \end{Vmatrix} \to$$

$$\to \begin{Vmatrix} 1 & 2 & 3 & -2 & 1 \\ 0 & 1 & -2 & 7 & -8 \\ 0 & 0 & -18 & 36 & -40 \\ 0 & 0 & -18 & 54 & -47 \end{Vmatrix} \to$$

$$\to \begin{Vmatrix} 1 & 2 & 3 & -2 & 1 \\ 0 & 1 & -2 & 7 & -8 \\ 0 & 0 & -18 & 36 & -40 \\ 0 & 0 & 0 & 18 & -7 \end{Vmatrix}.$$

Заданная система равносильна, таким образом, следующей:

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_3 - 2x_4 &= 1, \\ x_2 - 2x_3 + 7x_4 &= -8, \\ -18x_3 + 36x_4 &= -40, \\ 18x_4 &= -7. \end{aligned}\right\}$$

Эта система имеет единственное решение:

$$x_1 = \frac{2}{3}, \quad x_2 = -\frac{43}{18}, \quad x_3 = \frac{13}{9}, \quad x_4 = -\frac{7}{18}$$

(находим сначала x_4 из последнего уравнения, затем x_3 из третьего и т. д.).

Для сравнения рекомендуем решить систему этого примера при помощи формул Крамера.

Упражнения

1. Решить системы линейных уравнений:

$$\text{a)}\ \left.\begin{aligned} x_1 - x_2 + 2x_3 - 3x_4 &= 1,\\ x_1 + 4x_2 - x_3 - 2x_4 &= -2,\\ x_1 - 4x_2 + 3x_3 - 2x_4 &= -2,\\ x_1 - 8x_2 + 5x_3 - 2x_4 &= -2; \end{aligned}\right\}$$

$$\text{b)}\ \left.\begin{aligned} x_1 + 2x_2 - 3x_3 + 4x_4 &= 7,\\ 2x_1 + 5x_2 + x_3 - 2x_4 &= 5,\\ 3x_1 - 7x_2 + 4x_3 + 5x_4 &= -11,\\ 7x_1 + 2x_2 - x_3 + 11x_4 &= 6; \end{aligned}\right\}$$

$$\text{c)}\ \left.\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 7,\\ 2x_1 + x_2 + 2x_3 + 3x_4 &= 6,\\ 3x_1 + 2x_2 + x_3 + 2x_4 &= 7,\\ 4x_1 + 3x_2 + 2x_3 + x_4 &= 18. \end{aligned}\right\}$$

Отв. a) $x_1 = -8$, $x_2 = 3 + \alpha$, $x_3 = 6 + 2\alpha$, $x_4 = \alpha$ (α — произвольно); b) несовместна; c) $x_1 = 2$, $x_2 = 1$, $x_3 = 5$, $x_4 = -3$.

2. Решить систему при всевозможных значениях параметра t:

$$\left.\begin{aligned} 2x - y + 3z &= -7,\\ x + 2y - 6z &= t,\\ tx + 5y - 15z &= 8. \end{aligned}\right\}$$

Отв. При $t \neq 5$ и $t \neq -1$ система несовместна. Если $t = 5$, то $x = -\frac{9}{5}$, $y = \frac{15\alpha + 17}{5}$, $z = \alpha$; если же $t = -1$, то $x = -3$, $y = 3\alpha + 1$, $z = \alpha$ (α — произвольно).

3. Решить систему при всевозможных значениях параметра t:

$$\left.\begin{aligned} x + y - z &= 2,\\ x + ty - tz &= 2,\\ (1 + t)\,x + 2ty - tz &= 3 + 2t. \end{aligned}\right\}$$

Отв. При $t \neq 0$ и $t \neq 1$ система имеет единственное решение $x = 2$, $y = \frac{1}{t}$, $z = \frac{1}{t}$. При $t = 0$ система несовместна; при $t = 1$: $x = 3 - \alpha$, $y = \alpha$, $z = 1$.

4. Решить систему линейных уравнений

$$\left.\begin{array}{l} x_2+x_3+\ldots+x_{n-1}+x_n=0, \\ x_1 \qquad +x_3+\ldots+x_{n-1}+x_n=1, \\ x_1+x_2 \qquad +\ldots+x_{n-1}+x_n=2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1+x_2+x_3+\ldots+x_{n-1} \qquad =n-1. \end{array}\right\}$$

Отв. $x_k=\frac{n}{2}-k+1 \quad (k=1, 2, \ldots, n)$.

Указание. К системе приписать уравнение $0=0$ и к нему прибавить остальные уравнения.

§ 11. Условие совместности

В предыдущем параграфе нами был изложен довольно простой метод для практического решения системы линейных уравнений с числовыми коэффициентами, дающий возможность, в частности, ответить на вопрос о совместности системы. Однако этот метод не дает возможности сформулировать условие совместности системы в терминах ее коэффициентов и свободных членов. Кроме того, в случае существования решений он не дает явных формул для этих решений, подобных формулам Крамера. В то же время и то и другое необходимо при рассмотрении различных теоретических вопросов.

В этом параграфе мы исследуем в общем виде произвольную систему m линейных уравнений с n неизвестными:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m. \end{array}\right\} \qquad (1)$$

Рассмотрим матрицу A этой системы и ее расширенную матрицу B:

$$A=\left\|\begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{array}\right\|, \qquad B=\left\|\begin{array}{ccccc} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right\|.$$

Ясно, что ранги этих матриц связаны неравенством $r_B \geq r_A$; при этом, согласно замечанию 2 из § 9, ранг матрицы B может быть лишь на одну единицу больше r_A.

Вопрос о совместности системы (1) полностью решается следующей теоремой.

Теорема Кронекера — Капелли. *Для того чтобы система линейных уравнений была совместна, необходимо и достаточно, чтобы ранг матрицы этой системы был равен рангу ее расширенной матрицы, т. е. чтобы* $r_A = r_B$.

Доказательство. Необходимость. Пусть система (1) совместна и $\alpha_1, \alpha_2, \ldots, \alpha_n$ — какое-нибудь ее решение. Тогда имеют место равенства

$$\left.\begin{array}{l} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n = b_1, \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2n}\alpha_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n = b_m. \end{array}\right\}$$

В расширенной матрице B системы (1) выполним следующие элементарные преобразования. Из последнего столбца вычтем первый столбец, умноженный на α_1, второй, умноженный на α_2, и т. д., наконец, n-й столбец, умноженный на α_n. Получим матрицу

$$C = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 - a_{11}\alpha_1 - a_{12}\alpha_2 - \ldots - a_{1n}\alpha_n \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 - a_{21}\alpha_1 - a_{22}\alpha_2 - \ldots - a_{2n}\alpha_n \\ \cdot & \cdot & \cdot & \cdot & \cdots\cdots\cdots\cdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m - a_{m1}\alpha_1 - a_{m2}\alpha_2 - \ldots - a_{mn}\alpha_n \end{Vmatrix}.$$

В силу выписанных выше равенств последний столбец матрицы C состоит из нулей:

$$C = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} & 0 \end{Vmatrix}.$$

Так как матрица C получается из B при помощи элементарных преобразований, то $r_C = r_B$ по теореме 1 из § 9. С другой стороны, C отличается от A только столбцом из нулей, поэтому $r_C = r_A$ (замечание 1 из § 9). Следовательно, $r_B = r_A$, и необходимость условия теоремы доказана.

Достаточность. Предположим, что ранги матриц A и B равны, и докажем, что система (1) совместна.

Пусть r обозначает ранг матрицы A. Среди миноров r-го порядка матрицы A имеется минор, отличный от нуля. Не нарушая общности рассуждения, можно считать, что отличный от нуля минор r-го порядка расположен в левом верхнем углу:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

(Это всегда можно устроить посредством изменения нумерации неизвестных и перестановок уравнений в системе.)

Выделим в системе (1) первые r уравнений и составим из них новую систему

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rn}x_n = b_r. \end{array}\right\} \qquad (2)$$

Заметим, что в случае $r = m$ система (2) будет совпадать с системой (1).

Покажем, что системы (1) и (2) равносильны (при выполнении условия $r_B = r_A = r$).

Тот факт, что каждое решение системы (1) (если оно существует) является также и решением системы (2), очевиден. Мы должны проверить, что и, наоборот, каждое решение системы (2), если оно существует, является также решением системы (1).

Пусть $\alpha_1, \alpha_2, \dots, \alpha_n$ — произвольное решение системы (2), так что имеют место равенства

$$\left.\begin{array}{l} a_{11}\alpha_1 + a_{12}\alpha_2 + \dots + a_{1n}\alpha_n = b_1, \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \dots + a_{2n}\alpha_n = b_2, \\ \dots\dots\dots\dots\dots\dots \\ a_{r1}\alpha_1 + a_{r2}\alpha_2 + \dots + a_{rn}\alpha_n = b_r. \end{array}\right\} \qquad (3)$$

Числа $\alpha_1, \alpha_2, \dots, \alpha_n$ удовлетворяют, конечно, первым r

уравнениям системы (1). Мы должны убедиться, что они удовлетворяют и последним $m-r$ уравнениям, т. е. что при $s=r+1,\ r+2,\ \ldots,\ m$ справедливы равенства

$$a_{s1}\alpha_1+a_{s2}\alpha_2+\ldots+a_{sn}\alpha_n=b_s. \tag{4}$$

Рассмотрим расширенную матрицу

$$B=\begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \ldots & a_{rn} & b_r \\ a_{r+1,1} & a_{r+1,2} & \ldots & a_{r+1,n} & b_{r+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{Vmatrix}$$

и подвергнем ее, как и при доказательстве необходимости, элементарным преобразованиям: из последнего столбца вычтем все остальные столбцы, умноженные соответственно на $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$. Мы получим матрицу

$$C=\begin{Vmatrix} a_{11} & \ldots & a_{1n} & b_1-a_{11}\alpha_1-\ldots-a_{1n}\alpha_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{r1} & \ldots & a_{rn} & b_r-a_{r1}\alpha_1-\ldots-a_{rn}\alpha_n \\ a_{r+1,1} & \ldots & a_{r+1,n} & b_{r+1}-a_{r+1,1}\alpha_1-\ldots-a_{r+1,n}\alpha_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & \ldots & a_{mn} & b_m-a_{m1}\alpha_1-\ldots-a_{mn}\alpha_n \end{Vmatrix}.$$

Первые r элементов последнего столбца матрицы C равны нулю в силу равенств (3). Что касается остальных элементов последнего столбца, то о них нам пока ничего не известно. Нашей задачей является доказательство равенств (4) для всех $s=r+1,\ r+2,\ \ldots,\ m$. Но доказательство этих равенств, очевидно, равносильно доказательству того, что все прочие элементы последнего столбца матрицы C также равны нулю. Положим для краткости (при $s=r+1,\ \ldots,\ m$):

$$\gamma_s=b_s-a_{s1}\alpha_1-a_{s2}\alpha_2-\ldots-a_{sn}\alpha_n. \tag{5}$$

Тогда матрицу C можно будет переписать так:

$$C=\left\|\begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 \\ a_{12} & a_{22} & \ldots & a_{2n} & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \ldots & a_{rn} & 0 \\ a_{r+1,1} & a_{r+1,2} & \ldots & a_{r+1,n} & \gamma_{r+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} & \gamma_m \end{matrix}\right\|.$$

Матрица C получена из B при помощи элементарных преобразований, поэтому ранг C также равен r. Но в таком случае все ее миноры порядка $r+1$ равны нулю. В частности, равен нулю и минор $(r+1)$-го порядка, составленный из столбцов с номерами $1, 2, \ldots$ $\ldots, r, n+1$ и из строк с номерами $1, 2, \ldots, r, s$ (где s — любое из чисел $r+1, \ldots, m$), т. е.

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} & 0 \\ a_{21} & a_{22} & \ldots & a_{2r} & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \ldots & a_{rr} & 0 \\ a_{s1} & a_{s2} & \ldots & a_{sr} & \gamma_s \end{vmatrix}=0.$$

Разлагая этот определитель по элементам последнего столбца, получим

$$\Delta\cdot\gamma_s=0.$$

Но $\Delta\neq 0$, поэтому полученное равенство свидетельствует о том, что $\gamma_s=0$. Таким образом, все числа (5) равны нулю, а значит, справедливы равенства (4). Это доказывает, что всякое решение системы (2) удовлетворяет всем уравнениям системы (1).

Установлено, таким образом, что системы (1) и (2) равносильны. Вопрос о существовании решений у системы (1) сведен, следовательно, к вопросу о существовании решений у системы (2).

Если $r=n$, то система (2) есть система n уравнений с n неизвестными и с не равным нулю определителем ($\Delta\neq 0$). Согласно теореме Крамера она имеет единственное решение. Следовательно, система (1) также имеет единственное решение.

В случае $r<n$ неизвестным $x_{r+1}, \ldots, x_n$ придадим произвольные значения $\alpha_{r+1}, \ldots, \alpha_n$ и рассмотрим систему

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1r}x_r=b_1-a_{1,\,r+1}\alpha_{r+1}-\ldots-a_{1n}\alpha_n, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2r}x_r=b_2-a_{2,\,r+1}\alpha_{r+1}-\ldots-a_{2n}\alpha_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+a_{r2}x_2+\ldots+a_{rr}x_r=b_r-a_{r,\,r+1}\alpha_{r+1}-\ldots-a_{rn}\alpha_n. \end{array}\right\}$$

Определитель Δ этой системы r уравнений с r неизвестными не равен нулю, а потому она имеет единственное решение

$$x_1=\alpha_1, \quad x_2=\alpha_2, \ \ldots, \ x_r=\alpha_r.$$

Совокупность чисел $\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n$ будет, очевидно, решением системы (2). Так что в этом случае система (2) имеет бесконечно много решений (ведь значения для неизвестных $x_{r+1}, \ldots, x_n$ можно выбрать бесконечным числом способов). Следовательно, система (1) при $r<n$ также имеет бесконечно много решений.

Таким образом, если $r_A=r_B$, то система (1) совместна.

Теорема Кронекера — Капелли доказана полностью.

Из доказательства теоремы Кронекера — Капелли легко получить также ответ на вопрос о числе решений (в случае совместности системы).

Теорема о числе решений. *Пусть для системы m линейных уравнений с n неизвестными выполнено условие совместности, т. е. ранг r матрицы из коэффициентов системы равен рангу ее расширенной матрицы. Тогда, если ранг матрицы системы равен числу неизвестных ($r=n$), то система имеет единственное решение. Если ранг матрицы системы меньше числа неизвестных ($r<n$), то система имеет бесконечно много решений, а именно: некоторым $n-r$ неизвестным можно придавать произвольные значения, тогда оставшиеся r неизвестных определятся уже единственным образом.*

В самом деле, если $r=n$, то система (2) имеет единственное решение. Если же $r<n$, то система (2) имеет бесконечно много решений и все ее решения являются решениями системы (1).

Замечание 1. Если ранг матрицы системы линейных уравнений равен числу уравнений, т. е. если

$r=m$, то система совместна при любых свободных членах. Действительно, в этом случае ранг расширенной матрицы также равен m, так как ранг матрицы не может быть больше числа ее строчек.

Замечание 2. Доказательство теоремы Кронекера — Капелли позволяет также дать явные формулы для решений системы (в случае ее совместности). Если уже известно, что система совместна, то, чтобы найти ее решения, необходимо:

1) отыскать в матрице системы A ранга $r_A=r$ отличный от нуля минор r-го порядка Δ;

2) отбросить те уравнения, которые соответствуют строчкам матрицы A, не входящим в определитель Δ;

3) члены с коэффициентами, не входящими в определитель Δ, перенести в правую часть, а затем, придавая неизвестным, находящимся в правой части, произвольные значения, определить по формулам Крамера оставшиеся r неизвестных из системы r уравнений с отличным от нуля определителем Δ.

Пример. Решить систему

$$\left.\begin{aligned} x-2y+3z-5u&=2,\\ 2x+y+4z+u&=-3,\\ 3x-3y+8z-2u&=-1,\\ 2x-2y+5z-12u&=4, \end{aligned}\right\}$$

следуя доказательству теоремы Кронекера — Капелли.

Вычисляем ранг матрицы этой системы и ранг расширенной матрицы. В обоих случаях он равен 3. Следовательно, система совместна. Так как ранг матрицы системы меньше числа неизвестных, то система имеет бесконечно много решений: одно неизвестное может быть взято произвольно. Минор

$$\begin{vmatrix} 1 & -2 & 3\\ 2 & 1 & 4\\ 3 & -3 & 8 \end{vmatrix}=1$$

отличен от нуля, поэтому последнее уравнение отбрасываем и неизвестному u придаем произвольное значение α. Оставшиеся неизвестные определяются из системы

$$\left.\begin{aligned} x-2y+3z &= 2+5\alpha, \\ 2x+y+4z &= -3-\alpha, \\ 3x-3y+8z &= -1+2\alpha. \end{aligned}\right\}$$

Решая последнюю систему (по формулам Крамера или иным способом), находим $x=30+71\alpha$, $y=-7-15\alpha$, $z=-14-32\alpha$. Присоединяя сюда $u=\alpha$ и придавая α произвольные значения, получаем все решения заданной системы.

Заметим, что приведенный нами путь решения примера не является единственным. В качестве Δ можно было взять минор

$$\begin{vmatrix} 1 & 4 & 1 \\ -3 & 8 & -2 \\ -2 & 5 & -12 \end{vmatrix},$$

который также отличен от нуля. В этом случае мы отбросим первое уравнение и x придадим произвольное значение, а все остальные неизвестные найдем из трех последних уравнений. Ответ мы получим, конечно, в другой форме.

Замечание 3. Изложенное в настоящем параграфе применимо, разумеется, и к случаю, когда число уравнений в системе совпадает с числом неизвестных. Пусть Δ — определитель некоторой системы n линейных уравнений с n неизвестными. Если $\Delta \neq 0$, то ранг матрицы из коэффициентов r равен числу неизвестных n, и поэтому система совместна при любых свободных членах (см. замечание 1) и имеет единственное решение. Этот факт совпадает, как видим, с теоремой Крамера, которая только к этому случаю и применима. Если же $\Delta=0$, то $r<n$, и тогда возможны два случая: либо система несовместна (если $r_B \neq r$), либо она имеет бесконечно много решений (если $r_B=r$). Мы видим, таким образом, что система n линейных уравнений с n неизвестными имеет единственное решение тогда и только тогда, когда определитель этой системы отличен от нуля.

§ 12. Однородные системы линейных уравнений

Линейное уравнение вида

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

называется *однородным*, если в нем свободный член b равен нулю. В соответствии с этим систему линейных уравнений мы называем *однородной*, если все ее свободные члены равны нулю. Произвольная однородная система линейных уравнений имеет, таким образом, вид

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0. \end{aligned}\right\} \quad (1)$$

Однородная система всегда совместна, так как она имеет следующее очевидное решение:

$$x_1 = 0, \quad x_2 = 0, \ \ldots, \ x_n = 0.$$

Это решение (у которого все значения неизвестных равны нулю) называется *нулевым* или *тривиальным*. Всякое другое решение (если оно есть), у которого значение хоть одного неизвестного отлично от нуля, называется *ненулевым* или *нетривиальным*.

Совместность системы (1) можно, конечно, установить и с помощью теоремы Кронекера — Капелли: так как расширенная матрица B получается из матрицы системы A приписыванием столбца из нулей, то, по замечанию 1 из § 9, мы имеем $r_B = r_A$, а значит, система совместна.

Для однородных систем представляет особый интерес вопрос о существовании ненулевых решений.

Обозначим ранг матрицы системы (1) через r и применим к ней теорему § 11 о числе решений. Если $r = n$, то система имеет единственное решение, и так как этим единственным решением является нулевое решение, то ненулевых решений в этом случае нет. Если же $r < n$, то система имеет бесконечно много решений и среди них, понятно, имеется бесконечно много ненулевых решений (ведь нулевое решение только одно). Имеет место, таким образом, следующая основная теорема об однородных системах.

Т е о р е м а 1. *Для того чтобы однородная система линейных уравнений имела ненулевое решение, необходимо и достаточно, чтобы ранг матрицы системы был меньше числа неизвестных* ($r < n$).

В частном случае, когда число уравнений совпадает с числом неизвестных ($m = n$), условие $r < n$ равносильно, очевидно, тому, что определитель системы равен нулю ($\Delta = 0$). Для этого частного случая мы можем поэтому сформулировать следующую теорему.

Т е о р е м а 2. *Для того чтобы однородная система n линейных уравнений с n неизвестными имела ненулевое решение, необходимо и достаточно, чтобы определитель этой системы был равен нулю.*

Отметим еще один частный случай, когда $m < n$. Так как ранг матрицы не может быть больше числа ее строчек (как и числа столбцов), то $r \leqslant m$. Следовательно, если $m < n$, то и $r < n$, а значит, по теореме 1, система имеет ненулевое решение. Этим доказана следующая теорема.

Т е о р е м а 3. *Если число уравнений в однородной системе линейных уравнений меньше числа неизвестных, то эта система имеет ненулевые решения.*

П р и м е р. Покажем, что на плоскости с декартовой системой координат три точки (x_1, y_1), (x_2, y_2), (x_3, y_3) — лежат на одной прямой тогда и только тогда, когда определитель

$$\Delta = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

равен нулю. Пусть $ax + by + c = 0$ есть уравнение прямой, на которой лежат наши точки. Тогда

$$\left.\begin{aligned} ax_1 + by_1 + c = 0, \\ ax_2 + by_2 + c = 0, \\ ax_3 + by_3 + c = 0. \end{aligned}\right\} \qquad (2)$$

Так как a, b, c не равны 0 одновременно, то последние равенства означают, что однородная система трех уравнений с тремя неизвестными и с определителем Δ имеет ненулевое решение, а тогда, по теореме 2, $\Delta = 0$. Наоборот, если $\Delta = 0$, то существуют a, b, c, не равные 0 одновременно, для которых выполнены равенства (2). Если бы a и b вместе равнялись нулю, то и c равнялось бы нулю, т. е. все три числа a, b, c были бы нулями, а это не так. Следовательно, хотя бы одно из чисел a, b отлично от 0, а значит, $ax+by+c=0$ является уравнением прямой. Координаты наших точек удовлетворяют этому уравнению, поэтому все они лежат на одной прямой.

Упражнения

1. Решить систему

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

Отв. $x_1=-3\alpha+5\beta$, $x_2=\alpha-3\beta$, $x_3=\alpha$, $x_4=\beta$ (α и β произвольны).

2. Решить систему при всевозможных значениях параметра λ:

$$\left.\begin{aligned} (2-\lambda)x_1+x_2+x_3&=0,\\ x_1+(2-\lambda)x_2+x_3&=0,\\ x_1+x_2+(2-\lambda)x_3&=0.\end{aligned}\right\}$$

Отв. При $\lambda=1$ $x_1=\alpha$, $x_2=\beta$, $x_3=-\alpha-\beta$; при $\lambda=4$ $x_1=\alpha$, $x_2=\alpha$, $x_3=\alpha$; при остальных значениях λ система имеет только нулевое решение.

3. При каких λ однородная система

$$\left.\begin{aligned} -\lambda x_1+x_2+\dots+x_n&=0,\\ x_1-\lambda x_2+\dots+x_n&=0,\\ \dots\dots\dots&\dots\dots\\ x_1+x_2+\dots-\lambda x_n&=0\end{aligned}\right\}$$

имеет ненулевые решения?

Отв. При $\lambda=n-1$ и $\lambda=-1$ система имеет ненулевые решения.

ГЛАВА III

ДЕЙСТВИЯ НАД МАТРИЦАМИ

§ 13. Умножение матриц

В различных разделах математики часто приходится иметь дело с переменными $x_1, x_2, \ldots, x_m$, которые являются функциями других переменных $y_1, y_2, \ldots$ $\ldots, y_n$:

$$\left.\begin{array}{l} x_1 = \varphi_1(y_1, y_2, \ldots, y_n), \\ x_2 = \varphi_2(y_1, y_2, \ldots, y_n), \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_m = \varphi_m(y_1, y_2, \ldots, y_n). \end{array}\right\}$$

Такого рода выражение одной системы переменных через другую носит название *преобразования переменных*.

Одним из примеров преобразований переменных являются формулы преобразования координат в аналитической геометрии. Пусть на плоскости мы имеем две прямоугольные декартовы системы координат xOy и $x'Oy'$ с общим началом O, причем вторая система расположена под углом α относительно первой. Тогда, как известно, координаты x, y произвольной точки M в первой системе координат выражаются через координаты x', y' той же точки во второй системе по формулам

$$\left.\begin{array}{l} x = x'\cos\alpha - y'\sin\alpha, \\ y = x'\sin\alpha + y'\cos\alpha. \end{array}\right\}$$

Если одну из координатных систем, например $x'Oy'$, мы заменим на косоугольную и при этом изменим масштабы на каждой из ее координатных осей, то формулы преобразования координат примут более общий вид:

$$\left.\begin{aligned} x &= ax' + by', \\ y &= cx' + dy', \end{aligned}\right\}$$

где коэффициенты могут быть любыми вещественными числами, удовлетворяющими единственному условию $ad - bc \neq 0$.

Аналогичные формулы преобразования координат имеют место и для точек пространства:

$$\left.\begin{aligned} x &= a_1x' + b_1y' + c_1z', \\ y &= a_2x' + b_2y' + c_2z', \\ z &= a_3x' + b_3y' + c_3z'. \end{aligned}\right\}$$

Приведем еще один пример. Пусть через начало O некоторой системы координат $Oxyz$ в пространстве проведена плоскость P, на которой выбрана своя система координат uOv (с тем же началом O). Каждую точку пространства M спроектируем на плоскость P. Тогда координаты u, v проекции будут выражаться через координаты x, y, z проектируемой точки по формулам

$$\left.\begin{aligned} u &= a_1x + a_2y + a_3z, \\ v &= b_1x + b_2y + b_3z, \end{aligned}\right\}$$

где a_1, a_2, a_3, b_1, b_2, b_3 — некоторые постоянные вещественные числа.

Во всех рассмотренных нами примерах одна система переменных выражается через другую при помощи однородных многочленов первой степени. Такие преобразования носят название линейных. Дадим определение линейного преобразования в общем виде.

Определение. *Линейным преобразованием переменных называется выражение одной системы переменных x_1, x_2, ..., x_m через новую систему переменных y_1, y_2, ..., y_n при помощи линейных однородных функций:*

$$\left.\begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n, \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_m &= a_{m1}y_1 + a_{m2}y_2 + \ldots + a_{mn}y_n. \end{aligned}\right\} \qquad (1)$$

Здесь a_{ij} — постоянные числа (вещественные или комплексные). Заметим, что число старых переменных $x_1, x_2, \ldots, x_m$ и число новых переменных $y_1, y_2, \ldots, y_n$ не обязательно одно и то же.

Матрица

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix},$$

составленная из коэффициентов при переменных $y_1, y_2, \ldots, y_n$, называется *матрицей линейного преобразования* (1).

Если матрицы двух линейных преобразований одинаковы (т. е. имеют одинаковое число строчек и столбцов и соответствующие элементы обоих матриц равны), то эти преобразования могут отличаться друг от друга лишь наименованием переменных, что не имеет существенного значения. Мы условимся поэтому два линейных преобразования считать одинаковыми или равными, если их матрицы равны. Например, линейные преобразования

$$\left.\begin{aligned} x_1 &= 2y_1 + 3y_2, \\ x_2 &= y_1 - 2y_2 \end{aligned}\right\} \quad \text{и} \quad \left.\begin{aligned} x &= 2u + 3v, \\ y &= u - 2v \end{aligned}\right\}$$

одинаковы, так как оба преобразования имеют одну и ту же матрицу

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

Пусть помимо линейного преобразования (1) задано другое линейное преобразование, выражающее переменные $y_1, y_2, \ldots, y_n$ через третью систему переменных $z_1, z_2, \ldots, z_p$:

$$\left.\begin{aligned} y_1 &= b_{11}z_1 + b_{12}z_2 + \ldots + b_{1p}z_p, \\ y_2 &= b_{21}z_1 + b_{22}z_2 + \ldots + b_{2p}z_p, \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ y_n &= b_{n1}z_1 + b_{n2}z_2 + \ldots + b_{np}z_p. \end{aligned}\right\} \qquad (2)$$

Матрицу линейного преобразования (2) обозначим через B. Так как переменные $x_1, x_2, \ldots, x_m$ являются функциями переменных $y_1, y_2, \ldots, y_n$, а последние в свою очередь являются функциями $z_1, z_2, \ldots, z_p$, то переменные $x_1, x_2, \ldots, x_m$ можно рассматривать как функции от $z_1, z_2, \ldots, z_p$. Чтобы получить явное выражение переменных $x_1, x_2, \ldots, x_m$ через $z_1, z_2, \ldots, z_p$, следует выражения для $y_1, y_2, \ldots, y_n$ из (2) подставить в правую часть равенств (1). После приведения подобных членов получим

$$\left.\begin{aligned} x_1 &= c_{11}z_1 + c_{12}z_2 + \ldots + c_{1p}z_p, \\ x_2 &= c_{21}z_1 + c_{22}z_2 + \ldots + c_{2p}z_p, \\ &\cdots\cdots\cdots\cdots\cdots \\ x_m &= c_{m1}z_1 + c_{m2}z_2 + \ldots + c_{mp}z_p, \end{aligned}\right\} \qquad (3)$$

где c_{ij} — некоторые новые постоянные числа.

Таким образом, в результате последовательного выполнения двух линейных преобразований переменных мы снова получаем линейное преобразование. Операция замены двух последовательно выполненных линейных преобразований одним называется *умножением линейных преобразований*. Полученное при этом линейное преобразование называется *произведением* двух данных линейных преобразований.

Обозначим матрицу линейного преобразования (3) через C и найдем формулы, выражающие элементы c_{ij} матрицы C через элементы матриц A и B. Элемент c_{ij} есть коэффициент при z_j в выражении для x_i. Мы имеем

$$\begin{aligned} x_i &= a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n = \\ &= a_{i1}(\ldots + b_{1j}z_j + \ldots) + a_{i2}(\ldots + b_{2j}z_j + \ldots) + \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad + a_{in}(\ldots + b_{nj}z_j + \ldots) = \\ &= \ldots + (a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj})\, z_j + \ldots \end{aligned}$$

Следовательно,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Таким образом, элемент i-й строчки j-го столбца матрицы C равен сумме произведений элементов i-й строчки матрицы A на соответствующие элементы j-го столбца матрицы B.

Поскольку линейное преобразование (3) есть произведение линейных преобразований (1) и (2), то матрицу

C естественно назвать произведением матриц A и B. Мы приходим, таким образом, к следующему определению произведения матриц, заданных вне связи с линейными преобразованиями.

Определение умножения матриц. *Пусть даны две матрицы*

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}, \quad B = \begin{Vmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{Vmatrix}, \tag{4}$$

причем число столбцов матрицы A равно числу строчек матрицы B. Произведением матрицы A на матрицу B называется матрица

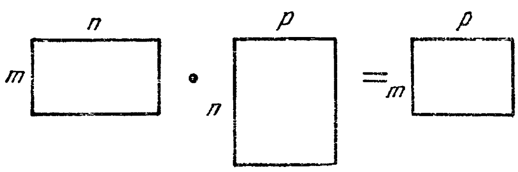
$$C = \begin{Vmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{Vmatrix},$$

обозначаемая через AB, элементы которой вычисляются по формулам

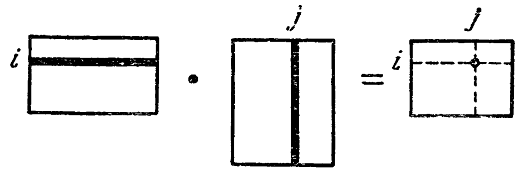
$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{s=1}^{n} a_{is}b_{sj}$$

$$(i = 1, 2, \dots, m; \; j = 1, 2, \dots, p).$$

Согласно нашему определению не всякие две матрицы можно перемножить. Произведение двух матриц имеет смысл тогда и только тогда, когда число столбцов первого множителя равно числу строчек второго множителя. При этом в произведении получается матрица, число строчек которой равно числу строчек первого множителя, а число столбцов равно числу столбцов второго множителя. Схематически последнее утверждение можно изобразить следующим образом:



Что касается правила для вычисления элементов в произведении двух матриц, то его схематически можно изобразить так:



Возвращаясь к линейным преобразованиям, можно сказать теперь, что матрица произведения двух линейных преобразований равна произведению матриц этих преобразований.

Пример 1. Перемножить матрицы

$$A=\begin{Vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{Vmatrix} \text{ и } B=\begin{Vmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{Vmatrix}.$$

Произведение AB имеет смысл, поскольку число столбцов матрицы A равно числу строчек матрицы B:

$$AB=$$
$$=\begin{Vmatrix} 1\cdot1+2\cdot1+3\cdot1 & 1\cdot1+2\cdot2+3\cdot3 & 1\cdot1+2\cdot4+3\cdot9 \\ 4\cdot1+5\cdot1+6\cdot1 & 4\cdot1+5\cdot2+6\cdot3 & 4\cdot1+5\cdot4+6\cdot9 \end{Vmatrix}=$$
$$=\begin{Vmatrix} 6 & 14 & 36 \\ 15 & 32 & 78 \end{Vmatrix}.$$

Произведение B на A здесь не имеет смысла.

Пример 2. Пусть

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

Тогда

$$AB=\begin{Vmatrix} 2 & -1 \\ 8 & 9 \end{Vmatrix}, \quad BA=\begin{Vmatrix} 2 & 7 & 9 \\ 2 & -5 & -11 \\ 2 & 10 & 14 \end{Vmatrix}.$$

Оба произведения, AB и BA, здесь имеют смысл, но являются различными матрицами (даже различных порядков).

Пример 3. Пусть

$$A=\begin{Vmatrix}2 & 5\\ 1 & 3\end{Vmatrix},\qquad B=\begin{Vmatrix}-2 & 3\\ 3 & -5\end{Vmatrix}.$$

Тогда

$$AB=\begin{Vmatrix}11 & -19\\ 7 & -12\end{Vmatrix},\qquad BA=\begin{Vmatrix}-1 & -1\\ 1 & 0\end{Vmatrix}.$$

Пример 4. Заданы два линейных преобразования:

$$\left.\begin{aligned} x_1 &= \quad y_1 - 2y_2,\\ x_2 &= \ 3y_1 + y_2,\\ x_3 &= -y_1 + y_2,\end{aligned}\right\}\qquad (a)$$

$$\left.\begin{aligned} y_1 &= z_1 + z_2 - z_3 - z_4,\\ y_2 &= 2z_1 + z_2 + z_3 - 3z_4.\end{aligned}\right\}\qquad (b)$$

Найти произведение линейных преобразований (a) и (b).

Матрица искомого линейного преобразования равна

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

следовательно,

$$\left.\begin{aligned} x_1 &= -3z_1 - z_2 - 3z_3 + 5z_4,\\ x_2 &= \ 5z_1 + 4z_2 - 2z_3 - 6z_4,\\ x_3 &= \quad z_1 \qquad\ + 2z_3 - 2z_4.\end{aligned}\right\}$$

Из примеров 2 и 3 следует, что умножение матриц не подчиняется перестановочному (коммутативному) закону, т. е., вообще говоря, $AB \neq BA$ (речь идет о случае, когда оба произведения, AB и BA, имеют смысл). Конечно, в некоторых отдельных случаях может оказаться, что $AB=BA$. Например,

$$\begin{Vmatrix}1 & 2\\ -1 & 3\end{Vmatrix}\begin{Vmatrix}-1 & 8\\ -4 & 7\end{Vmatrix}=\begin{Vmatrix}-1 & 8\\ -4 & 7\end{Vmatrix}\begin{Vmatrix}1 & 2\\ -1 & 3\end{Vmatrix}=\begin{Vmatrix}-9 & 22\\ -11 & 13\end{Vmatrix}.$$

Если для двух матриц A и B имеем $AB=BA$, то матрицы A и B называются *перестановочными*. Очевидно, что это может иметь место только в случае, когда A и B — квадратные матрицы одного и того же порядка.

Матрица

$$\|a_1,\ a_2,\ \ldots,\ a_n\|,$$

состоящая из одной строчки, называется просто *строчкой*. Матрица

$$\begin{Vmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{Vmatrix},$$

состоящая из одного столбца, называется *столбцом*. Произведение строчки на столбец имеет смысл, очевидно, только при условии, что «длина» строчки равна «высоте» столбца. При этом в произведении получается квадратная матрица 1-го порядка (имеющая одну строчку и один столбец), т. е. просто число. Например,

$$\|1\quad 2\quad 3\| \begin{Vmatrix} 2 \\ -4 \\ 1 \end{Vmatrix} = \|-3\| = -3.$$

Произведение столбца на строчку всегда имеет смысл. Например,

$$\begin{Vmatrix} a \\ b \end{Vmatrix} \|x\quad y\quad z\| = \begin{Vmatrix} ax & ay & az \\ bx & by & bz \end{Vmatrix}.$$

Каждую матрицу мы можем рассматривать как совокупность строчек (записанных в определенном порядке) или как совокупность столбцов. Например, если для матриц (4) мы положим

$$S_i = \|a_{i1}\quad a_{i2}\ \ldots\ a_{in}\|, \qquad T_j = \begin{Vmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{Vmatrix},$$

то эти матрицы можно будет записать в виде

$$A = \begin{Vmatrix} S_1 \\ S_2 \\ \vdots \\ S_m \end{Vmatrix}, \qquad B = \|T_1\quad T_2\ \ldots\ T_p\|.$$

Замечая теперь, что

$$S_iT_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

(согласно правилу умножения строчки на столбец), мы можем произведение AB представить в виде

$$AB = \begin{Vmatrix} S_1T_1 & S_1T_2 & \ldots & S_1T_p \\ S_2T_1 & S_2T_2 & \ldots & S_2T_p \\ \ldots & \ldots & \ldots & \ldots \\ S_mT_1 & S_mT_2 & \ldots & S_mT_p \end{Vmatrix}.$$

Если длина строчек матрицы A равна высоте столбца T, то имеет смысл произведение AT, которое является, очевидно, столбцом (число строчек этого столбца равно числу строчек матрицы A). Легко видеть, что для матриц (4) произведение AB может быть записано также и следующим образом:

$$AB = A \,\| T_1 \quad T_2 \ldots T_p \| = \| AT_1 \quad AT_2 \ldots AT_p \|. \quad (5)$$

Аналогично

$$AB = \begin{Vmatrix} S_1 \\ S_2 \\ \vdots \\ S_m \end{Vmatrix} B = \begin{Vmatrix} S_1B \\ S_2B \\ \vdots \\ S_mB \end{Vmatrix}.$$

Пользуясь правилом умножения матриц, можно линейное преобразование (1) записать в виде одного матричного равенства. Действительно, введем в рассмотрение столбцы

$$X = \begin{Vmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{Vmatrix}, \qquad Y = \begin{Vmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Vmatrix}$$

из переменных $x_1, x_2, \ldots, x_m$ и $y_1, y_2, \ldots, y_n$

соответственно. Тогда

$$AY = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix} \begin{Vmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Vmatrix} =$$

$$= \begin{Vmatrix} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \\ \dots\dots\dots\dots\dots \\ a_{m1}y_1 + a_{m2}y_2 + \dots + a_{mn}y_n \end{Vmatrix} = \begin{Vmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{Vmatrix} = X.$$

Мы видим, таким образом, что система равенств (1) равносильна матричному равенству

$$X = AY.$$

(Здесь мы использовали условие равенства матриц: две матрицы равны тогда и только тогда, когда они имеют одинаковое строение и равны их соответствующие элементы.)

Как мы уже отмечали, умножение матриц некоммутативно. Однако для матриц справедлив сочетательный (ассоциативный) закон умножения.

Теорема 1. *Пусть A, B и C — три матрицы, для которых произведения AB и BC имеют смысл. Тогда произведения $(AB)C$ и $A(BC)$ также имеют смысл и имеет место равенство*

$$(AB)C = A(BC).$$

Доказательство. По условию число столбцов матрицы A равно числу строчек матрицы B, а число столбцов матрицы B равно числу строчек матрицы C. Следовательно,

$$A = {}_m\overset{n}{\boxed{}}, \quad B = {}_n\overset{p}{\boxed{}}, \quad C = {}_p\overset{q}{\boxed{}}$$

(эта схематическая запись означает, например, что A имеет m строчек и n столбцов). Тогда

$$AB = G = {}_m\overset{p}{\boxed{}}, \quad BC = H = {}_n\overset{q}{\boxed{}}.$$

Отсюда легко усматриваем, что произведения $(AB)C$ и $A(BC)$ имеют смысл; при этом

$$(AB)C = GC = U = {}_m\overset{q}{\boxed{}}, \quad A(BC) = AH = V = {}_m\overset{q}{\boxed{}}.$$

Таким образом, матрицы $(AB)C = U$ и $A(BC) = V$ имеют одинаковые строения. Для доказательства теоремы надо теперь убедиться, что соответствующие элементы этих матриц равны.

Будем обозначать элементы всех встречающихся в доказательстве матриц соответствующими строчными буквами с двумя, как обычно, индексами внизу (первый индекс — номер строчки, второй — номер столбца). Тогда, согласно правилу умножения матриц, для любого $i = 1, 2, \ldots, m$ и любого $j = 1, 2, \ldots, q$ мы будем иметь

$$\begin{aligned} u_{ij} = g_{i1}c_{1j} + g_{i2}c_{2j} + \ldots + g_{ip}c_{pj} = \\ = (a_{i1}b_{11} + a_{i2}b_{21} + \ldots + a_{in}b_{n1})\,c_{1j} + \\ + (a_{i1}b_{12} + a_{i2}b_{22} + \ldots + a_{in}b_{n2})\,c_{2j} + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + (a_{i1}b_{1p} + a_{i2}b_{2p} + \ldots + a_{in}b_{np})\,c_{pj}, \end{aligned} \qquad (6)$$

$$\begin{aligned} v_{ij} = a_{i1}h_{1j} + a_{i2}h_{2j} + \ldots + a_{in}h_{nj} = \\ = a_{i1}(b_{11}c_{1j} + b_{12}c_{2j} + \ldots + b_{1p}c_{pj}) + \\ + a_{i2}(b_{21}c_{1j} + b_{22}c_{2j} + \ldots + b_{2p}c_{pj}) + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + a_{in}(b_{n1}c_{1j} + b_{n2}c_{2j} + \ldots + b_{np}c_{pj}). \end{aligned} \qquad (7)$$

Выражение (6) легко преобразуется в (7); для этого надо в (6) раскрыть скобки и привести подобные члены относительно $a_{i1}, a_{i2}, \ldots, a_{in}$. Следовательно, $u_{ij} = v_{ij}$, т. е. соответствующие элементы матриц U и V равны, а значит, и сами матрицы равны. Таким образом, $(AB)C = A(BC)$, и теорема 1 доказана.

Замечание 1. Вторую часть доказательства теоремы 1 можно записать короче, если воспользоваться знаком суммирования $\sum$ и использовать один простой факт, относящийся к двойным суммам. Пусть индексы i и j независимо друг от друга пробегают значения $i = 1, 2, \ldots, m$ и $j = 1, 2, \ldots, p$, и пусть каждой паре i, j этих индексов сопоставлено некоторое число d_{ij}.

Числа d_{ij} мы можем, конечно, расположить в виде прямоугольной матрицы D с m строчками и p столбцами. Просуммировав в матрице D элементы каждой строчки, мы получим выражения

$$\sigma_1 = \sum_{j=1}^{p} d_{1j}, \quad \sigma_2 = \sum_{j=1}^{p} d_{2j}, \ \ldots, \ \sigma_m = \sum_{j=1}^{p} d_{mj}.$$

Сумма σ всех чисел d_{ij} тогда примет вид

$$\sigma = \sum_{i=1}^{m} \sigma_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{p} d_{ij} \right).$$

С другой стороны, мы можем определить сначала сумму элементов каждого столбца матрицы D и сложить затем все полученные значения. Следовательно, сумму σ мы можем записать также в виде

$$\sigma = \sum_{j=1}^{p} \left(\sum_{i=1}^{m} d_{ij} \right).$$

Сравнивая оба результата, мы приходим к формуле

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{p} d_{ij} \right) = \sum_{j=1}^{p} \left(\sum_{i=1}^{m} d_{ij} \right), \tag{8}$$

означающей, что в двойной сумме можно изменить порядок суммирования.

Вернемся к доказательству теоремы 1. В силу отмеченного свойства двойных сумм равенство $u_{ij} = v_{ij}$ может быть установлено следующей выкладкой:

$$u_{ij} = \sum_{s=1}^{p} g_{is} c_{sj} = \sum_{s=1}^{p} \left(\sum_{r=1}^{n} a_{ir} b_{rs} \right) c_{sj} = \sum_{s=1}^{p} \left(\sum_{r=1}^{n} a_{ir} b_{rs} c_{sj} \right) =$$

$$= \sum_{r=1}^{n} \left(\sum_{s=1}^{p} a_{ir} b_{rs} c_{sj} \right) = \sum_{r=1}^{n} a_{ir} \left(\sum_{s=1}^{p} b_{rs} c_{sj} \right) = \sum_{r=1}^{n} a_{ir} h_{rj} = v_{ij}.$$

З а м е ч а н и е 2. Пусть последовательность матриц $A_1, A_2, \ldots, A_m$ такова, что произведение $A_i A_{i+1}$ любых двух соседних матриц имеет смысл. Тогда индуктивно может быть определено произведение $A_1 A_2 \ldots A_m$, а именно, считая, что для случая $m-1$ сомножителей произведение уже определено, полагаем

$$A_1A_2\ldots A_m = (A_1A_2\ldots A_{m-1})\,A_m.$$

Число строчек в этом произведении равно числу строчек первой матрицы A_1, а число столбцов равно числу столбцов последней матрицы A_m. Пользуясь теоремой 1, индукцией по числу сомножителей m легко можно показать, что для любого $i=1, 2, \ldots, m-1$ имеет место равенство

$$A_1A_2\ldots A_m = (A_1A_2\ldots A_i)(A_{i+1}\ldots A_m). \tag{9}$$

В начале § 4 для случая квадратных матриц нами было введено понятие транспонированной матрицы. Перенесем это понятие на случай произвольных прямоугольных матриц.

Определение. *Пусть дана матрица*

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot\,\cdot\,\cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix}.$$

Матрица

$$A' = \begin{Vmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \cdot & \cdot & \cdot\,\cdot\,\cdot & \cdot \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{Vmatrix},$$

получающаяся из A заменой строчек на столбцы, называется транспонированной с A. Операция замены матрицы A на A' называется транспонированием.

При транспонировании, как видим, меняется строение матрицы (если $m \neq n$), а именно:

$$\text{если } A = {}_m\overset{n}{\boxed{}}\,, \quad \text{то } A' = {}_n\overset{m}{\boxed{}}\,.$$

Например, если

$$A = \begin{Vmatrix} a & b & c \\ a_1 & b_1 & c_1 \end{Vmatrix}, \qquad S = \begin{Vmatrix} x & y & z \end{Vmatrix},$$

то

$$A' = \begin{Vmatrix} a & a_1 \\ b & b_1 \\ c & c_1 \end{Vmatrix}, \quad S' = \begin{Vmatrix} x \\ y \\ z \end{Vmatrix}.$$

При транспонировании строчка превращается в столбец, а столбец — в строчку.

Если операцию транспонирования применить дважды, то мы вернемся, очевидно, к исходной матрице:

$$A'' = (A')' = A.$$

Связь между элементами матриц A и A' выражается формулой

$$a'_{ij} = a_{ji}, \tag{10}$$

где a'_{ij} обозначает элемент i-й строчки j-го столбца матрицы A'.

Теорема 2. *Пусть A и B — две матрицы. Если произведение AB имеет смысл, то*

$$(AB)' = B'A'.$$

Другими словами, чтобы транспонировать произведение матриц, надо транспонировать сомножители и записать их в обратном порядке.

Доказательство. Пусть

$$A = {}_{m}\overset{n}{\boxed{}}, \quad B = {}_{n}\overset{p}{\boxed{}}.$$

Тогда

$$B' = {}_{p}\overset{n}{\boxed{}}, \quad A' = {}_{n}\overset{m}{\boxed{}},$$

так что произведение $B'A'$ имеет смысл. Далее,

$$AB = G = {}_{m}\overset{p}{\boxed{}}, \quad (AB)' = G' = {}_{p}\overset{m}{\boxed{}},$$

$$B'A' = U = {}_{p}\overset{m}{\boxed{}}.$$

Матрицы $(AB)'$ и $B'A'$ имеют, таким образом, одно и то же строение. Для доказательства теоремы остается теперь проверить, что соответствующие элементы матриц G' и U совпадают.

Обозначим элементы всех встречающихся здесь матриц соответствующими строчными буквами с двумя индексами (из которых первый обозначает номер строчки, а второй — номер столбца). Тогда, согласно формуле (10), мы будем иметь

$$u_{ij} = \sum_{s=1}^{n} b'_{is}a'_{sj} = \sum_{s=1}^{n} a_{js}b_{si} = g_{ji} = g'_{ij}.$$

Таким образом, соответствующие элементы матриц U и G' равны между собой, а значит, $G' = U$, что и требовалось доказать.

Замечание. Теорема 2 легко может быть обобщена на случай произвольного числа сомножителей. Например, если AB и BC имеют смысл, то $(ABC)' = C'B'A'$.

Среди квадратных матриц n-го порядка особую роль играет матрица

$$\begin{Vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{Vmatrix},$$

у которой все элементы, расположенные на главной диагонали, равны единице, а все прочие равны нулю. Эта матрица называется *единичной матрицей n-го порядка* и обозначается через E_n или, если ясно, чему равен ее порядок, через E. Свое название матрица E получила благодаря следующему свойству.

Теорема 3. *Для любой прямоугольной матрицы A с m строчками и n столбцами мы имеем*

$$E_m A = A, \qquad AE_n = A.$$

Действительно, элемент i-й строчки j-го столбца матрицы $E_m A = U$ равен

$$u_{ij} = 0 \cdot a_{1j} + \dots + 1 \cdot a_{ij} + \dots + 0 \cdot a_{mj} = a_{ij},$$

а значит, $U = A$. Аналогичной проверкой устанавливается и второе равенство.

Для единичной матрицы $E = E_n$ соответствующее линейное преобразование переменных имеет вид

$$\left.\begin{aligned} x_1 &= y_1, \\ x_2 &= \quad y_2, \\ &\cdots\cdots \\ x_n &= \qquad y_n. \end{aligned}\right\} \tag{11}$$

Это преобразование оставляет, по существу, все переменные без изменений; меняется только их наименование. Преобразование (11) называется *тождественным преобразованием*.

Очевидно, что линейное преобразование с матрицей A не изменится, если до него или вслед за ним выполнено тождественное преобразование (меняется только наименование старых или соответственно новых переменных). Этим мы еще раз другим способом убедились в справедливости теоремы 3.

Две квадратные матрицы одного и того же порядка всегда могут быть перемножены, при этом в произведении получится опять-таки квадратная матрица того же порядка. В силу этого мы можем определить степени квадратных матриц.

Пусть A — произвольная квадратная матрица n-го порядка. По определению полагаем

$$A^0 = E, \qquad A^1 = A, \qquad A^2 = AA, \qquad A^3 = AAA, \ldots$$

(здесь E — единичная матрица n-го порядка). В силу ассоциативности умножения матриц произведение

$$\underbrace{AA\ldots A}_{m \text{ раз}}$$

не зависит от способа расстановки в нем скобок (см. замечание 2 к теореме 1), поэтому степень A^m определена однозначно.

Из определения легко следует (см. равенство (9)), что для любых целых неотрицательных k и l имеют место формулы

$$\begin{gathered} A^k A^l = A^{k+l}, \\ (A^k)^l = A^{kl}. \end{gathered} \tag{12}$$

Из первой формулы в качестве следствия получаем, что

$$A^k A^l = A^l A^k, \tag{13}$$

т. е. две степени одной и той же квадратной матрицы всегда перестановочны.

Упражнения

1. Перемножить матрицы

a) $\begin{Vmatrix} 1 & 2 & 3 \\ -2 & 3 & 1 \\ 4 & 2 & 4 \end{Vmatrix} \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & -2 & 3 \\ -2 & -1 & 5 & -2 \end{Vmatrix}$; b) $\begin{Vmatrix} 2 & 3 & 4 \end{Vmatrix} \begin{Vmatrix} -2 \\ 1 \\ 5 \end{Vmatrix}$;

c) $\begin{Vmatrix} 1 & 2 & -1 \end{Vmatrix} \begin{Vmatrix} 2 & 4 & -2 & 1 \\ 1 & 2 & 1 & 2 \\ 3 & 1 & -3 & 1 \end{Vmatrix}$; d) $\begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & -2 \end{Vmatrix} \begin{Vmatrix} 1 & 1 \\ -3 & 3 \\ 3 & -5 \\ -1 & 1 \end{Vmatrix}$.

Отв. a) $\begin{Vmatrix} -1 & 4 & 12 & 1 \\ 2 & 6 & -3 & 5 \\ 0 & 6 & 20 & 2 \end{Vmatrix}$; b) 19; c) $\begin{Vmatrix} 1 & 7 & 3 & 4 \end{Vmatrix}$; d) $\begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}$.

2. Найти x, y, z из равенства

$$\begin{Vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{Vmatrix} \begin{Vmatrix} x \\ y \\ z \end{Vmatrix} = \begin{Vmatrix} 4 \\ 1 \\ 3 \end{Vmatrix}.$$

Отв. $x=2$, $y=-1$, $z=1$.

3. Найти двумя способами произведение

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

Отв. $\begin{Vmatrix} 16 & 80 & 54 \\ 17 & 8 & 15 \end{Vmatrix}$.

4. Найти все матрицы, перестановочные с матрицей $\begin{Vmatrix} 3 & 5 \\ 1 & 2 \end{Vmatrix}$.

Отв.

$$\begin{Vmatrix} \alpha+\beta & 5\alpha \\ \alpha & \beta \end{Vmatrix},$$

где α и β — произвольные числа.

5. Для произвольной матрицы A обозначим через $\bar{A}$ матрицу, получающуюся из A заменой всех ее элементов комплексно сопряженными числами.

Доказать, что если $AB = C$, то $\bar{A}\bar{B} = \bar{C}$.

6. Квадратная матрица n-порядка называется симметричной, если она совпадает со своей транспонированной матрицей. Доказать, что если две симметричные матрицы перестановочны, то их произведение есть симметричная матрица.

7. Найти степени

$$\text{a)} \begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix}^n; \quad \text{b)} \begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix}^n.$$

Отв. a) $\begin{Vmatrix} 1 & n \\ 0 & 1 \end{Vmatrix}$; b) $\begin{Vmatrix} \cos n\varphi & -\sin n\varphi \\ \sin n\varphi & \cos n\varphi \end{Vmatrix}$.

8. Пусть A — квадратная матрица n-го порядка, у которой все элементы, расположенные на главной диагонали и ниже ее, равны нулю. Доказать, что тогда A^n есть нулевая матрица (т. е. состоит сплошь из нулей).

9. Найти n-ю степень матрицы

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

Отв. A.

10. Пусть A и B — две квадратные матрицы. Доказать, что сумма элементов, стоящих на главной диагонали, для матриц AB и BA одна и та же.

§ 14. Сложение матриц и умножение матрицы на число

Определение. *Пусть даны две матрицы A и B одинакового строения. Их суммой называется матрица $C = A + B$ того же строения, элементы которой равны суммам соответствующих элементов матриц A и B.*

Элементы c_{ij} матрицы $C = A + B$ определяются, стало быть, формулами

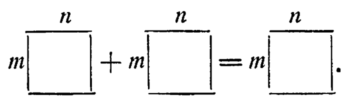
$$c_{ij} = a_{ij} + b_{ij},$$

где a_{ij} и b_{ij} — элементы матриц A и B соответственно.

Например,

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

Подчеркнем еще раз, что складывать можно только матрицы с одинаковым числом строчек и с одинаковым числом столбцов. Выразим этот факт схематически.



Очевидно, что сложение матриц подчиняется переместительному (коммутативному) и сочетательному (ассоциативному) законам, т. е.

$$A+B=B+A, \tag{1}$$

$$(A+B)+C=A+(B+C). \tag{2}$$

Действие сложения матриц может быть распространено, разумеется, и на случай любого числа слагаемых.

Вычитание для матриц (как и для чисел) определяется как действие, обратное сложению. Разностью $B-A$ матриц B и A (одинакового строения) называется такая матрица X, что

$$A+X=B.$$

Легко видеть, что матрица X, удовлетворяющая этому условию, всегда существует, и притом только одна. Ее элементы x_{ij} определяются равенствами

$$x_{ij}=b_{ij}-a_{ij}.$$

Таким образом, при вычитании матриц вычитаются соответствующие элементы этих матриц. Например,

$$\begin{Vmatrix} 1 & 7 & 5 \\ 2 & 9 & -7 \end{Vmatrix} - \begin{Vmatrix} 3 & 3 & 3 \\ 5 & -2 & -9 \end{Vmatrix} = \begin{Vmatrix} -2 & 4 & 2 \\ -3 & 11 & 2 \end{Vmatrix}.$$

Матрица O, все элементы которой равны нулю, называется *нулевой матрицей*.

Очевидно, что

$$A+O=A, \qquad A-A=O.$$

Разность $O-A$ обозначается через $-A$, так что $A+(-A)=O$. Например, если

$$A=\begin{Vmatrix} 2 & 3 \\ -5 & 1 \end{Vmatrix}, \text{ то } -A=\begin{Vmatrix} -2 & -3 \\ 5 & -1 \end{Vmatrix}.$$

Матрица $-A$ называется *противоположной матрице* A.

Теорема 1. *Сложение матриц связано с умножением распределительными (дистрибутивными) законами*:

$$(A+B)C=AC+BC,$$
$$D(A+B)=DA+DB$$

(*при этом предполагается, что выражение слева или справа имеет смысл*).

Доказательство. Предположим, что левая часть первого равенства имеет смысл. Тогда A и B имеют одинаковое строение. Пусть они имеют по m строчек и по n столбцов. Такое же строение имеет и матрица $A+B=H$. Так как HC имеет смысл, то C состоит из n строчек и, скажем, из p столбцов. Произведение $HC=X$ имеет m строчек и p столбцов. Обратимся теперь к правой части. Так как число столбцов в матрицах A и B совпадает с числом строчек в C, то существуют произведения $AC=U$ и $BC=V$; при этом оба они имеют m строчек и p столбцов. Следовательно, их сумма $U+V=Y$ имеет смысл и также имеет m строчек и p столбцов. Установлено, таким образом, что матрицы $X=(A+B)C$ и $Y=$ $=AC+BC$ имеют одинаковое строение. Сравним соответствующие элементы этих матриц.

Обозначим, как и прежде, элементы матриц соответствующими строчными буквами. Тогда для любого $i=1, 2, \ldots, m$ и любого $j=1, 2, \ldots, p$ мы будем иметь

$$x_{ij}=\sum_{s=1}^{n} h_{is}c_{sj}=\sum_{s=1}^{n}(a_{is}+b_{is})c_{sj}=\sum_{s=1}^{n}(a_{is}c_{sj}+b_{is}c_{sj})=$$
$$=\sum_{s=1}^{n} a_{is}c_{sj}+\sum_{s=1}^{n} b_{is}c_{sj}=u_{ij}+v_{ij}=y_{ij}.$$

Таким образом, $X=Y$ и первое равенство теоремы 1 доказано (мы предполагали, что левая часть имеет смысл; можно было исходить из предположения, что правая часть имеет смысл).

Второе равенство доказывается аналогично.

Замечание. Доказанный нами дистрибутивный закон умножения может быть распространен, конечно, на случай произвольного числа слагаемых, а также на случай разности. Например,

$$A(B_1+B_2+\ldots+B_m)=AB_1+AB_2+\ldots+AB_m,$$
$$(A_1+A_2)(B_1+B_2+B_3)=$$
$$=A_1B_1+A_1B_2+A_1B_3+A_2B_1+A_2B_2+A_2B_3,$$
$$(A-B)C=AC-BC.$$

Определение. *Произведением матрицы A на число α называется матрица, получающаяся из A умножением всех ее элементов на α.*

Обозначается это произведение через αA или $A\alpha$.

Пример.

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

При умножении матрицы на число ее строение, как видим, сохраняется.

Очевидно, что

$$1\cdot A=A, \qquad (-1)A=-A, \qquad 0\cdot A=O.$$

Легко проверяются также следующие свойства:

$$\alpha(\beta A)=(\alpha\beta)A, \tag{3}$$
$$(\alpha+\beta)A=\alpha A+\beta A, \tag{4}$$
$$\alpha(A+B)=\alpha A+\alpha B. \tag{5}$$

Теорема 2. *Если произведение AB имеет смысл, то*

$$\alpha(AB)=(\alpha A)B=A(\alpha B).$$

Другими словами, чтобы умножить произведение матриц на число, достаточно один из сомножителей умножить на это число.

Доказательство. Пусть матрица A имеет m строчек и n столбцов, а матрица B — n строчек и p столбцов. Ясно, что все три произведения имеют одно и то же строение. Далее, для любого $i=1, 2, \ldots, m$ и любого $j=1, 2, \ldots, p$ мы имеем

$$\alpha\sum_{s=1}^{n} a_{is}b_{sj}=\sum_{s=1}^{n}(\alpha a_{is})b_{sj}=\sum_{s=1}^{n} a_{is}(\alpha b_{sj}).$$

Следовательно, элементы, стоящие на одинаковых местах, у всех трех матриц совпадают, а значит, сами эти матрицы равны.

Квадратная матрица вида

$$\begin{Vmatrix} \alpha & 0 \dots 0 \\ 0 & \alpha \dots 0 \\ \cdot & \cdot \; \cdot \; \cdot \; \cdot \\ 0 & 0 \dots \alpha \end{Vmatrix},$$

у которой элементы, стоящие на главной диагонали, равны α, а все прочие равны нулю, называется *скалярной матрицей.*

Пользуясь правилом умножения матрицы на число, скалярную матрицу можно представить в виде

$$\alpha E,$$

где E — единичная матрица соответствующего порядка. Так как

$$\alpha A = \alpha (EA) = (\alpha E) A,$$

$$\alpha A = \alpha (AE) = A (\alpha E),$$

то, следовательно, умножение матрицы на число равносильно умножению этой матрицы слева или справа на соответствующую скалярную матрицу. Например,

$$\begin{Vmatrix} \alpha & 0 \\ 0 & \alpha \end{Vmatrix} \begin{Vmatrix} a & b & c \\ x & y & z \end{Vmatrix} = \begin{Vmatrix} \alpha a & \alpha b & \alpha c \\ \alpha x & \alpha y & \alpha z \end{Vmatrix} =$$

$$= \alpha \begin{Vmatrix} a & b & c \\ x & y & z \end{Vmatrix} = \begin{Vmatrix} a & b & c \\ x & y & z \end{Vmatrix} \begin{Vmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{Vmatrix}.$$

Пусть нам заданы матрицы

$$A_1, \; A_2, \dots, \; A_k$$

одного и того же строения. Тогда для любой системы чисел $\alpha_1, \; \alpha_2, \dots, \; \alpha_k$ имеет смысл выражение

$$\alpha_1 A_1 + \alpha_2 A_2 + \dots + \alpha_k A_k, \qquad (6)$$

являющееся, очевидно, матрицей того же строения. Это выражение называется *линейной комбинацией* матриц $A_1, A_2, \dots, \; A_k$ с коэффициентами $\alpha_1, \alpha_2, \dots, \; \alpha_k$.

Из свойств (3) — (5) следует, что сложение двух линейных комбинаций вида (6) при фиксированных матрицах $A_1, \; A_2, \dots, \; A_k$ сводится к сложению соответствующих коэффициентов (приведению подобных членов), а умножение линейной комбинации (6) на число α сводится к умножению всех ее коэффициентов на α.

В § 13 мы показали, как линейное преобразование (1) из § 13 может быть записано в матричной форме. Точно так же, разумеется, мы можем записать в матричной форме и любую систему линейных уравнений. Если матрицу системы

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \cdots\cdots\cdots\cdots\cdots&\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}\right\} \qquad (7)$$

мы обозначим через A, столбец из неизвестных x_1, $x_2, \ldots, x_n$ — через X и столбец из свободных членов b_1, $b_2, \ldots, b_m$ — через R, то эта система примет следующую матричную форму:

$$AX = R. \qquad (8)$$

Если $\alpha_1, \alpha_2, \ldots, \alpha_n$ — какое-нибудь решение системы (7), то составленный из этих чисел столбец будет удовлетворять уравнению (8) (с неизвестным столбцом X). Наоборот, всякий столбец, удовлетворяющий уравнению (8), дает нам систему чисел, составляющих решение системы (7). Таким образом, для решения системы (7) в матричной интерпретации надо найти все столбцы, удовлетворяющие уравнению (8). Существование таких столбцов равносильно совместности системы (7).

Если в системе (7) все свободные члены мы заменим нулями, то получим однородную систему. В соответствии с изложенным эта однородная система может быть записана в виде матричного уравнения

$$AX = O, \qquad (9)$$

где O — нулевой столбец. Так как для столбцов (являющихся прямоугольными матрицами частного вида) определено сложение и умножение на числа, то можно говорить о сумме решений уравнения (9), о произведении решения на число, а также о линейной комбинации решений.

Теорема 3. *Произвольная линейная комбинация решений уравнения* (9) *является также решением этого* *уравнения. В частности, сумма двух решений и произведение решения на число являются решениями.*

Доказательство. Достаточно ограничиться рассмотрением линейной комбинации двух решений. Пусть столбцы U_1 и U_2 удовлетворяют уравнению (9), и пусть α и β — произвольные числа. Так как $AU_1 = O$ и $AU_2 = O$, то

$$A(\alpha U_1 + \beta U_2) = A(\alpha U_1) + A(\beta U_2) = \alpha(AU_1) + \beta(AU_2) = \\ = \alpha O + \beta O = O.$$

Таким образом, вместе с U_1 и U_2 столбец $\alpha U_1 + \beta U_2$ также удовлетворяет уравнению (9).

Теорема 4. *Пусть T_0 — какое-нибудь фиксированное решение уравнения* (8) (*мы предполагаем, стало быть, систему* (7) *совместной*). *Тогда всякое решение уравнения* (8) *имеет вид*

$$T_0 + U, \tag{10}$$

где U — некоторое решение соответствующего однородного уравнения (9). *Наоборот, сумма* (10) *при любом решении U уравнения* (9) *является решением уравнения* (8).

Доказательство. Пусть столбец T удовлетворяет уравнению (8), так что $AT = R$. Положим $U = T - T_0$. Тогда $AU = A(T - T_0) = AT - AT_0 = R - R = O$. Таким образом, $T = T_0 + U$, причем $AU = O$. Наоборот, если $AU = O$, то $A(T_0 + U) = AT_0 + AU = R + O = R$. Теорема 4 доказана.

В заключение параграфа введем понятие многочлена от матрицы. Пусть A — квадратная матрица n-го порядка, и пусть

$$f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_m x^m$$

— многочлен от переменной x. Выражение

$$\alpha_0 E + \alpha_1 A + \alpha_2 A^2 + \ldots + \alpha_m A^m$$

называется *многочленом от матрицы* A и обозначается через $f(A)$. Здесь $E = A^0$ — единичная матрица n-го порядка. Ясно, что $f(A)$ является также квадратной матрицей n-го порядка.

Пример. Пусть

$$A=\begin{Vmatrix} 1 & 2 \\ -1 & 1 \end{Vmatrix}, \quad f(x)=x^2-5x+7, \quad g(x)=x^2-2x+3.$$

Тогда

$$f(A)=\begin{Vmatrix} -1 & 4 \\ -2 & -1 \end{Vmatrix}-5\begin{Vmatrix} 1 & 2 \\ -1 & 1 \end{Vmatrix}+7\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}=\begin{Vmatrix} 1 & -6 \\ 3 & 1 \end{Vmatrix},$$

$$g(A)=\begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}.$$

Теорема 5. *Пусть $f(x)$ и $g(x)$ — два многочлена и A — квадратная матрица. Тогда, если*

$$f(x)+g(x)=\varphi(x),$$
$$f(x)g(x)=\psi(x),$$

то

$$f(A)+g(A)=\varphi(A),$$
$$f(A)g(A)=\psi(A).$$

Доказательство. Вспомним правила сложения и умножения многочленов. Сложение многочленов сводится к выполнению действий вида

$$\alpha x^k+\beta x^k=(\alpha+\beta)x^k$$

(приведение подобных членов). Умножение многочленов происходит на основе распределительного закона: сначала перемножают одночлены по правилу

$$\alpha x^k\cdot\beta x^l=\alpha\beta x^{k+l},$$

а затем приводят подобные члены. Но в силу свойства (4) мы имеем

$$\alpha A^k+\beta A^k=(\alpha+\beta)A^k.$$

Далее, согласно теореме 2 и формуле (12) из § 13 мы имеем также

$$\alpha A^k\cdot\beta A^l=\alpha\beta(A^kA^l)=\alpha\beta A^{k+l}.$$

Мы видим, таким образом, что матрица A ведет себя по отношению ко всем рассматриваемым операциям точно так же, как и переменная x. Следовательно, если мы будем, например, перемножать многочлены $f(A)$ и

$g(A)$ по правилам умножения многочленов, то в результате получим многочлен от матрицы A, коэффициенты которого будут такие же, как и у $\psi(x)$, т. е. получим $\psi(A)$.

Следствие. *Два многочлена от одной и той же матрицы всегда перестановочны, т. е.*

$$f(A)\,g(A)=g(A)\,f(A). \tag{11}$$

Действительно, обе части равенства равны $\psi(A)$. (Равенство (11) является обобщением формулы (13) из § 13.)

Пример. Так как

$$x^3-1=(x-1)(x^2+x+1),$$

то

$$A^3-E=(A-E)(A^2+A+E)=(A^2+A+E)(A-E).$$

Упражнения

1. Представить строчку $S=\|10\ \ -5\ \ -3\|$ в виде линейной комбинации строчек

$$S_1=\|1\ \ 2\ \ -1\|,\quad S_2=\|4\ \ 3\ \ 5\|,\quad S_3=\|3\ \ -1\ \ 2\|.$$

Отв. $S=3S_1-2S_2+5S_3$.

2. Показать, что формула $A^2-B^2=(A-B)(A+B)$ имеет место тогда и только тогда, когда матрицы A и B перестановочны.

3. Доказать, что если $A^2=A$, то матрица $B=2A-E$ удовлетворяет условию $B^2=E$.

4. Найти $f(A)$ и $g(A)$, если $f(x)=x^3-7x+9$, $g(x)=x^3-3x^2+3x-1$,

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

Отв. $f(A)=\begin{Vmatrix}3 & 4 & -5\\ 0 & 3 & 4\\ 0 & 0 & 3\end{Vmatrix},\quad g(A)=\begin{Vmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{Vmatrix}.$

§ 15. Определитель произведения квадратных матриц

Если A и B — две квадратные матрицы n-го порядка, то их произведение AB также является квадратной матрицей n-го порядка. Как связаны между собой определители $|A|$, $|B|$ и $|AB|$ этих трех матриц?

Если мы возьмем, например,

$$A=\begin{Vmatrix} 3 & 2 \\ 1 & 2 \end{Vmatrix}, \qquad B=\begin{Vmatrix} 2 & 1 \\ 1 & 3 \end{Vmatrix},$$

то

$$|A|=4, \qquad |B|=5, \qquad |AB|=\begin{vmatrix} 8 & 9 \\ 4 & 7 \end{vmatrix}=20.$$

Бросающаяся в этом примере в глаза закономерность выражается следующей теоремой.

Теорема. *Определитель произведения двух квадратных матриц n-го порядка равен произведению определителей перемножаемых матриц*:

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

Доказательство. Рассмотрим следующий определитель порядка $2n$:

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix}=\begin{vmatrix} A & O \\ -E & B \end{vmatrix}.$$

Так как матрица, стоящая здесь под знаком определителя, есть ступенчатая матрица с диагональными клетками A и B, то, по теореме 1 из § 6,

$$\Delta=|A|\cdot|B|. \tag{1}$$

Вычислим определитель Δ другим способом. Для этого преобразуем его так, чтобы все элементы матрицы B (в правом нижнем углу) превратились в нули. Это легко сделать с помощью диагональных элементов левой нижней клетки $-E$. Чтобы, например, избавиться от b_{11}, следует к $(n+1)$-му столбцу прибавить первый, умноженный на b_{11}. В общем же случае, чтобы удалить b_{kj}, надо к столбцу с номером $n+j$ прибавить k-й столбец, умноженный на b_{kj}. В результате этих n^2 операций левая часть определителя останется без изменений, правый нижний угол займут нули, а на месте нулей в правом верхнем углу окажутся некоторые новые элементы. Так как при наших преобразованиях величина определителя не меняется (свойство 9 на § 4), то мы получим

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & c_{11} & c_{12} & \dots & c_{1n}\\ a_{21} & a_{22} & \dots & a_{2n} & c_{21} & c_{22} & \dots & c_{2n}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ a_{n1} & a_{n2} & \dots & a_{nn} & c_{n1} & c_{n2} & \dots & c_{nn}\\ -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0\\ 0 & -1 & \dots & 0 & 0 & 0 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & -1 & 0 & 0 & \dots & 0 \end{vmatrix}.$$

Обозначим квадратную матрицу n-го порядка, появившуюся здесь в правом верхнем углу, через C.

Чтобы преобразовать полученный определитель в ступенчатый, мы поменяем местами строчки с номерами $n+1$, $n+2, \dots, 2n$ соответственно с 1-й, 2-й, ..., n-й строчкой. В результате получим

$$\Delta=(-1)^n\begin{vmatrix} -E & O\\ A & C \end{vmatrix}=(-1)^n|-E|\cdot|C|=$$
$$=(-1)^n\cdot(-1)^n|C|.$$

Таким образом, для Δ мы нашли другое выражение:

$$\Delta=|C|. \qquad (2)$$

Разберемся теперь, что же представляет собой матрица C. Перепишем наш исходный определитель Δ еще раз, выделив в нем лишь интересующие нас места:

$$\Delta=\begin{array}{cc} \begin{matrix} & & & & & (j) & \end{matrix} & \\ \begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \vdots & \\ a_{i1} & a_{i2} & \dots & a_{in} & \dots & \dots & \dots\\ \cdot & \cdot & \cdot & \cdot & & \vdots & \\ -1 & 0 & \dots & 0 & \dots & b_{1j} & \dots\\ 0 & -1 & \dots & 0 & \dots & b_{2j} & \dots\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & -1 & \dots & b_{nj} & \dots \end{vmatrix} & \begin{matrix} \\ (i) \\ \\ \\ \\ \\ \\ \end{matrix} \end{array}$$

До наших первоначальных преобразований на пересечении пунктирных линий в правом верхнем углу стоял нуль. При выполнении преобразований к этому нулю мы прибавили произведения

$$a_{i1}b_{1j},\ a_{i2}b_{2j}, \ldots,\ a_{in}b_{nj}.$$

После всех преобразований там должен оказаться элемент c_{ij}. Следовательно,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Таким образом, элементы матрицы C составляются по тому же правилу, что и соответствующие элементы произведения AB; значит, $C = AB$. Сопоставляя этот факт с равенствами (1) и (2), мы и получаем утверждение теоремы.

З а м е ч а н и е. Теорема справедлива, конечно, и для случая любого числа сомножителей.

Доказанная теорема в некоторых случаях может оказаться полезной при вычислении определителей. Пусть, например, требуется найти определитель матрицы

$$A = \begin{Vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{Vmatrix}.$$

Составим транспонированную с ней матрицу

$$A' = \begin{Vmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{Vmatrix}$$

и найдем произведение AA':

$$AA' = \begin{Vmatrix} a^2+b^2+c^2+d^2 & 0 & 0 & 0 \\ 0 & a^2+b^2+c^2+d^2 & 0 & 0 \\ 0 & 0 & a^2+b^2+c^2+d^2 & 0 \\ 0 & 0 & 0 & a^2+b^2+c^2+d^2 \end{Vmatrix}.$$

Имеем $|AA'| = |A|\,|A'| = |A|^2 = (a^2+b^2+c^2+d^2)^4$, поэтому

$$|A| = \pm(a^2+b^2+c^2+d^2)^2.$$

Чтобы установить, какой здесь нужно взять знак — плюс или минус, обращаем внимание на тот факт, что слагаемое a^4 входит в определитель со знаком плюс как произведение элементов, стоящих на главной диагонали. Следовательно,

$$|A| = (a^2+b^2+c^2+d^2)^2.$$

Упражнение

Вычислить определитель матрицы

$$\begin{Vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{Vmatrix},$$

умножив ее предварительно на матрицу

$$\begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{Vmatrix}.$$

Отв. $(a+b+c+d)(a+b-c-d)(a-b+c-d)\cdot$
$\cdot(a-b-c+d)$.

§ 16. Обратная матрица

Как известно, для каждого числа $a \neq 0$ существует такое b, что $ab = 1$. Число b называется обратным для a. Если мы зафиксируем натуральное число n и будем рассматривать квадратные матрицы n-го порядка, то в этом множестве матриц единичная матрица E будет играть роль единицы (теорема 3 из § 13). Естественно поставить вопрос о существовании обратной матрицы, т. е. такой матрицы, которая в произведении с данной дает единичную матрицу E. Этому вопросу мы и посвятим настоящий параграф.

Определение. *Пусть A — квадратная матрица n-го порядка. Квадратная матрица X (того же порядка n) называется обратной для A, если*

$$AX = XA = E.$$

По поводу данного определения следует сделать следующее пояснение. Поскольку умножение матриц, вообще говоря, некоммутативно, мы должны учитывать возможность того, что $AX \neq XA$. В нашем определении на матрицу X наложено два условия: $AX=E$ и $XA=E$. Ввиду этого матрицу X можно было бы назвать *двусторонней обратной* для A. Если же о некоторых матрицах Y и Z нам известно лишь, что $AY=E$ и $ZA=E$, то Y естественно назвать *правой обратной*, а Z — *левой обратной* для A. В дальнейшем, однако, мы увидим, что это разнообразие обратных матриц фактически отсутствует: либо обратной матрицы вообще нет, либо она только одна (и притом двусторонняя).

Предположим, что для матрицы A существует правая обратная матрица Y, так что

$$AY=E.$$

Перейдем в этом равенстве к определителям. Так как $|E|=1$ и $|AY|=|A|\cdot|Y|$ по теореме § 15, то

$$|A|\cdot|Y|=1.$$

Отсюда заключаем, что $|A|\neq 0$ (в противном случае левая часть равнялась бы нулю). Тот же вывод, конечно, мы получим, если предположим, что для A существует левая обратная матрица.

Этим доказано, что если $|A|=0$, то для A не существует ни левой, ни правой обратной матрицы.

Определение. *Квадратная матрица A n-го порядка называется особенной, если ее определитель равен нулю. Если же $|A|\neq 0$, то A называется неособенной матрицей.*

Особенные матрицы, как доказано нами, обратных не имеют. Займемся изучением неособенных матриц.

Пусть нам дана неособенная матрица

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

определитель которой мы обозначим через Δ. Рассмотрим соответствующее ей линейное преобразование переменных

$$\left.\begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n, \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n, \\ &\dots\dots\dots\dots\dots\dots \\ x_n &= a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n. \end{aligned}\right\} \tag{2}$$

Предположим, что для матрицы A существует правая обратная. Обозначим ее через B и рассмотрим соответствующее ей линейное преобразование переменных

$$y_i = \sum_{j=1}^{n} b_{ij} z_j \qquad (i = 1, 2, \ldots, n). \tag{3}$$

Так как $AB = E$, то произведение линейных преобразований (2) и (3) является тождественным преобразованием:

$$x_i = z_i \qquad (i = 1, 2, \ldots, n).$$

Следовательно, если в (3) вместо z_j мы подставим x_j, а затем полученные выражения

$$y_i = \sum_{j=1}^{n} b_{ij} x_j \tag{4}$$

подставим в (2), то после выполнения всех необходимых действий все равенства (2) превратятся в тождества вида $x_i = x_i$ $(i = 1, 2, \ldots, n)$. Но это значит, что значения (4) составляют решение системы линейных уравнений (2) с неизвестными $y_1, y_2, \ldots, y_n$ и со свободными членами $x_1, x_2, \ldots, x_n$.

Таким образом, если линейное преобразование (2) мы рассмотрим как систему линейных уравнений относительно неизвестных $y_1, y_2, \ldots, y_n$ и решим ее, то, поскольку решение должно иметь вид (4), мы найдем элементы матрицы B.

По условию $|A| = \Delta \neq 0$, поэтому к системе (2) применимы формулы Крамера:

$$y_1 = \frac{\Delta_1}{\Delta}, \quad y_2 = \frac{\Delta_2}{\Delta}, \ldots, \quad y_n = \frac{\Delta_n}{\Delta},$$

где Δ_j — определитель матрицы, получающейся из A заменой j-го столбца на свободные члены.

Обозначим, как обычно, через A_{ij} алгебраическое дополнение элемента a_{ij} матрицы A. Тогда, раскладывая каждый из определителей Δ_j по элементам j-го столбца (где стоят переменные $x_1, x_2, \ldots, x_n$), мы получим

$$\left.\begin{aligned} y_1 &= \frac{A_{11}}{\Delta} x_1 + \frac{A_{21}}{\Delta} x_2 + \ldots + \frac{A_{n1}}{\Delta} x_n, \\ y_2 &= \frac{A_{12}}{\Delta} x_1 + \frac{A_{22}}{\Delta} x_2 + \ldots + \frac{A_{n2}}{\Delta} x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n &= \frac{A_{1n}}{\Delta} x_1 + \frac{A_{2n}}{\Delta} x_2 + \ldots + \frac{A_{nn}}{\Delta} x_n. \end{aligned}\right\} \qquad (5)$$

Так как выражения (4) и (5) должны совпадать, то фактически мы нашли вид матрицы B (в предположении ее существования).

После этих предварительных рассуждений исследуем теперь вопрос об обратных матрицах, не опираясь на линейные преобразования переменных.

Определение. *Пусть дана квадратная матрица* (1), *и пусть A_{ij} обозначает алгебраическое дополнение элемента a_{ij}. Матрица*

$$\tilde{A} = \begin{Vmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{Vmatrix},$$

в строчках которой расположены алгебраические дополнения элементов соответствующих столбцов матрицы A, называется взаимной для A.

Взаимная матрица $\tilde{A}$ получается, таким образом, из матрицы A заменой каждого элемента на его алгебраическое дополнение с последующим транспонированием. Ясно, что понятие взаимной матрицы имеет смысл для любой квадратной матрицы (как неособенной, так и особенной).

Теорема 1. *Для взаимной матрицы $\tilde{A}$ квадратной матрицы A n-го порядка мы имеем*

$$A\tilde{A} = \tilde{A}A = \Delta E,$$

где $\Delta = |A|$ и E — единичная матрица n-го порядка.

Доказательство. Вычислим произведение $A\tilde{A}$:

$$C = A\tilde{A} = \begin{Vmatrix} \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \end{Vmatrix} \begin{Vmatrix} \ldots & A_{j1} & \ldots \\ \ldots & A_{j2} & \ldots \\ \cdot & \cdot & \cdot \\ \ldots & A_{jn} & \ldots \end{Vmatrix}.$$

(Напомним, что в j-м столбце матрицы $\tilde{A}$ стоят алгебраические дополнения элементов j-й строчки матрицы A.) Согласно правилу умножения матриц, для элементов c_{ij} матрицы C имеем

$$c_{ij}=a_{i1}A_{j1}+a_{i2}A_{j2}+\ldots+a_{in}A_{jn}. \qquad (6)$$

Но согласно свойствам 4 и 5 из § 4 правая часть равенства (6) равна Δ при $i=j$ и равна нулю при $i\neq j$. Следовательно,

$$c_{ii}=\Delta, \quad c_{ij}=0 \quad \text{при } i\neq j.$$

Таким образом,

$$A\tilde{A}=C=\begin{Vmatrix}\Delta & 0 & \ldots & 0\\ 0 & \Delta & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & \Delta\end{Vmatrix}=\Delta\begin{Vmatrix}1 & 0 & \ldots & 0\\ 0 & 1 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & 1\end{Vmatrix}=\Delta E.$$

Аналогично проверяется и второе равенство $\tilde{A}A=\Delta E$. Теорема 1 доказана.

Предположим опять, что A — неособенная матрица, так что $|A|=\Delta\neq 0$. Положим

$$X=\frac{1}{\Delta}\tilde{A}. \qquad (7)$$

Тогда

$$AX=A\left(\frac{1}{\Delta}\tilde{A}\right)=\frac{1}{\Delta}(A\tilde{A})=\frac{1}{\Delta}\cdot\Delta E=E$$

и аналогично

$$XA=E.$$

Таким образом, матрица (7) является двусторонней обратной для A. Этим мы установили, что всякая неособенная матрица имеет обратную.

Пусть теперь Y — какая-нибудь правая (в частности, двусторонняя) обратная для A, так что $AY=E$. Умножив это равенство слева на матрицу (7), получим

$$X(AY)=XE, \qquad (XA)Y=X, \qquad EY=X, \qquad Y=X.$$

Аналогично из $ZA=E$ также следует, что $Z=X$. Таким образом, для неособенной матрицы существуют только одна правая обратная и только одна левая обратная; обе они совпадают между собой и совпадают с единственной двусторонне обратной матрицей (7).

Объединим полученные результаты об обратной матрице.

Теорема 2. *Особенные матрицы обратных матриц не имеют. Всякая неособенная матрица A имеет единственную обратную матрицу, обозначаемую через A^{-1}. Если $|A| = \Delta \neq 0$, то обратная матрица для A имеет вид*

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

Пример. Найти обратную матрицу для

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

Так как определитель $|A| = \Delta = 3$ отличен от нуля, то обратная матрица A^{-1} существует. Вычисляем миноры всех элементов матрицы A:

$$\Delta_{11} = \begin{vmatrix} 3 & -6 \\ -2 & 3 \end{vmatrix} = -3, \quad \Delta_{12} = \begin{vmatrix} 5 & -6 \\ -1 & 3 \end{vmatrix} = 9,$$

$$\Delta_{21} = \begin{vmatrix} -1 & 0 \\ -2 & 3 \end{vmatrix} = -3, \quad \Delta_{22} = \begin{vmatrix} 2 & 0 \\ -1 & 3 \end{vmatrix} = 6,$$

$$\Delta_{31} = \begin{vmatrix} -1 & 0 \\ 3 & -6 \end{vmatrix} = 6 \quad \Delta_{32} = \begin{vmatrix} 2 & 0 \\ 5 & -6 \end{vmatrix} = -12,$$

$$\Delta_{13} = \begin{vmatrix} 5 & 3 \\ -1 & -2 \end{vmatrix} = -7,$$

$$\Delta_{23} = \begin{vmatrix} 2 & -1 \\ -1 & -2 \end{vmatrix} = -5,$$

$$\Delta_{33} = \begin{vmatrix} 2 & -1 \\ 5 & 3 \end{vmatrix} = 11.$$

Следовательно, взаимная матрица для A

$$\tilde{A} = \begin{Vmatrix} -3 & 3 & 6 \\ -9 & 6 & 12 \\ -7 & 5 & 11 \end{Vmatrix},$$

а значит,

$$A^{-1}=\frac{1}{3}\tilde{A}=\begin{Vmatrix} -1 & 1 & 2\\ -3 & 2 & 4\\ -\frac{7}{3} & \frac{5}{3} & \frac{11}{3}\end{Vmatrix}.$$

Теорема 3. *Произведение AB неособенных матриц A и B является неособенной матрицей, при этом*

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

т. е. обратная матрица произведения двух матриц равна произведению их обратных матриц в обратном порядке. Обратная матрица A^{-1} для неособенной матрицы A также неособенна, и

$$(A^{-1})^{-1}=A.$$

Доказательство. Если $|A|\neq 0$ и $|B|\neq 0$, то, согласно теореме § 15, определитель $|AB|=|A|\cdot|B|$ также отличен от нуля, так что AB — неособенная матрица. Проверим, что матрица $B^{-1}A^{-1}$ удовлетворяет матричному уравнению $(AB)X=E$. Действительно,

$$(AB)(B^{-1}A^{-1})=ABB^{-1}A^{-1}=A(BB^{-1})A^{-1}=$$
$$=AEA^{-1}=AA^{-1}=E.$$

Таким образом, произведение $B^{-1}A^{-1}$ является правой обратной матрицей для AB, а значит (в силу единственности), оно совпадает с $(AB)^{-1}$. Этим доказано первое утверждение теоремы. Далее, если $|A|\neq 0$ и, значит, A^{-1} существует, то, переходя в равенстве $AA^{-1}=E$ к определителям, получим $|AA^{-1}|=|A||A^{-1}|=1$, откуда следует, что $|A^{-1}|\neq 0$; при этом

$$|A^{-1}|=\frac{1}{|A|}=|A|^{-1}. \tag{8}$$

Матрица A удовлетворяет, очевидно, матричному уравнению $A^{-1}X=E$, поэтому она является обратной матрицей для A^{-1}. Теорема 3 доказана полностью.

Замечание. Первая часть теоремы 3 справедлива, очевидно, и в случае произвольного числа сомножителей. Например, если A, B и C — три неособенные матрицы n-го порядка, то

$$(ABC)^{-1}=C^{-1}B^{-1}A^{-1}.$$

Вернемся теперь снова к линейным преобразованиям переменных.

Определение. *Линейное преобразование переменных с квадратной матрицей A называется неособенным, если матрица A неособенная, и называется особенным, если матрица A особенная.*

Если преобразование (2) неособенное, то для него, как мы видели, существует обратное преобразование, выражающее переменные $y_1, y_2, \ldots, y_n$ через $x_1, x_2, \ldots$ $\ldots, x_n$. Этим обратным преобразованием является преобразование (5). Таким образом, имеет место следующая теорема.

Теорема 4. *Для всякого неособенного линейного преобразования переменных с матрицей A существует обратное преобразование, которое является также линейным, и его матрица равна A^{-1}.*

Связь между обратным линейным преобразованием и обратной матрицей может служить средством для нахождения обратной матрицы. Именно, если для заданной неособенной матрицы A мы выпишем линейное преобразование переменных с матрицей A, а затем найдем для него обратное преобразование, то тем самым мы найдем и A^{-1}.

Пример 1. Найти обратную матрицу для

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

Прежде всего проверяем, что $|A| \neq 0$, так что обратная матрица существует. Линейное преобразование

$$\left.\begin{aligned} x_1 &= 2y_1 + y_2 + y_3, \\ x_2 &= \qquad 2y_2 + y_3, \\ x_3 &= 3y_1 + y_2 + 2y_3 \end{aligned}\right\}$$

имеет матрицу A. Находим для него обратное преобразование, решая систему относительно неизвестных y_1, y_2, y_3:

$$\left.\begin{aligned} y_1 &= x_1 - \frac{1}{3}x_2 - \frac{1}{3}x_3, \\ y_2 &= x_1 + \frac{1}{3}x_2 - \frac{2}{3}x_3, \\ y_3 &= -2x_1 + \frac{1}{3}x_2 + \frac{4}{3}x_3. \end{aligned}\right\}$$

Следовательно,

$$A^{-1}=\left\|\begin{matrix} 1 & -\frac{1}{3} & -\frac{1}{3} \\ 1 & \frac{1}{3} & -\frac{2}{3} \\ -2 & \frac{1}{3} & \frac{4}{3} \end{matrix}\right\|.$$

Пример 2. Найти обратную матрицу для

$$A=\left\|\begin{matrix} 0 & 1 \dots 1 \\ 1 & 0 \dots 1 \\ \dots & \dots \\ 1 & 1 \dots 0 \end{matrix}\right\|$$

(на главной диагонали — нули, а все остальные элементы — единицы). Так как $|A|=(-1)^{n-1}(n-1)\neq 0$ при $n\geqslant 2$ (n — порядок матрицы), то A^{-1} существует. Составляем соответствующее преобразование переменных:

$$\left.\begin{aligned} x_1 &= \phantom{y_1+{}} y_2+y_3+\dots+y_{n-1}+y_n, \\ x_2 &= y_1 \phantom{{}+y_2} +y_3+\dots+y_{n-1}+y_n, \\ x_3 &= y_1+y_2 \phantom{{}+y_3} +\dots+y_{n-1}+y_n, \\ &\dots\dots\dots\dots\dots\dots \\ x_n &= y_1+y_2+y_3+\dots+y_{n-1}. \end{aligned}\right\}$$

Сложив все эти равенства и разделив затем на $n-1$, мы получим

$$y_1+y_2+\dots+y_n=\frac{x_1+x_2+\dots+x_n}{n-1}.$$

Сопоставив теперь полученное равенство с каждым из равенств нашего преобразования, мы найдем

$$\left.\begin{aligned} y_1 &= \frac{(2-n)x_1+x_2+\dots+x_n}{n-1}, \\ y_2 &= \frac{x_1+(2-n)x_2+\dots+x_n}{n-1}, \\ &\dots\dots\dots\dots\dots\dots \\ y_n &= \frac{x_1+x_2+\dots+(2-n)x_n}{n-1}. \end{aligned}\right\}$$

Следовательно,

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

Рассмотрим систему n линейных уравнений с n неизвестными:

$$\left.\begin{aligned}a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\dots+a_{2n}x_n&=b_2,\\ \dots\dots\dots\dots\dots\dots&\dots\\ a_{n1}x_1+a_{n2}x_2+\dots+a_{nn}x_n&=b_n.\end{aligned}\right\} \qquad (9)$$

Запишем эту систему в матричной форме:

$$AX=R, \qquad (10)$$

где A — матрица системы, X — столбец из неизвестных, R — столбец из свободных членов. Предположим, что определитель $|A|=\Delta$ системы (9) отличен от нуля. Тогда, по теореме Крамера, система имеет единственное решение, и, с другой стороны, для A существует обратная матрица A^{-1}. Будем под X в (10) подразумевать столбец из решения системы (9). Умножим обе части равенства (10) слева на A^{-1} (произведение имеет смысл, так как A^{-1} — квадратная матрица n-го порядка, а AX и R — столбцы высоты n). Мы получим

$$A^{-1}(AX)=(A^{-1}A)X=EX=X=A^{-1}R.$$

Таким образом, для столбца решений системы (9) имеем

$$X=A^{-1}R. \qquad (11)$$

Пример. Решить систему

$$\left.\begin{aligned}7x_1+2x_2+3x_3&=13,\\ 9x_1+3x_2+4x_3&=15,\\ 5x_1+x_2+3x_3&=14.\end{aligned}\right\}$$

Определитель системы здесь отличен от нуля. Для матрицы системы находим обратную и применяем формулу (11):

$$\begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = \frac{1}{3} \begin{Vmatrix} 5 & -3 & -1 \\ -7 & 6 & -1 \\ -6 & 3 & 3 \end{Vmatrix} \begin{Vmatrix} 13 \\ 15 \\ 14 \end{Vmatrix} = \frac{1}{3} \begin{Vmatrix} 6 \\ -15 \\ 9 \end{Vmatrix} = \begin{Vmatrix} 2 \\ -5 \\ 3 \end{Vmatrix}.$$

Следовательно, $x_1 = 2$, $x_2 = -5$, $x_3 = 3$.

Формула (11) является, по существу, матричной записью формул Крамера. Действительно, в j-й строчке взаимной матрицы $\tilde{A}$ стоят алгебраические дополнения $A_{1j}, A_{2j}, \ldots, A_{nj}$ элементов j-го столбца матрицы A. Следовательно, матричное равенство

$$X = A^{-1}R = \frac{1}{\Delta}\tilde{A}R$$

равносильно следующим n равенствам:

$$x_j = \frac{1}{\Delta}(A_{1j}b_1 + A_{2j}b_2 + \ldots + A_{nj}b_n) \qquad (j = 1, 2, \ldots, n).$$

Но выражение в скобкак равно определителю Δ_j, получающемуся из Δ заменой j-го столбца свободными членами. Следовательно, формула (11) равносильна формулам $x_j = \frac{\Delta_j}{\Delta}$ $(j = 1, 2, \ldots, n)$, что и утверждалось.

Упражнения

1. Найти обратные матрицы для следующих матриц:

a) $\begin{Vmatrix} 7 & 2 \\ 3 & 1 \end{Vmatrix}$; b) $\begin{Vmatrix} 4 & 5 & -5 \\ 1 & 2 & 2 \\ 5 & 7 & -2 \end{Vmatrix}$; c) $\begin{Vmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{Vmatrix}$;

d) $\begin{Vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$; e) $\begin{Vmatrix} 1 & 1 & 1 \ldots & 1 \\ 0 & 1 & 1 \ldots & 1 \\ 0 & 0 & 1 \ldots & 1 \\ \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ 0 & 0 & 0 \ldots & 1 \end{Vmatrix}$.

Отв.

a) $\begin{Vmatrix} 1 & -2 \\ -3 & 7 \end{Vmatrix}$; b) $\frac{1}{3}\begin{Vmatrix} -18 & -25 & 20 \\ 12 & 17 & -13 \\ -3 & -3 & 3 \end{Vmatrix}$; c) $\frac{1}{9}\begin{Vmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{Vmatrix}$;

d) $\begin{Vmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$; e) $\begin{Vmatrix} 1 & -1 & 0 \dots 0 & 0 \\ 0 & 1 & -1 \dots 0 & 0 \\ 0 & 0 & 1 \dots 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 \dots 1 & -1 \\ 0 & 0 & 0 \dots 0 & 1 \end{Vmatrix}$.

2. Решить, пользуясь формулой (11), систему линейных уравнений

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 3, \\ 2x_1 + x_2 + 2x_3 + 3x_4 &= -2, \\ 3x_1 + 2x_2 + x_3 + 2x_4 &= 3, \\ 4x_1 + 3x_2 + 2x_3 + x_4 &= 2. \end{aligned}\right\}$$

Отв. $x_1 = -2$, $x_2 = 5$, $x_3 = -3$, $x_4 = 1$.

3. Показать, что если одна из квадратных матриц n-го порядка A и B особенная, то их произведение AB есть также особенная матрица.

4. Две квадратные матрицы n-го порядка A и B называются *подобными*, если существует такая неособенная матрица C того же порядка n, что $C^{-1}AC = B$. Доказать, что определители подобных матриц равны.

5. Пусть квадратная матрица A порядка n имеет ранг $\leqslant n-2$. Чему равна ее взаимная матрица $\tilde{A}$?

6. Для определителей матрицы A порядка n и ее взаимной матрицы $\tilde{A}$ доказать соотношение

$$|\tilde{A}| = |A|^{n-1}.$$

7. Пусть квадратная матрица A такова, что $A^m = E$ (E — единичная матрица). Показать, что матрица A неособенная. Чему равна ее обратная матрица?

8. Пусть $A^m = O$ (O — нулевая матрица). Доказать, что матрица $E + A + A^2 + \dots + A^{m-1}$ является обратной для $E - A$.

9. Матрица называется *целочисленной*, если все ее элементы — целые числа. Доказать, что для квадратной целочисленной матрицы C обратная матрица существует и является также целочисленной тогда и только тогда, когда $|C| = \pm 1$.

10. Решить матричные уравнения

$$AX = B \text{ и } YA = B, \text{ где } A = \begin{Vmatrix} 2 & 5 \\ 1 & 3 \end{Vmatrix}, \quad B = \begin{Vmatrix} 1 & 2 \\ 0 & 3 \end{Vmatrix}.$$

Отв. $X = \begin{Vmatrix} 3 & -9 \\ -1 & 4 \end{Vmatrix}$, $Y = \begin{Vmatrix} 1 & -1 \\ -3 & 6 \end{Vmatrix}$.

Указание. Найти обратную матрицу A^{-1}.

§ 17. Характеристический многочлен

Пусть A — квадратная матрица порядка n. Если X — столбец высоты n, то произведение AX имеет смысл и также является столбцом высоты n. При рассмотрении ряда вопросов для данной матрицы A приходится разыскивать ненулевые столбцы X, для которых умножение на A слева равносильно умножению на некоторое число λ, т. е. для которых имеет место равенство

$$AX = \lambda X. \tag{1}$$

Нулевой столбец, конечно, при любом λ удовлетворяет этому соотношению. Однако ненулевые столбцы, удовлетворяющие условию (1), существуют далеко не при всяком λ.

Определение. *Число λ называется собственным числом квадратной матрицы A, если существует ненулевой столбец X такой, что $AX = \lambda X$. Если λ — собственное число матрицы A, то всякий столбец X (в том числе и нулевой), удовлетворяющий условию $AX = \lambda X$, называется собственным столбцом матрицы A, соответствующим собственному числу λ.*

Займемся выяснением вопроса, что же представляют собой собственные числа данной квадратной матрицы (и существуют ли они вообще).

Пусть

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix}, \qquad X = \begin{Vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Vmatrix}.$$

Матричное равенство (1) равносильно системе n равенств

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= \lambda x_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= \lambda x_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= \lambda x_n, \end{aligned}\right\} \tag{2}$$

которые можно также переписать в виде

$$\left.\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \dots + a_{2n}x_n &= 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + (a_{nn} - \lambda)x_n &= 0. \end{aligned}\right\} \tag{3}$$

Существование ненулевого столбца X, удовлетворяющего условию (1), равносильно, таким образом, существованию ненулевого решения у системы n линейных однородных уравнений (3) с n неизвестными. Но согласно теореме 2 из § 12 система (3) имеет ненулевое решение тогда и только тогда, когда ее определитель равен нулю, т. е. когда

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0. \qquad (4)$$

Следовательно, собственные числа λ матрицы A характеризуются тем, что для них обращается в нуль определитель (4).

Определение. *Пусть t — переменная. Определитель*

$$\varphi_A(t) = \begin{vmatrix} a_{11}-t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-t \end{vmatrix} = |A - tE| \qquad (5)$$

называется характеристическим многочленом матрицы A.

Тот факт, что определитель (5) является многочленом от переменной t, очевиден (любое из $n!$ произведений, из которых составлен этот определитель, является многочленом от t). Произведение элементов главной диагонали определителя (5) является, очевидно, многочленом от t степени n; при этом коэффициент при t^n равен $(-1)^n$. Все же другие произведения, входящие в состав этого определителя, будут многочленами от t степени, строго меньшей, чем n. Следовательно, степень характеристического многочлена $\varphi_A(t)$ равна порядку n матрицы A и его старший коэффициент равен $(-1)^n$.

Предыдущие наши рассуждения позволяют сформулировать следующую теорему.

Теорема 1. *Все собственные числа квадратной матрицы совпадают с корнями ее характеристического многочлена $\varphi_A(t)$.*

Действительно, если λ — собственное число матрицы A, то, как мы видели, обращается в нуль определитель (4), а значит, λ — корень многочлена (5). Наоборот, если λ — корень характеристического многочлена, т. е. $\varphi_A(\lambda) = 0$, то ввиду (4) система (3) имеет ненулевое решение, и, следовательно, существует ненулевой столбец, удовлетворяющий условию (1); последнее же означает, что λ — собственное число матрицы A.

Согласно теореме 1, чтобы найти собственные числа квадратной матрицы, надо составить ее характеристический многочлен и найти его корни.

Чтобы найти все собственные столбцы матрицы A, соответствующие данному собственному числу λ, надо, очевидно, найти все решения системы (3). Эти решения будут удовлетворять и системе (2), а значит, столбцы из решений будут собственными столбцами матрицы A.

Пример. Найти собственные числа и собственные столбцы матрицы

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

Находим характеристический многочлен $\varphi(t)=|A-tE|$. Он равен $(t-1)(t-2)^3$. Значит, матрица имеет два собственных числа: $\lambda_1=1$ и $\lambda_2=2$. Для каждого λ составляем теперь систему (3):

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

Решая эти системы, получаем, что собственными столбцами матрицы A для собственных чисел $\lambda=1$ и $\lambda=2$ соответственно являются столбцы

$$\begin{Vmatrix}\alpha\\ 0\\ 0\\ 0\end{Vmatrix},\qquad \begin{Vmatrix}\alpha\\ -\alpha\\ \beta\\ -\beta\end{Vmatrix},$$

где α и β — произвольные числа.

До сих пор на элементы матрицы A мы не накладывали никаких ограничений, они могли быть любыми вещественными или комплексными числами. Если мы предположим, что элементы матрицы A вещественны, то отсюда еще нельзя сделать никакого вывода о ее собственных числах. Характеристический многочлен, конечно, будет иметь вещественные коэффициенты, но его корни могут быть комплексными. Например, для матрицы

$$\begin{Vmatrix}1 & -2\\ 1 & -1\end{Vmatrix}$$

характеристический многочлен равен t^2+1. Однако если помимо условия вещественности предположить еще, что матрица симметрична, т. е. совпадает со своей транспонированной, то тогда все ее собственные числа обязательно будут вещественными. К доказательству этого неожиданного факта мы сейчас и переходим.

Для любого комплексного числа $\alpha = a + bi$ через $\bar{\alpha}$ мы будем обозначать комплексно сопряженное число $a - bi$. Имеют место следующие формулы:

$$\overline{\alpha+\beta} = \bar{\alpha} + \bar{\beta},$$
$$\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}.$$

Действительно, если $\alpha = a + bi$, $\beta = c + di$, то

$$\alpha + \beta = (a + c) + (b + d)\, i, \quad \alpha\beta = (ac - bd) + (ad + bc)\, i$$

и

$$\overline{\alpha+\beta} = (a + c) - (b + d)\, i = (a - bi) + (c - di) = \bar{\alpha} + \bar{\beta},$$
$$\overline{\alpha\beta} = (ac - bd) - (ad + bc)\, i = (a - bi)(c - di) = \bar{\alpha}\bar{\beta}.$$

Условие вещественности α может быть выражено, очевидно, в виде равенства

$$\bar{\alpha} = \alpha.$$

Для произвольной матрицы A с комплексными элементами через $\bar{A}$ будем обозначать матрицу, получающуюся из A заменой всех ее элементов на комплексно сопряженные числа. Матрица A вещественна, очевидно, тогда и только тогда, когда

$$\bar{A} = A.$$

Предполагая, что для прямоугольных матриц A и B произведение AB имеет смысл, докажем формулу

$$\overline{AB} = \bar{A}\bar{B}. \tag{6}$$

Пусть $c_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj}$ — элемент i-й строчки и j-го столбца матрицы AB. Тогда элемент i-й строчки и j-го столбца матрицы $\overline{AB}$ равен

$$\bar{c}_{ij} = \overline{a_{i1}b_{1j} + \ldots + a_{in}b_{nj}} = \overline{a_{i1}b_{1j}} + \ldots + \overline{a_{in}b_{nj}} =$$
$$= \bar{a}_{i1}\bar{b}_{1j} + \ldots + \bar{a}_{in}\bar{b}_{nj},$$

а это совпадает с элементом i-й строчки и j-го столбца матрицы $\bar{A}\bar{B}$. Равенство (6), таким образом, установлено.

Справедлива, очевидно, и следующая формула:

$$\overline{(\alpha A)} = \bar{\alpha}\bar{A}. \tag{7}$$

Теорема 2. *Все собственные числа вещественной симметричной матрицы вещественны.*

Доказательство. Пусть A — вещественная симметричная матрица и λ — ее произвольное собственное число. Тогда, согласно определению собственного числа, существует ненулевой столбец X, для которого

$$AX = \lambda X. \tag{8}$$

Пока еще не доказано, что λ вещественно, мы должны с ним обращаться как с комплексным числом. Элементы $x_1, x_2, \ldots, x_n$ столбца X мы также считаем комплексными. Составим выражение

$$X' A\bar{X}. \tag{9}$$

Так как X' — строчка длины n (транспонированный столбец X), A — квадратная матрица порядка n и $\bar{X}$ — столбец высоты n, то произведение (9) имеет смысл и является квадратной матрицей 1-го порядка, т. е. просто числом. Вычислим произведение (9) двумя способами, по-разному расставив скобки. Так как по условию матрица A вещественна и, значит, $\bar{A} = A$, то

$$X'A\bar{X} = X'(\bar{A}\bar{X}) = X'(\overline{AX}) = X'(\overline{\lambda X}) = X'(\bar{\lambda}\bar{X}) = \bar{\lambda}(X'\bar{X}).$$

При выкладках мы воспользовались последовательно формулой (6), равенством (8), формулой (7) и теоремой 2 из § 14.

Воспользовавшись теперь симметричностью матрицы A, т. е. наличием равенства $A' = A$, получим

$$X'A\bar{X} = (X'A')\bar{X} = (AX)'\bar{X} = (\lambda X)'\bar{X} = (\lambda X')\bar{X} = \lambda(X'\bar{X}).$$

Здесь мы применили теорему 2 из § 13, равенство (8), очевидную формулу $(\alpha A)' = \alpha A'$ и теорему 2 из § 14. Сравнивая теперь оба результата мы приходим к равенству

$$\bar{\lambda}(X'\bar{X}) = \lambda(X'\bar{X}). \tag{10}$$

Вычислим произведение в скобках:

$$X'\overline{X} = \|x_1\ x_2\ \dots\ x_n\| \begin{Vmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{Vmatrix} = x_1\bar{x}_1 + x_2\bar{x}_2 + \dots + x_n\bar{x}_n =$$

$$= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2$$

(произведение $\alpha\bar{\alpha}$ равно квадрату модуля $|\alpha|$ комплексного числа α). Так как столбец X ненулевой, то модули $|x_1|$, $|x_2|$, ..., $|x_n|$ не все нули, а потому $X'\overline{X} \neq 0$. Сокращая (10) на этот отличный от нуля множитель, приходим к равенству

$$\bar{\lambda} = \lambda,$$

из которого и вытекает, что собственное число λ вещественно. Теорема 2 доказана.

Установим также одно свойство собственных столбцов симметричной (не обязательно вещественной) матрицы.

Теорема 3. *Если собственные столбцы X и Y симметричной матрицы A соответствуют различным собственным числам λ и μ, то они удовлетворяют соотношению*

$$X'Y = 0.$$

Доказательство. По условию $AX = \lambda X$ и $AY = \mu Y$. Составляем произведение $X'AY$ и вычисляем его двумя способами:

$$X'AY = X'(AY) = X'(\mu Y) = \mu(X'Y),$$

$$X'AY = (X'A')Y = (AX)'Y = (\lambda X)'Y = (\lambda X')Y = \lambda(X'Y).$$

Следовательно, $\mu(X'Y) = \lambda(X'Y)$, откуда, ввиду условия $\mu \neq \lambda$, и вытекает равенство $X'Y = 0$. Теорема 3, таким образом, доказана.

Теоремы 2 и 3 будут нами использованы в главе IV.

Рассмотрим характеристический многочлен матрицы второго порядка

$$A = \begin{Vmatrix} a & b \\ c & d \end{Vmatrix}.$$

Он равен

$$\varphi(t) = \begin{vmatrix} a-t & b \\ c & d-t \end{vmatrix} = t^2 - (a+d)t + (ad - bc).$$

Вычислим теперь $\varphi(A)$:

$$\varphi(A)=A^2-(a+d)A+(ad-bc)E=$$

$$=\left\|\begin{matrix} a^2+bc & ab+bd \\ ca+cd & cb+d^2 \end{matrix}\right\|-\left\|\begin{matrix} a^2+ad & ab+bd \\ ac+cd & ad+d^2 \end{matrix}\right\|+$$

$$+\left\|\begin{matrix} ad-cb & 0 \\ 0 & ad-cb \end{matrix}\right\|=\left\|\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}\right\|.$$

Мы видим, таким образом, что $\varphi(A)$ равно нулевой матрице, т. е. матрица A является как бы корнем своего характеристического многочлена. Этот факт справедлив, оказывается, и для матриц произвольного порядка.

Теорема 4 (теорема Гамильтона — Кэли). *Если $\varphi(t)$ есть характеристический многочлен квадратной матрицы A порядка n, то*

$$\varphi(A)=O,$$

где O — нулевая матрица n-го порядка.

Прежде чем доказывать эту теорему, сделаем несколько замечаний о так называемых полиномных матрицах.

Полиномной матрицей называется матрица, элементами которой являются многочлены от одной переменной t. Например:

$$\left\|\begin{matrix} 3t^2+5t-7 & t^2+5 \\ -t+1 & t^2+3t \end{matrix}\right\|.$$

Действия над полиномными матрицами производятся по тем же правилам, что и над числовыми (с сохранением обычных свойств). Полиномную матрицу в приведенном примере можно представить в виде

$$\left\|\begin{matrix} -7 & 5 \\ 1 & 0 \end{matrix}\right\|+\left\|\begin{matrix} 5 & 0 \\ -1 & 3 \end{matrix}\right\|t+\left\|\begin{matrix} 3 & 1 \\ 0 & 1 \end{matrix}\right\|t^2.$$

Ясно, что всякая полиномная матрица может быть записана аналогичным образом:

$$B_0+B_1t+B_2t^2+\ldots+B_mt^m, \tag{11}$$

где B_0, B_1, ..., B_m — числовые матрицы.

Два многочлена от переменной t называются равными, если у них равны коэффициенты при одинаковых степенях t. Так как равенство двух матриц сводится к равенствам соответствующих элементов, то, как легко видеть, две полиномные матрицы равны тогда и только тогда, когда у них в представлении вида (11) равны матричные коэффициенты при одинаковых степенях t.

Доказательство теоремы 4. Рассмотрим полиномную матрицу

$$B = A - Et.$$

Ее определитель $\Delta = |A - Et|$ равен, разумеется, характеристическому многочлену $\varphi(t)$ матрицы A. Положим

$$\Delta = \varphi(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n$$

(для симметрии мы здесь положили $a_n = (-1)^n$).

Составим для матрицы B взаимную матрицу $\tilde{B}$. Каждый элемент матрицы $\tilde{B}$ с точностью до знака является минором порядка $n-1$ матрицы B. Но поскольку все элементы в B содержат t в степени не выше первой, то любой минор матрицы B (порядка $n-1$) является многочленом степени не выше $n-1$. Таким образом, все элементы взаимной матрицы $\tilde{B}$ — многочлены степени $\leqslant n-1$. Представление (11) для матрицы $\tilde{B}$ имеет, следовательно, вид

$$\tilde{B} = B_0 + B_1 t + B_2 t^2 + \ldots + B_{n-1} t^{n-1}.$$

Согласно теореме 1 из § 16 мы имеем

$$B\tilde{B} = \Delta E = \varphi(t) E. \qquad (12)$$

Найдем представление вида (11) для левой части:

$$\begin{aligned} B\tilde{B} = (A - Et)(B_0 + B_1 t + B_2 t^2 + \ldots + B_{n-1} t^{n-1}) = \\ = AB_0 + AB_1 t + AB_2 t^2 + \ldots + AB_{n-1} t^{n-1} - \\ - B_0 t - B_1 t^2 - \ldots - B_{n-2} t^{n-1} - B_{n-1} t^n = \\ = AB_0 + (AB_1 - B_0) t + (AB_2 - B_1) t^2 + \ldots \\ \ldots + (AB_{n-1} - B_{n-2}) t^{n-1} - B_{n-1} t^n. \end{aligned}$$

С другой стороны, представление (11) для правой части равенства (12) имеет вид

$$\varphi(t) E = a_0 E + a_1 E t + a_2 E t^2 + \ldots + a_{n-1} E t^{n-1} + a_n E t^n.$$

Следовательно, по условию равенства двух полиномных матриц из (12) вытекает следующая система равенств:

$$\begin{aligned}
&AB_0 && = a_0E,\\
&AB_1 && - B_0 = a_1E,\\
&AB_2 && - B_1 = a_2E,\\
&\dots\dots\dots\dots\\
&AB_{n-1} && - B_{n-2} = a_{n-1}E,\\
& && - B_{n-1} = a_nE.
\end{aligned}$$

Домножив эти равенства слева соответственно на E, A, $A^2, \dots, A^{n-1}$, A^n, мы получим новую систему равенств:

$$\begin{aligned}
&AB_0 && = a_0E,\\
&A^2B_1 && - AB_0 = a_1A,\\
&A^3B_2 && - A^2B_1 = a_2A^2,\\
&\dots\dots\dots\dots\dots\dots\\
&A^nB_{n-1} && - A^{n-1}B_{n-2} = a_{n-1}A^{n-1},\\
& && - A^nB_{n-1} = a_nA^n.
\end{aligned}$$

Сложим теперь полученные матричные равенства. Слева, как легко заметить, все члены попарно сокращаются, т. е. мы получим в результате нулевую матрицу O порядка n. Справа же мы получим сумму $a_0E + a_1A + a_2A^2 + \dots + a_nA^n$, которая, согласно определению многочлена от матрицы, равна $\varphi(A)$. Таким образом, мы пришли к равенству $\varphi(A) = O$, и теорема Гамильтона — Кэли доказана.

Упражнения

1. Найти собственные числа и собственные столбцы матриц

a) $\begin{Vmatrix} 2 & 1 \\ 2 & 3 \end{Vmatrix}$; b) $\begin{Vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{Vmatrix}$; c) $\begin{Vmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{Vmatrix}$.

Отв. a) $\lambda = 1$, $\begin{Vmatrix} \alpha \\ -\alpha \end{Vmatrix}$, $\lambda = 4$, $\begin{Vmatrix} \alpha \\ 2\alpha \end{Vmatrix}$; b) $\lambda = 1$, $\begin{Vmatrix} \alpha \\ 0 \\ 0 \end{Vmatrix}$;

c) $\lambda = 2$, $\begin{Vmatrix} 2\alpha - 2\beta \\ \beta \\ \alpha \end{Vmatrix}$, $\lambda = -7$, $\begin{Vmatrix} \alpha \\ 2\alpha \\ -2\alpha \end{Vmatrix}$

(α и β произвольны).

2. Доказать, что любая линейная комбинация собственных столбцов матрицы, соответствующих одному и тому же собственному числу λ, является также собственным столбцом для того же собственного числа λ.

3. Квадратная матрица A называется *кососимметричной*, если $A' = -A$. Доказать, что все собственные числа вещественной кососимметричной матрицы имеют вид $\lambda = bi$ с вещественным b (т. е. расположены на мнимой оси).

4. Квадратная матрица A (с комплексными элементами) называется *эрмитовской*, если $A' = \overline{A}$. Доказать, что все собственные числа эрмитовской матрицы вещественны.

5. Квадратная матрица C (с комплексными элементами) называется *унитарной*, если $C'\overline{C} = E$. Доказать, что все собственные числа унитарной матрицы по модулю равны единице ($|\lambda| = 1$).

Указание. Для собственного столбца X рассмотреть произведение $X'C'\overline{C}\overline{X}$.

6. Пусть X является собственным столбцом для двух квадратных матриц n-го порядка A и B (собственные числа, вообще говоря, различны). Доказать, что тогда X является собственным столбцом и для матриц AB, $A+B$, αA. Чему равны соответствующие собственные числа?

7. Пусть X — собственный столбец квадратной матрицы A, λ — соответствующее собственное число и $f(t)$ — произвольный многочлен. Основываясь на предшествующей задаче, показать, что X является также собственным столбцом матрицы $f(A)$, соответствующим собственному числу $f(\lambda)$.

ГЛАВА IV

КВАДРАТИЧНЫЕ ФОРМЫ

§ 18. Приведение квадратичной формы к диагональному виду

Теория квадратичных форм имеет ряд приложений в математике и в смежных с нею дисциплинах (например, в теоретической механике). Особенно важна эта теория для исследования линий и поверхностей 2-го порядка. Общее уравнение линии 2-го порядка на плоскости имеет вид

$$F(x,\ y)=0, \tag{1}$$

где $F(x, y)$ — многочлен второй степени. Если (1) является уравнением центральной кривой (эллипса или гиперболы), то после параллельного переноса начала координатной системы в центр кривой ее уравнение примет вид

$$ax^2+2bxy+cy^2=h \tag{2}$$

(новые координаты мы здесь обозначили теми же буквами x и y, что и старые). В левой части уравнения (2) стоит однородный многочлен второй степени (степень каждого его члена равна двум). Такой многочлен называется *квадратичной формой* от переменных x и y.

Сделаем теперь поворот координатных осей на некоторый угол α. Старые координаты $(x,\ y)$ произвольной точки плоскости выражаются через новые координаты этой же точки, как известно, по формулам

$$\left.\begin{aligned} x&=x'\cos\alpha-y'\sin\alpha,\\ y&=x'\sin\alpha+y'\cos\alpha. \end{aligned}\right\} \tag{3}$$

Угол α, оказывается, всегда можно подобрать так, чтобы в новых координатах уравнение кривой (2) приняло вид

$$\lambda x'^2 + \mu y'^2 = h. \tag{4}$$

Особенностью уравнения (4) является то, что в нем стоящая слева квадратичная форма не содержит члена с произведением $x'y'$.

При преобразовании уравнения (2) в уравнение (4) преобразуется фактически только стоящая в левой части квадратичная форма. Именно, выражения для x и y из (3) мы подставляем в форму $ax^2 + 2bxy + cy^2$ и после выполнения всех необходимых действий приходим к форме $\lambda x'^2 + \mu y'^2$. Формулы (3) мы можем рассматривать при этом как преобразование переменных (неособенное, поскольку определитель из коэффициентов равен 1). Таким образом, с алгебраической точки зрения преобразование уравнения (2) в уравнение (4) является преобразованием квадратичной формы при помощи неособенного линейного преобразования переменных.

Вот эту задачу приведения квадратичной формы при помощи линейного неособенного преобразования переменных к наиболее простому виду мы и рассмотрим в этом параграфе в самом общем виде.

Определение. *Квадратичной формой называется однородный многочлен $f = f(x_1, x_2, \ldots, x_n)$ второй степени от n переменных.*

Примерами квадратичных форм являются

$$x_1^2 + x_2^2 + \ldots + x_n^2, \qquad \sum_{1 \leq i < j \leq n} x_i x_j,$$
$$x^2 + 2y^2 + 5z^2 - 3xy + 7xz - 10yz.$$

Согласно данному определению коэффициентами квадратичной формы могут быть любые числа (в том числе и комплексные). Если все коэффициенты квадратичной формы f вещественны, то она называется *вещественной*. Если же в качестве коэффициентов формы f допускаются комплексные числа, то f называется *комплексной* квадратичной формой.

Укажем для квадратичных форм одну специальную форму записи. Считая, что в форме $f = f(x_1, x_2, \ldots, x_n)$ уже выполнено приведение подобных членов, обозначим коэффициент при x_i^2 через a_{ii}, а коэффициент при $x_i x_j = x_j x_i$ $(i \neq j)$ через

$$2a_{ij} = a_{ij} + a_{ji} = 2a_{ji},$$

так что

$$a_{ij} = a_{ji}. \tag{5}$$

Для члена, содержащего $x_i x_j$, мы получим тогда следующую симметричную форму записи:

$$2a_{ij}x_i x_j = a_{ij}x_i x_j + a_{ji}x_j x_i.$$

Вся квадратичная форма f может быть записана теперь в виде

$$\left.\begin{array}{l} f(x_1, x_2, \ldots, x_n) = \\ \quad = a_{11}x_1^2 + a_{12}x_1x_2 + \ldots + a_{1n}x_1x_n + \\ \quad + a_{21}x_2x_1 + a_{22}x_2^2 + \ldots + a_{2n}x_2x_n + \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \ldots + a_{nn}x_n^2. \end{array}\right\} \tag{6}$$

Ясно, что коэффициенты a_{ij} формы f в записи (6) определены однозначно. Составленная из них матрица

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix}$$

называется *матрицей квадратичной формы* $f(x_1, x_2, \ldots, x_n)$. Ввиду условия (5) элементы матрицы A, расположенные симметрично относительно главной диагонали, равны между собой. Следовательно, $A' = A$, т. е. A — симметричная матрица.

Очевидно, что для любой симметричной матрицы A всегда можно указать такую квадратичную форму, что ее матрица совпадает с A. Если две квадратичные формы имеют одну и ту же матрицу, то эти формы могут отличаться друг от друга только обозначением переменных, что не имеет существенного значения. Две такие квадратичные формы мы можем считать одинаковыми. Таким образом, квадратичные формы вполне определяются своими матрицами.

Пример. Для квадратичной формы $f(x_1, x_2, x_3) = 2x_1^2 + 4x_1x_2 + 2x_1x_3 + x_2x_3$ запись (6) имеет вид

$$\begin{aligned} f = {} & 2x_1^2 + 2x_1x_2 + x_1x_3 + \\ & + 2x_2x_1 + 0 \cdot x_2^2 + \frac{1}{2}x_2x_3 + \\ & + x_3x_1 + \frac{1}{2}x_3x_2 + 0 \cdot x_3^2. \end{aligned}$$

Следовательно, ее матрица равна

$$\begin{Vmatrix} 2 & 2 & 1 \\ 2 & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 \end{Vmatrix}.$$

Для квадратичной формы может быть дана компактная матричная запись, а именно, пользуясь правилом умножения прямоугольных матриц, мы можем выражение (6) записать в виде произведения строчки на столбец, а затем в виде произведения трех матриц:

$$\begin{aligned} & f(x_1, x_2, \ldots, x_n) = \\ & = x_1(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) + \\ & + x_2(a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n) + \\ & \ldots \ldots \ldots \ldots \ldots \ldots \\ & + x_n(a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n) = \end{aligned}$$

$$= \| x_1\ x_2 \ldots x_n \| \begin{Vmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots \ldots \ldots \ldots \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n \end{Vmatrix} =$$

$$= \| x_1\ x_2\ \ldots\ x_n \| \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Vmatrix}.$$

Таким образом,

$$f(x_1, x_2, \ldots, x_n) = X'AX, \qquad (7)$$

где X — столбец из переменных $x_1, x_2, \ldots, x_n$, а X' — строчка из тех же переменных (транспонированный столбец).

Легко видеть, что если A — произвольная симметричная матрица порядка n и X — столбец из переменных высоты n, то произведение $X'AX$ является квадратичной формой и матрица этой квадратичной формы равна A.

Пусть теперь нам задано линейное преобразование переменных

$$\left.\begin{aligned} x_1 &= c_{11}y_1 + c_{12}y_2 + \ldots + c_{1n}y_n, \\ x_2 &= c_{21}y_1 + c_{22}y_2 + \ldots + c_{2n}y_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= c_{n1}y_1 + c_{n2}y_2 + \ldots + c_{nn}y_n, \end{aligned}\right\} \tag{8}$$

выражающее переменные $x_1, x_2, \ldots, x_n$ через новую систему переменных $y_1, y_2, \ldots, y_n$ (число переменных в обеих системах одно и то же). Если через C мы обозначим матрицу преобразования (8), а через X и Y — столбцы из старых и новых переменных соответственно, то преобразование (8) примет следующую матричную форму:

$$X = CY \tag{9}$$

(см. § 13, стр. 111). Мы будем рассматривать исключительно неособенные линейные преобразования переменных, т. е. будем предполагать, что $|C| \neq 0$. В этом случае для преобразования (9) существует обратное преобразование

$$Y = C^{-1}X, \tag{10}$$

выражающее новые переменные через старые (§ 16, теорема 4).

Подставив выражения для $x_1, x_2, \ldots, x_n$ из преобразования (8) в квадратичную форму $f(x_1, x_2, \ldots, x_n)$, мы получим

$$f(x_1, x_2, \ldots, x_n) = f\left(\sum_{j=1}^{n} c_{1j}y_j,\ \sum_{j=1}^{n} c_{2j}y_j,\ \ldots,\ \sum_{j=1}^{n} c_{nj}y_j\right).$$

После выполнения всех необходимых действий правая часть превратится, очевидно, в квадратичную форму $g(y_1, y_2, \ldots, y_n)$ от новых переменных $y_1, y_2, \ldots, y_n$. Будем говорить в этом случае, что форма $g(y_1, y_2, \ldots, y_n)$ получена из $f(x_1, x_2, \ldots, x_n)$ в результате линейного преобразования переменных (8).

Очевидно, если полученную форму $g(y_1, y_2, \ldots, y_n)$ подвергнуть обратному линейному преобразованию переменных (10), то мы вернемся к исходной форме $f(x_1, x_2, \ldots, x_n)$.

Пример. Подвергнем форму $f = x_1^2 + x_1x_2 + x_2^2$ преобразованию

$$\left.\begin{aligned} x_1 &= y_1 + y_2, \\ x_2 &= y_1 - y_2. \end{aligned}\right\}$$

Мы получим форму

$$g = (y_1 + y_2)^2 + (y_1 + y_2)(y_1 - y_2) + (y_1 - y_2)^2 = 3y_1^2 + y_2^2.$$

Подвергая ее обратному преобразованию

$$y_1 = \frac{x_1 + x_2}{2}, \quad y_2 = \frac{x_1 - x_2}{2},$$

мы приходим к исходной форме

$$3\left(\frac{x_1 + x_2}{2}\right)^2 + \left(\frac{x_1 - x_2}{2}\right)^2 = x_1^2 + x_1x_2 + x_2^2.$$

Выясним, как меняется матрица квадратичной формы при линейном преобразовании переменных.

Подвергнем форму (7) преобразованию (9). Так как $X' = (CY)' = Y'C'$ (теорема 2 § 13), то мы получим

$$f = X'AX = Y'C'ACY = g(y_1, y_2, \ldots, y_n).$$

Рассмотрим квадратную матрицу $B = C'AC$. Так как

$$B' = (C'AC)' = C'A'C'' = C'AC = B,$$

то матрица B симметрична, а значит, она и является матрицей квадратичной формы $g = Y'BY$. Таким образом, можно сформулировать следующую теорему.

Теорема 1. *Если в квадратичной форме с матрицей A сделано линейное преобразование переменных с матрицей C, то полученная квадратичная форма будет иметь матрицу $C'AC$.*

Пример. Квадратичная форма

$$7x_1^2 + 5x_2^2 + 2x_3^2 - 8x_1x_2 + 2x_1x_3 - 6x_2x_3$$

после преобразования

$$\left.\begin{aligned} x_1 &= y_1 + y_2 + y_3, \\ x_2 &= y_1 + 2y_2 + 2y_3, \\ x_3 &= y_1 + y_2 + 2y_3 \end{aligned}\right\}$$

перейдет в форму с матрицей

$$C'AC=\begin{Vmatrix}1&1&1\\1&2&1\\1&2&2\end{Vmatrix}\begin{Vmatrix}7&-4&1\\-4&5&-3\\1&-3&2\end{Vmatrix}\begin{Vmatrix}1&1&1\\1&2&2\\1&1&2\end{Vmatrix}=$$

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

т. е. в форму $2y_1^2+3y_2^2-y_3^2$.

Квадратная матрица вида

$$\begin{Vmatrix}\alpha_1&0&\dots&0\\0&\alpha_2&\dots&0\\ \dots&\dots&\dots&\dots\\0&0&\dots&\alpha_n\end{Vmatrix},$$

у которой все элементы вне главной диагонали равны нулю, называется *диагональной матрицей*. В соответствии с этим мы даем следующее определение.

Определение. *Квадратичная форма вида*

$$\alpha_1x_1^2+\alpha_2x_2^2+\dots+\alpha_nx_n^2,$$

не содержащая членов с произведениями различных переменных и имеющая поэтому диагональную матрицу, называется диагональной квадратичной формой.

Диагональную квадратичную форму часто называют также *канонической*.

Нашей ближайшей задачей является доказательство того, что всякая квадратичная форма при помощи неособенного линейного преобразования переменных всегда может быть приведена к диагональному виду. Сформулируем предварительно два вспомогательных утверждения.

Лемма 1. *Произведение двух (или нескольких) последовательно выполненных неособенных линейных преобразований переменных является также неособенным преобразованием.*

Утверждение этой леммы, по существу, содержится в теореме 3 из § 16. Действительно, неособенность преобразований $X=CY$ и $Y=DZ$ означает, что матрицы

C и D неособенные. Но тогда матрица CD, т. е. матрица линейного преобразования $X = C(DZ) = (CD)Z$, также неособенная.

Лемма 2. *Если у квадратичной формы* (6) *имеется хотя бы один ненулевой коэффициент, то надлежащим неособенным линейным преобразованием переменных $X = CY$ она может быть преобразована в форму, у которой коэффициент при y_1^2 отличен от нуля.*

Доказательство. 1) Пусть $a_{11} \neq 0$. В этом случае сама форма f обладает требуемым свойством и никакого преобразования делать не надо (можно сделать тождественное преобразование).

2) Предположим, что $a_{11} = 0$, но при некотором $i \geqslant 2$ отличен от нуля коэффициент при x_i^2, т. е. $a_{ii} \neq 0$. В этом случае достаточно изменить нумерацию переменных, т. е. сделать преобразование вида

$$x_1 = y_i, \quad x_i = y_1,$$
$$x_k = y_k \text{ при } k \neq 1, \ k \neq i.$$

Это преобразование, очевидно, неособенное (здесь легко усматривается обратное преобразование), и после его выполнения получим

$$f = \ldots + a_{ii}x_i^2 + \ldots = a_{ii}y_1^2 + \ldots$$

Выписанный член $a_{ii}y_1^2$ не имеет себе подобных, а потому он сократиться не может. Таким образом, поскольку $a_{ii} \neq 0$, для рассматриваемого случая лемма доказана.

3) Остается рассмотреть случай, когда $a_{11} = a_{22} = \ldots$ $\ldots = a_{nn} = 0$, т. е. когда все диагональные коэффициенты равны нулю. По условию форма имеет хоть один ненулевой коэффициент. Пусть $a_{ij} \neq 0$ $(i \neq j)$. Чтобы свести рассматриваемый случай к случаю 2, нам достаточно сделать какое-нибудь неособенное преобразование, только бы появился квадрат одной из переменных с ненулевым коэффициентом. Сделаем, например, преобразование

$$x_j = y_j + y_i,$$
$$x_k = y_k \text{ при } k \neq j$$

(в частности, $x_i = y_i$). Преобразование, как легко видеть, неособенное (обратным для него будет преобразование

$y_j = x_j - x_i$, $y_k = x_k$ при $k \neq j$). После выполнения преобразования получим

$$f = \ldots + 2a_{ij}x_ix_j + \ldots = \ldots + 2a_{ij}y_i(y_j + y_i) + \ldots =$$
$$= \ldots + 2a_{ij}y_iy_j + 2a_{ij}y_i^2 + \ldots$$

Член $2a_{ij}y_i^2$ является здесь единственным членом с y_i^2, поэтому после приведения подобных членов он не сократится. Полученная нами форма содержит, таким образом, квадрат переменной с ненулевым коэффициентом ($2a_{ij} \neq 0$), и, значит, к ней применимо рассуждение пункта 2. Для завершения доказательства леммы 2 остается сослаться на лемму 1.

Теорема 2 (теорема Лагранжа). *Всякая квадратичная форма при помощи неособенного линейного преобразования переменных может быть приведена к диагональному виду.*

Доказательство теоремы проведем индукцией по числу переменных n. При $n = 1$ утверждение теоремы справедливо тривиальным образом: всякая квадратичная форма от одной переменной имеет вид ax_1^2 и является, следовательно, диагональной (всякая матрица 1-го порядка диагональна).

Предположим теперь, что $n \geqslant 2$ и что для форм от $n-1$ переменных теорема уже доказана. Пусть $f = f(x_1, x_2, \ldots, x_n)$ — квадратичная форма от n переменных. Если все коэффициенты формы f — нули, то доказывать нечего (сама форма f уже диагональна с нулевыми коэффициентами при квадратах). Пусть не все ее коэффициенты нули. Если $a_{11} = 0$ (в обозначениях (6)), то согласно лемме 2 можно совершить неособенное линейное преобразование переменных так, чтобы после преобразования формы коэффициент при квадрате первой переменной был отличен от нуля. Можно считать поэтому, что уже с самого начала $a_{11} \neq 0$.

Выделим в форме (6) все члены, содержащие x_1:

$$f = a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + g(x_2, \ldots, x_n).$$

Здесь g является, очевидно, формой от $n-1$ переменных $x_2, \ldots, x_n$. Преобразуем теперь выписанную сумму так, чтобы все члены с x_1 вошли в квадрат линейного выражения:

$$f=a_{11}\left(x_1^2+2x_1\left(\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n\right)\right)+g(x_2,\ \ldots,\ x_n)=$$
$$=a_{11}\left(x_1+\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n\right)^2-$$
$$-a_{11}\left(\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n\right)^2+g(x_2,\ \ldots,\ x_n)=$$
$$=a_{11}\left(x_1+\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n\right)^2+f_1(x_2,\ \ldots,\ x_n).$$

Здесь f_1, как и g, опять-таки является квадратичной формой от $n-1$ переменных $x_2,\ \ldots,\ x_n$.

Сделаем преобразование переменных

$$\left.\begin{array}{l} y_1=x_1+\frac{a_{12}}{a_{11}}x_2+\ldots+\frac{a_{1n}}{a_{11}}x_n, \\ y_2=\qquad\quad x_2, \\ \cdots\cdots\cdots\cdots\cdots \\ y_n=\qquad\qquad\qquad\quad x_n. \end{array}\right\} \tag{11}$$

Преобразование (11), очевидно, неособенное. В нем новые переменные выражены через старые. Обычно же мы всегда старые переменные выражаем через новые. Найдем поэтому для (11) обратное преобразование:

$$\left.\begin{array}{l} x_1=y_1-\frac{a_{12}}{a_{11}}y_2-\ldots-\frac{a_{1n}}{a_{11}}y_n, \\ x_2=\qquad\quad y_2, \\ \cdots\cdots\cdots\cdots\cdots \\ x_n=\qquad\qquad\qquad\quad y_n. \end{array}\right\} \tag{12}$$

После выполнения преобразования (12) (или (11)) форма примет вид

$$f=a_{11}y_1^2+f_1(y_2,\ \ldots,\ y_n).$$

По индуктивному предположению существует неособенное линейное преобразование

$$\left.\begin{array}{l} y_2=c_{22}z_2+\ldots+c_{2n}z_n, \\ \cdots\cdots\cdots\cdots \\ y_n=c_{n2}z_2+\ldots+c_{nn}z_n, \end{array}\right\}$$

при котором форма f_1 приводится к диагональному виду

$$f_1(y_2,\ \ldots,\ y_n)=\alpha_2 z_2^2+\ldots+\alpha_n z_n^2.$$

Для нашей исходной формы f вслед за преобразованием (12) выполним следующее преобразование переменных:

$$\left.\begin{array}{l} y_1 = z_1, \\ y_2 = \quad c_{22}z_2 + \ldots + c_{2n}z_n, \\ \cdots\cdots\cdots\cdots \\ y_n = \quad c_{n2}z_2 + \ldots + c_{nn}z_n. \end{array}\right\} \qquad (13)$$

Определитель этого преобразования

$$\begin{vmatrix} 1 & 0 & \ldots & 0 \\ 0 & c_{22} & \ldots & c_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & c_{n2} & \ldots & c_{nn} \end{vmatrix} = \begin{vmatrix} c_{22} & \ldots & c_{2n} \\ \cdot & \cdot & \cdot \\ c_{n2} & \ldots & c_{nn} \end{vmatrix} \neq 0,$$

следовательно, преобразование (13) неособенное. После его выполнения вслед за преобразованием (12) форма f приобретает вид

$$f(x_1, x_2, \ldots, x_n) = a_{11}z_1^2 + \alpha_2 z_2^2 + \ldots + \alpha_n z_n^2.$$

Таким образом, после выполнения двух (а может быть, и трех, если у первоначальной формы коэффициент при квадрате первой переменной был равен нулю) неособенных линейных преобразований форма f приняла диагональный вид. Значит, теорема 2 доказана, так как по лемме 1 несколько последовательно выполненных неособенных линейных преобразований переменных равносильны одному неособенному преобразованию.

Метод доказательства теоремы 2 указывает нам также прием для фактического приведения квадратичной формы к диагональному виду. В доказательстве мы выделили только один квадрат линейного выражения и сослались затем на индуктивное предположение. В конкретном примере для приведения формы к диагональному виду надо процесс выделения квадратов продолжать до тех пор, пока не придем либо к форме с нулевыми коэффициентами, либо к форме от одной переменной (которая всегда является диагональной).

Пример 1. Привести к диагональному виду форму

$$f(x_1, x_2, x_3) = x_1x_2 + 2x_1x_3 + 4x_2x_3.$$

Так как здесь нет членов с квадратом переменной, то делаем вспомогательное преобразование с целью получить квадрат:

$$\left.\begin{aligned} x_1 &= y_1, \\ x_2 &= y_1 + y_2, \\ x_3 &= \qquad\quad y_3. \end{aligned}\right\} \qquad (14)$$

В полученной после этого преобразования форме выделяем квадрат линейного выражения, в котором содержатся все члены с y_1:

$$\begin{aligned} f &= y_1(y_1 + y_2) + 2y_1y_3 + 4(y_1 + y_2)y_3 = \\ &= y_1^2 + y_1y_2 + 6y_1y_3 + 4y_2y_3 = \\ &= \left(y_1 + \frac{1}{2}y_2 + 3y_3\right)^2 - \left(\frac{1}{2}y_2 + 3y_3\right)^2 + 4y_2y_3 = \\ &= \left(y_1 + \frac{1}{2}y_2 + 3y_3\right)^2 - \frac{1}{4}y_2^2 + y_2y_3 - 9y_3^2. \end{aligned}$$

Делаем преобразование

$$\left.\begin{aligned} z_1 &= y_1 + \frac{1}{2}y_2 + 3y_3, \\ z_2 &= \qquad\quad y_2, \\ z_3 &= \qquad\qquad\quad y_3. \end{aligned}\right\}$$

Обратным для него будет

$$\left.\begin{aligned} y_1 &= z_1 - \frac{1}{2}z_2 - 3z_3, \\ y_2 &= \qquad\quad z_2, \\ y_3 &= \qquad\qquad\quad z_3. \end{aligned}\right\} \qquad (15)$$

В полученной форме выделяем квадрат, содержащий все члены с z_2:

$$f = z_1^2 - \frac{1}{4}z_2^2 + z_2z_3 - 9z_3^2 = z_1^2 - \frac{1}{4}(z_2 - 2z_3)^2 - 8z_3^2.$$

Этим определилось преобразование

$$\left.\begin{aligned} u_1 &= z_1, \\ u_2 &= \quad z_2 - 2z_3, \\ u_3 &= \qquad\quad z_3, \end{aligned}\right\}$$

для которого сразу же находим обратное:

$$\left.\begin{aligned} z_1 &= u_1, \\ z_2 &= \quad u_2 + 2u_3, \\ z_3 &= \qquad\quad u_3. \end{aligned}\right\} \tag{16}$$

В результате трех последовательно выполненных неособенных линейных преобразований (14), (15) и (16) форма приняла диагональный вид:

$$f = u_1^2 - \frac{1}{4}u_2^2 - 8u_3^2. \tag{17}$$

Чтобы найти матрицу результирующего преобразования, мы должны перемножить матрицы преобразований (14), (15) и (16):

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

$$= \begin{Vmatrix} 1 & -\frac{1}{2} & -4 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{Vmatrix} = C.$$

Следовательно, неособенным линейным преобразованием, приводящим форму f к диагональному виду (17), является

$$\left.\begin{aligned} x_1 &= u_1 - \frac{1}{2}u_2 - 4u_3, \\ x_2 &= u_1 + \frac{1}{2}u_2 - 2u_3, \\ x_3 &= \qquad\qquad\quad u_3. \end{aligned}\right\} \tag{18}$$

Для проверки вычислим произведение $C'AC$, где A — матрица заданной формы:

$$C'AC = \begin{Vmatrix} 1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -4 & -2 & 1 \end{Vmatrix} \begin{Vmatrix} 0 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & 2 \\ 1 & 2 & 0 \end{Vmatrix} \begin{Vmatrix} 1 & -\frac{1}{2} & -4 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -8 \end{Vmatrix}.$$

Как и следовало ожидать, мы получили матрицу квадратичной формы (17).

Пример 2. Привести к диагональному виду форму

$$f = x_1^2 + 3x_2^2 + x_3^2 + 2x_1x_2 - 4x_1x_3.$$

Имеем

$$f = (x_1 + x_2 - 2x_3)^2 + 2x_2^2 + 4x_2x_3 - 3x_3^2 = \\ = (x_1 + x_2 - 2x_3)^2 + 2(x_2 + x_3)^2 - 5x_3^2.$$

Следовательно, при помощи преобразования

$$\left.\begin{aligned} y_1 &= x_1 + x_2 - 2x_3, \\ y_2 &= \phantom{x_1 + {}} x_2 + x_3, \\ y_3 &= \phantom{x_1 + x_2 + {}} x_3, \end{aligned}\right\} \quad \left.\begin{aligned} x_1 &= y_1 - y_2 + 3y_3, \\ x_2 &= \phantom{y_1 - {}} y_2 - y_3, \\ x_3 &= y_3 \end{aligned}\right\}$$

форма приводится к виду

$$y_1^2 + 2y_2^2 - 5y_3^2.$$

Простота выкладок во втором примере обусловлена тем, что в нем ни в начале, ни по пути не приходилось применять вспомогательное преобразование леммы 2. Ввиду этого матрица линейного преобразования переменных оказалась треугольной.

Замечание. В процессе доказательства теоремы 2 при приведении квадратичной формы к диагональному виду мы производили над ее коэффициентами только рациональные действия (сложение, вычитание, умножение, деление). Следовательно, если заданная квадратичная форма вещественна (т. е. имеет вещественные коэффициенты), то при приведении ее к диагональному виду изложенным приемом все встречающиеся по пути коэффициенты будут также вещественными. В частности, вещественными будут коэффициенты неособенного линейного преобразования, приводящего форму к диагональному виду, и коэффициенты самой диагональной формы.

Ясно также, что квадратичная форма с рациональными коэффициентами приводится к диагональному виду неособенным линейным преобразованием, все коэффициенты которого рациональны (см. рассмотренные выше примеры).

Упражнения

1. Привести форму $x_1x_2+x_2x_3$ (от трех переменных) к диагональному виду.

Отв. $y_1^2-\frac{1}{4}y_2^2$; $\left.\begin{aligned} x_1&=y_1-\frac{1}{2}y_2-y_3,\\ x_2&=y_1+\frac{1}{2}y_2,\\ x_3&=\qquad\qquad\quad y_3\end{aligned}\right\}$

(здесь коэффициент при y_3^2 равен нулю).

2. Привести к диагональному виду квадратичную форму

$$x_1^2+x_2^2+x_3^2+x_4^2+x_1x_2+x_1x_3+x_1x_4+x_2x_3+x_2x_4+x_3x_4.$$

Отв. $y_1^2+\frac{3}{4}y_2^2+\frac{2}{3}y_3^2+\frac{5}{8}y_4^2$ при преобразовании переменных с матрицей

$$\begin{Vmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\\ 0 & 1 & \frac{1}{3} & \frac{1}{3}\\ 0 & 0 & 1 & \frac{1}{4}\\ 0 & 0 & 0 & 1\end{Vmatrix}.$$

3. Привести к диагональному виду квадратичную форму от $2n$ переменных

$$x_1x_2+x_3x_4+\ldots+x_{2n-1}x_{2n}.$$

Отв. $y_1^2-y_2^2+y_3^2-y_4^2+\ldots+y_{2n-1}^2-y_{2n}^2$. Матрица преобразования — клеточно-диагональная матрица с n одинаковыми клетками второго порядка

$$\begin{Vmatrix} 1 & -1\\ 1 & 1\end{Vmatrix}$$

на главной диагонали.

§ 19. Вещественные квадратичные формы

Начиная с этого параграфа и до конца главы мы будем рассматривать исключительно вещественные квадратичные формы. Именно этот случай наиболее важен в приложениях.

В конце § 18 мы видели, что вещественная квадратичная форма может быть приведена к диагональному виду при помощи вещественного неособенного линейного преобразования переменных. Это приведение может быть осуществлено бесконечным числом способов, т. е. существует бесконечно много преобразований, приводящих одну и ту же форму к диагональному виду. При этом коэффициенты при квадратах в диагональной форме также определены не однозначно.

Так, в примере 1 из § 18 (стр. 165) форма f при помощи преобразования (18) приведена нами к виду (17). Легко проверить, что та же самая форма f приводится к виду

$$u_1^2 - u_2^2 - 2u_3^2, \tag{1}$$

если сделать следующее преобразование переменных:

$$\left.\begin{aligned} x_1 &= u_1 - u_2 - 2u_3, \\ x_2 &= u_1 + u_2 - u_3, \\ x_3 &= \tfrac{1}{2}u_3. \end{aligned}\right\}$$

Сравнивая диагональные формы (1) и (17) из § 18, легко подмечаем, что, хотя коэффициенты при квадратах переменных у них и не совпадают, в обоих случаях число положительных и число отрицательных коэффициентов одно и то же. Это является, оказывается, общей закономерностью.

Теорема 1 (закон инерции квадратичных форм). *Если вещественная квадратичная форма вещественными неособенными линейными преобразованиями переменных приведена двумя способами к диагональному виду, то в обоих случаях число положительных коэффициентов, число отрицательных коэффициентов и число нулевых коэффициентов при квадратах новых переменных одно и то же.*

Прежде чем приступать к доказательству теоремы 1, сделаем одно замечание о значениях квадратичных форм. Если в квадратичной форме $f(x_1, x_2, \ldots, x_n)$ вместо переменных мы подставим какие-нибудь числовые значения $x_1^*, x_2^*, \ldots, x_n^*$ и произведем все необходимые вычисления, то в результате получим некоторое число. Это число называется *значением* квадратичной формы f при заданных значениях переменных и обозначается через $f(x_1^*, x_2^*, \ldots, x_n^*)$.

Пусть теперь для формы $f(x_1, x_2, \ldots, x_n)$ выполнено неособенное линейное преобразование переменных

$$x_i = \sum_{j=1}^{n} c_{ij} y_j \qquad (i = 1, 2, \ldots, n) \tag{2}$$

с матрицей C, в результате которого форма f перешла в форму $g(y_1, y_2, \ldots, y_n)$. В силу неособенности матрицы C для произвольных значений $x_1^*, x_2^*, \ldots, x_n^*$ существует единственный набор значений $y_1^*, y_2^*, \ldots, y_n^*$ такой, что

$$x_i^* = \sum_{j=1}^{n} c_{ij} y_j^* \qquad (i = 1, 2, \ldots, n).$$

Значения $y_1^*, y_2^*, \ldots, y_n^*$ мы будем называть значениями, соответствующими значениям $x_1^*, x_2^*, \ldots, x_n^*$ при преобразовании (2). Утверждается, что

$$f(x_1^*, x_2^*, \ldots, x_n^*) = g(y_1^*, y_2^*, \ldots, y_n^*). \tag{3}$$

В самом деле, форма g определяется равенством

$$g(y_1, \ldots, y_n) = f\left(\sum c_{1j} y_j, \ldots, \sum c_{nj} y_j\right),$$

поэтому

$$g(y_1^*, \ldots, y_n^*) = f\left(\sum c_{1j} y_j^*, \ldots, \sum c_{nj} y_j^*\right) = f(x_1^*, \ldots, x_n^*).$$

Ясно, что значения $x_1^*, \ldots, x_n^*$ и $y_1^*, \ldots, y_n^*$ однозначно определяются друг через друга не только преобразованием (2), но и обратным ему преобразованием.

Доказательство теоремы 1. Пусть вещественная квадратичная форма $f(x_1, x_2, \ldots x_n)$ при помощи вещественных неособенных линейных преобразований переменных двумя способами приведена к диагональному виду:

$$f(x_1, x_2, \ldots, x_n) = \alpha_1 y_1^2 + \ldots + \alpha_p y_p^2 - \beta_1 y_{p+1}^2 - \ldots - \beta_q y_{p+q}^2, \tag{4}$$

$$f(x_1, x_2, \ldots, x_n) = \gamma_1 z_1^2 + \ldots + \gamma_r z_r^2 - \delta_1 z_{r+1}^2 - \ldots - \delta_s z_{r+s}^2. \tag{5}$$

Мы считаем, что здесь все α_i, β_j, γ_k, δ_l строго положительны, так что форма (4) при квадратах новых переменных имеет p положительных коэффициентов, q отрицательных и $n-p-q$ нулевых, а форма (5) — r положительных, s отрицательных и $n-r-s$ нулевых. Конечно, может случиться, что положительные, отрицательные и нулевые коэффициенты при квадратах новых переменных расположены вперемежку, но, поскольку изменение нумерации переменных является неособенным линейным преобразованием, мы всегда можем коэффициенты при квадратах расположить в любом порядке, не меняя при этом самих коэффициентов. Заметим еще, что в конкретных случаях некоторые из коэффициентов (положительные, отрицательные или нулевые) могут фактически отсутствовать.

Мы должны показать, что

$$p = r, \quad q = s, \quad n-p-q = n-r-s. \tag{6}$$

Выпишем соответствующие линейные преобразования переменных (вещественные и неособенные), выразив при этом, против нашего обыкновения, новые переменные через старые:

$$y_i = \sum_{j=1}^{n} c_{ij} x_j \qquad (i = 1, 2, \ldots, n), \tag{7}$$

$$z_i = \sum_{j=1}^{n} d_{ij} x_j \qquad (i = 1, 2, \ldots, n). \tag{8}$$

Приступая к доказательству первого из равенств (6), предположим, вопреки утверждению теоремы, что $p \neq r$. Для определенности можно, конечно, считать, что $p < r$. Мы покажем, что тогда для старых переменных можно найти такие вещественные значения $x_1^*, x_2^*, \ldots, x_n^*$, не равные одновременно нулю, что соответствующие им значения новых переменных $y_1^*, y_2^*, \ldots, y_n^*$ и $z_1^*, z_2^*, \ldots, z_n^*$ будут удовлетворять условиям

$$y_1^* = 0, \ldots, y_p^* = 0, \tag{9}$$

$$z_{r+1}^* = 0, \ldots, z_n^* = 0. \tag{10}$$

Вычислив затем по формуле (3) значение $f(x_1^*, \ldots, x_n^*)$ двумя способами с помощью диагональных форм (4) и (5), мы получим несовместимые между собой неравенства $f(x_1^*, \ldots, x_n^*) \leqslant 0$ и $f(x_1^*, \ldots, x_n^*) > 0$.

Итак, пусть $p < r$. Рассмотрим следующую систему линейных однородных уравнений:

$$\left.\begin{array}{l} c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ c_{p1}x_1 + c_{p2}x_2 + \ldots + c_{pn}x_n = 0, \\ d_{r+1,1}x_1 + d_{r+1,2}x_2 + \ldots + d_{r+1,n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ d_{n1}x_1 + d_{n2}x_2 + \ldots + d_{nn}x_n = 0. \end{array}\right\} \tag{11}$$

Система эта получена приравниванием к нулю первых p линейных выражений справа в преобразовании (7) и последних $n - r$ линейных выражений в преобразовании (8). Число уравнений в системе (11) равно

$$p + (n - r) = n - (r - p),$$

что строго меньше n (поскольку $r > p$). Но раз число уравнений в однородной системе меньше числа неизвестных, то, по теореме 3 из § 12, она имеет ненулевое решение. Ввиду вещественности преобразований (7) и (8) коэффициенты системы (11) также вещественны, а значит, ненулевое решение этой системы мы можем выбрать вещественным. Пусть это будет $x_1^*, x_2^*, \ldots, x_n^*$. Этим мы и нашли искомые значения для старых переменных. Если теперь в соответствии с преобразованиями (7) и (8) мы найдем соответствующие значения для новых переменных $y_1^*, y_2^*, \ldots, y_n^*$ и $z_1^*, z_2^*, \ldots, z_n^*$, то для этих значений будут выполнены, очевидно, условия (9) и (10).

Согласно равенствам (4), (5) и формуле (3) мы имеем

$$f(x_1^*, \ldots, x_n^*) = -\beta_1 y_{p+1}^{*2} - \ldots - \beta_q y_{p+q}^{*2}, \tag{12}$$

$$f(x_1^*, \ldots, x_n^*) = \gamma_1 z_1^{*2} + \ldots + \gamma_r z_r^{*2}. \tag{13}$$

Квадрат вещественного числа неотрицателен, поэтому из (12) следует, что

$$f(x_1^*, \ldots, x_n^*) \leqslant 0. \tag{14}$$

(Знак равенства здесь возможен ввиду того, что в случае $p+q<n$ ненулевые значения y_i^* могут оказаться только среди значений $y_{p+q+1}^*, \ldots, y_n^*$.) С другой стороны, в силу неособенности преобразования (8) не все z_i^* равны нулю (в противном случае все x_i^* были бы равны нулю); поэтому, учитывая (10), мы заключаем, что хоть одно из значений $z_1^*, \ldots, z_r^*$ отлично от нуля, а значит, из (13) вытекает неравенство

$$f(x_1^*, \ldots, x_n^*) > 0. \tag{15}$$

Таким образом, предположение о том, что $p \neq r$, привело нас к противоречащим друг другу неравенствам (14) и (15). Следовательно, $p = r$, т. е. число положительных коэффициентов в обеих диагональных формах (4) и (5) одно и то же.

Второе из равенств (6) может быть доказано аналогичным способом. Но можно поступить и следующим образом. Рассмотрим форму $-f(x_1, \ldots, x_n)$. При помощи тех же неособенных вещественных преобразований (7) и (8) форма $-f$ приводится к двум диагональным формам, получающимся из (4) и (5) умножением на -1. Первая из них будет иметь q, а вторая s положительных коэффициентов. Но по доказанному число положительных коэффициентов в обоих способах приведения одно и то же. Следовательно, $q = s$.

Третье из равенств (6) является очевидным следствием первых двух. Теорема 1 доказана.

Среди вещественных квадратичных форм особую роль играют формы, выделяемые следующим определением.

Определение. *Вещественная квадратичная форма называется положительно определенной, если положительны все ее значения при вещественных значениях переменных, не равных нулю одновременно.*

Положительная определенность вещественной формы $f(x_1, \ldots, x_n)$ означает, стало быть, что

$$f(x_1^*, \ldots, x_n^*) > 0,$$

если только среди вещественных x_i^* хоть одно значение отлично от нуля. Если все значения переменных равны нулю, то значение квадратичной формы, конечно, всегда равно нулю.

Примерами положительно определенных квадратичных форм являются

$$x_1^2+x_2^2+\ldots+x_n^2,$$
$$2x^2+5y^2+7z^2,$$
$$x^2+xy+y^2,$$
$$(x_1+x_2+x_3)^2+(x_2+x_3)^2+x_3^2.$$

Теорема 2. *Для того чтобы вещественная квадратичная форма была положительно определенной, необходимо и достаточно, чтобы при приведении ее к диагональному виду вещественным неособенным линейным преобразованием переменных все коэффициенты при квадратах новых переменных были положительны.*

Доказательство. Пусть вещественная форма $f(x_1,\ldots,x_n)$ вещественным неособенным преобразованием (2) приведена к диагональному виду

$$\alpha_1y_1^2+\alpha_2y_2^2+\ldots+\alpha_ny_n^2.$$

Если $x_1^*, x_2^*,\ldots, x_n^*$ — какие-нибудь значения старых переменных и $y_1^*, y_2^*,\ldots, y_n^*$ — соответствующие им значения новых переменных; то, согласно формуле (3), мы имеем

$$f(x_1^*,\ldots,x_n^*)=\alpha_1y_1^{*2}+\alpha_2y_2^{*2}+\ldots+\alpha_ny_n^{*2}. \qquad (16)$$

Пусть все коэффициенты $\alpha_1, \alpha_2,\ldots,\alpha_n$ положительны. Тогда, если вещественные значения $x_1^*,\ldots,x_n^*$ не все нули, то вещественные же значения $y_1^*,\ldots,y_n^*$ также не все нули (в силу неособенности преобразования (2)), и поэтому правая часть равенства (16), а значит, и левая его часть положительны. Таким образом, при положительных α_i форма f положительно определенная.

Наоборот, пусть форма f положительно определенная. Положим

$$y_1^*=0,\ \ldots,\ y_{i-1}^*=0,\ y_i^*=1,\ y_{i+1}^*=0,\ldots,\ y_n^*=0. \qquad (17)$$

Соответствующие значения $x_1^*,\ldots,x_n^*$ старых переменных, очевидно, вещественны и не равны нулю одновременно, так что $f(x_1^*,\ldots,x_n^*)>0$. Но согласно (16) и (17) $f(x_1^*, \ldots, x_n^*) = \alpha_i$, следовательно, $\alpha_i > 0$ при всех $i = 1, 2, \ldots, n$.

Квадратичную форму

$$x_1^2 + x_2^2 + \ldots + x_n^2$$

мы будем называть *чистой суммой квадратов*. Чистая сумма квадратов в качестве своей матрицы имеет, очевидно, единичную матрицу.

Теорема 3. *Вещественная квадратичная форма является положительно определенной тогда и только тогда, когда она вещественным неособенным линейным преобразованием переменных может быть приведена к чистой сумме квадратов.*

Доказательство. Если форма приводится к чистой сумме квадратов, то, согласно теореме 2, она положительно определенная. Наоборот, пусть $f(x_1, \ldots, x_n)$ — положительно определенная квадратичная форма. Приведем ее вещественным неособенным преобразованием (2) к диагональному виду

$$f = \alpha_1 y_1^2 + \alpha_2 y_2^2 + \ldots + \alpha_n y_n^2. \tag{18}$$

По теореме 2 все коэффициенты α_i здесь положительны. Сделаем вслед за преобразованием (2) следующее преобразование переменных:

$$\left.\begin{array}{l} y_1 = \frac{1}{\sqrt{\alpha_1}} z_1, \\ y_2 = \qquad \frac{1}{\sqrt{\alpha_2}} z_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = \qquad\qquad \frac{1}{\sqrt{\alpha_n}} z_n. \end{array}\right\} \tag{19}$$

При выполнении этого вещественного и неособенного преобразования форма (18) приобретает вид

$$z_1^2 + z_2^2 + \ldots + z_n^2,$$

т. е. переходит в чистую сумму квадратов. Для завершения доказательства теоремы 3 остается только заметить, что последовательное выполнение преобразований (2) и (19) равносильно одному вещественному неособенному преобразованию (см. лемму 1 из § 18).

Приведем без доказательства критерий, позволяющий устанавливать положительную определенность вещественных квадратичных форм непосредственно через их коэффициенты, без приведения к диагональному виду.

Теорема 4. *Для того чтобы вещественная квадратичная форма от n переменных была положительно определенной, необходимо и достаточно, чтобы все n миноров ее матрицы, расположенные в левом верхнем углу (порядков 1, 2, ..., n), были положительны.*

Другими словами, если для вещественной квадратичной формы $f(x_1, x_2, \ldots, x_n)$ с матрицей

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

мы положим

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix} \qquad (k = 1, 2, \ldots, n),$$

то эта форма f будет положительно определенной тогда и только тогда, когда

$$\Delta_1 > 0, \quad \Delta_2 > 0, \ \ldots, \ \Delta_n > 0.$$

Например, форма

$$3x_1^2 + 4x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_2x_3$$

является положительно определенной, так как для нее

$$\Delta_1 = 3 > 0, \qquad \Delta_2 = \begin{vmatrix} 3 & 2 \\ 2 & 4 \end{vmatrix} = 8 > 0,$$

$$\Delta_3 = \begin{vmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{vmatrix} = 28 > 0.$$

Заметим, что наряду с положительно определенными формами иногда рассматривают также и *отрицательно определенные формы.* Так называются вещественные квадратичные формы, все значения которых при вещественных значениях переменных, не равных одновременно нулю, отрицательны. Отрицательно определенные формы характеризуются тем, что при приведении их вещественными неособенными линейными преобразованиями переменных к диагональному виду все коэффициенты при квадратах новых переменных отрицательны. Вещественная квадратичная форма называется *неопределенной*, если она обладает как положительными, так и отрицательными значениями (при вещественных значениях переменных). Наконец, если среди значений формы нет отрицательных или нет положительных, но нулевые значения возникают не только при нулевых значениях переменных, то такая форма называется *полуопределенной*.

Упражнения

1. Показать, что всякая вещественная квадратичная форма от n переменных вещественным неособенным линейным преобразованием переменных может быть приведена к виду

$$y_1^2+y_2^2+\ldots+y_s^2-y_{s+1}^2-\ldots-y_r^2 \quad (r\leqslant n).$$

2. Показать, что для любой положительно определенной квадратичной формы (не обязательно диагональной) все коэффициенты при квадратах переменных положительны.

3. Доказать, что вещественная симметричная матрица A является матрицей положительно определенной квадратичной формы тогда и только тогда, когда ее можно представить в виде $A=C'C$, где C — вещественная неособенная матрица.

4. Выяснить, какая из квадратичных форм

$$2x_1^2+3x_2^2+4x_3^2-4x_1x_2-2x_1x_3+4x_2x_3,$$
$$2x_1^2+3x_2^2+4x_3^2-4x_1x_2-2x_1x_3-4x_2x_3$$

является положительно определенной.

Отв. Первая форма положительно определенная, вторая —нет.

5. Доказать необходимость условия теоремы 4.

Указание. Установить, что для положительно определенной квадратичной формы $f(x_1, \ldots, x_n)$ «частичная» форма $f(x_1, \ldots, x_k, 0, \ldots, 0)$ также положительно определенная. Использовать затем упражнение 3.

§ 20. Ортогональные преобразования переменных

В начале § 18 мы упоминали о задаче приведения общего уравнения кривой 2-го порядка к каноническому виду путем надлежащего выбора новой системы координат. Именно эта задача и послужила для нас отправным моментом в построении теории приведения квадратичных форм к диагональному виду. Изучая вопросы приведения квадратичных форм, мы допускали произвольные неособенные линейные преобразования переменных. Однако в случае декартовых прямоугольных систем координат с одинаковым масштабом по всем координатным осям формулы преобразования координат являются преобразованиями весьма специального вида. Так, в случае поворота координатных осей на угол α вокруг неподвижного начала координат старые координаты x, y произвольной точки плоскости выражаются через ее новые координаты x', y' по формулам (3) из § 18. Среди преобразований вида

$$\left.\begin{aligned} x &= \lambda_1 x' + \lambda_2 y', \\ y &= \mu_1 x' + \mu_2 y' \end{aligned}\right\} \tag{1}$$

формулы (3) из § 18 выделяются, как можно показать, условиями

$$\lambda_1^2 + \mu_1^2 = 1, \quad \lambda_2^2 + \mu_2^2 = 1, \tag{2}$$

$$\lambda_1 \lambda_2 + \mu_1 \mu_2 = 0, \tag{3}$$

$$\begin{vmatrix} \lambda_1 & \lambda_2 \\ \mu_1 & \mu_2 \end{vmatrix} = 1. \tag{4}$$

Если мы рассмотрим преобразование декартовых прямоугольных координат плоскости, при котором новая система получается из старой поворотом на угол α с последующим изменением направления одной из координатных осей на противоположное, то в этом случае формулы преобразования координат, записанные в виде (1), будут характеризоваться условиями (2), (3) и вместо (4) условием

$$\begin{vmatrix} \lambda_1 & \lambda_2 \\ \mu_1 & \mu_2 \end{vmatrix} = -1. \tag{5}$$

Легко показать, что при выполнении (2) и (3) автоматически имеет место одно из условий (4) или (5). Следовательно, объединяя оба случая вместе, можно сказать, что преобразование (1) выражает формулы преобразования прямоугольных координат на плоскости (с сохранением начала и масштаба) тогда и только тогда, когда выполнены условия (2) и (3). Равенства (4) и (5) свидетельствуют при этом о сохранении или, соответственно, об изменении ориентации координатных осей.

Рассмотрим теперь с такой же точки зрения формулы преобразования координат в пространстве.

Пусть в пространстве заданы две декартовы прямоугольные системы координат с общим началом и с одинаковым масштабом по всем координатным осям. Координаты x, y, z произвольной точки пространства в одной системе координат выражаются через ее координаты x', y', z' во второй системе по формулам вида

$$\left.\begin{aligned} x &= \lambda_1 x' + \lambda_2 y' + \lambda_3 z', \\ y &= \mu_1 x' + \mu_2 y' + \mu_3 z', \\ z &= \nu_1 x' + \nu_2 y' + \nu_3 z'. \end{aligned}\right\} \qquad (6)$$

В аналитической геометрии доказывается, что преобразование (6) выражает формулы преобразования декартовых прямоугольных координат в пространстве (с сохранением начала и масштаба) тогда и только тогда, когда его коэффициенты удовлетворяют условиям

$$\left.\begin{aligned} \lambda_1^2 + \mu_1^2 + \nu_1^2 &= 1, \\ \lambda_2^2 + \mu_2^2 + \nu_2^2 &= 1, \\ \lambda_3^2 + \mu_3^2 + \nu_3^2 &= 1, \end{aligned}\right\} \qquad (7)$$

$$\left.\begin{aligned} \lambda_1\lambda_2 + \mu_1\mu_2 + \nu_1\nu_2 &= 0, \\ \lambda_1\lambda_3 + \mu_1\mu_3 + \nu_1\nu_3 &= 0, \\ \lambda_2\lambda_3 + \mu_2\mu_3 + \nu_2\nu_3 &= 0. \end{aligned}\right\} \qquad (8)$$

При этом

$$\begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{vmatrix} = \pm 1. \qquad (9)$$

Обе координатные системы имеют одну и ту же ориентацию координатных осей тогда и только тогда, когда определитель (9) равен $+1$.

Из коэффициентов формул преобразования координат (6) составим матрицу

$$\begin{Vmatrix} \lambda_1 & \lambda_2 & \lambda_3 \\ \mu_1 & \mu_2 & \mu_3 \\ \nu_1 & \nu_2 & \nu_3 \end{Vmatrix} \qquad (10)$$

и рассмотрим векторы в пространстве, координатами которых в первой координатной системе являются столбцы матрицы (10). Равенства (7) геометрически означают, что эти три вектора все единичной длины, а равенства (8) — что они попарно ортогональны (рассматриваемые векторы образуют координатный базис второй системы координат).

Обобщим теперь линейные преобразования вида (1) и (6) (с условиями (2) — (3) и (7) — (8) соответственно) на случай произвольного числа переменных. Для этой цели введем следующее вспомогательное определение.

Определение. *Вещественный столбец T высоты n называется нормированным, если сумма квадратов его элементов равна* 1. *Два вещественных столбца T_1 и T_2 одной и той же высоты n называются ортогональными, если сумма произведений их соответствующих элементов равна* 0.

Равенства (7) и (8) означают, например, что столбцы матрицы (10) нормированы и попарно ортогональны.

Если воспользоваться правилом умножения строчки на столбец, то условие нормированности столбца T может быть выражено в виде равенства

$$T'T = 1, \tag{11}$$

а условие ортогональности столбцов T_1 и T_2 — равенством

$$T_1'T_2 = 0 \tag{12}$$

(или, что то же самое, равенством $T_2'T_1 = 0$).

Всякий ненулевой вещественный столбец

$$T = \begin{Vmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{Vmatrix}$$

может быть превращен в нормированный столбец путем умножения его на надлежащий множитель. Действительно, нормированность столбца αT означает, что

$$\alpha^2 a_1^2 + \alpha^2 a_2^2 + \ldots + \alpha^2 a_n^2 = 1,$$

поэтому надо взять

$$\alpha = \pm \frac{1}{\sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}}$$

(сумма квадратов под знаком корня отлична от нуля, так как столбец ненулевой).

Определение. *Вещественная квадратная матрица называется ортогональной, если все ее столбцы нормированы и попарно ортогональны. Линейное преобразование переменных называется ортогональным, если его матрица ортогональна.*

Условие ортогональности вещественной матрицы

$$C = \begin{Vmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{Vmatrix}$$

выражается, согласно определению, равенствами

$$\left.\begin{aligned} c_{1i}^2 + c_{2i}^2 + \ldots + c_{ni}^2 = 1 \quad (i = 1, 2, \ldots, n), \\ c_{1i}c_{1j} + c_{2i}c_{2j} + \ldots + c_{ni}c_{nj} = 0 \quad (1 \leqslant i < j \leqslant n). \end{aligned}\right\} \qquad (13)$$

При $n = 2$ и $n = 3$ эти равенства превращаются, очевидно, в условия (2)—(3) и (7)—(8) соответственно. Следовательно, при $n = 2$ и $n = 3$ ортогональные преобразования переменных, и только они, выражают формулы преобразования декартовых прямоугольных координат (при переходе от одной системы к другой с сохранением начала и с одинаковым масштабом по всем координатным осям).

Теорема 1. *Для того чтобы вещественная квадратная матрица C была ортогональной, необходимо и достаточно, чтобы*

$$C'C = E$$

(E — единичная матрица).

Доказательство. Представим матрицу C в виде совокупности столбцов $C = \| T_1 \; T_2 \ldots T_n \|$ и положим $C'C = X$. Согласно правилу умножения матриц элемент i-й строчки и j-го столбца матрицы X равен

$$x_{ij} = T_i' T_j.$$

С другой стороны, условие ортогональности матрицы C

равносильно тому, что

$$T_i'T_j=\begin{cases}1 & \text{при} \quad i=j,\\ 0 & \text{при} \quad i\neq j\end{cases}$$

(см. формулы (11) и (12)). Следовательно, матрица C ортогональна тогда и только тогда, когда $X=E$.

Теорема 2. *Для того чтобы линейное преобразование переменных с вещественной квадратной матрицей было ортогональным, необходимо и достаточно, чтобы при этом преобразовании чистая сумма квадратов переменных переходила опять в чистую сумму квадратов новых переменных.*

Доказательство. Чистая сумма квадратов переменных $x_1^2+\ldots+x_n^2$, как квадратичная форма, в качестве своей матрицы имеет единичную матрицу E. После линейного преобразования переменных с матрицей C форма $x_1^2+\ldots+x_n^2$ перейдет, согласно теореме 1 из § 18, в квадратичную форму с матрицей $C'EC=C'C$. Следовательно, полученная форма будет чистой суммой квадратов тогда и только тогда, когда $C'C=E$, т. е. когда матрица C ортогональна.

Теоремы 1 и 2 мы можем рассматривать, разумеется, как другие определения ортогональной матрицы и ортогонального преобразования переменных.

Теорема 3. *Ортогональные матрицы обладают следующими свойствами:*

1) *определитель ортогональной матрицы C равен ± 1;*

2) *произведение ортогональных матриц есть ортогональная матрица;*

3) *единичная матрица E ортогональна;*

4) *для ортогональной матрицы C обратная матрица существует и равна транспонированной матрице C';*

5) *для ортогональной матрицы C обратная матрица C^{-1} также ортогональна.*

Доказательство. 1) Перейдем в равенстве $C'C=E$ к определителям. Тогда, в силу теоремы из § 15 и свойства 1 из § 4, получим $|C'C|=|C'||C|=|C|^2=1$, откуда $|C|=\pm 1$.

2) Пусть C_1 и C_2 — две ортогональные матрицы (одного и того же порядка), и пусть $C=C_1C_2$. Тогда, согласно теореме 2 из § 13, мы имеем $C'C=(C_1C_2)'C_1C_2=C_2'C_1'C_1C_2=C_2'EC_2=C_2'C_2=E$, а значит, матрица C ортогональна.

Свойства 3) и 4) очевидны.

5) Пусть C — ортогональная матрица. Тогда

$$(C^{-1})' C^{-1} = (C')' C^{-1} = CC^{-1} = E,$$

т. е. матрица C^{-1} ортогональна.

Следствие. *Произведение двух (или нескольких) последовательно выполненных ортогональных преобразований переменных является ортогональным преобразованием. Для ортогонального преобразования переменных обратное преобразование также ортогонально.*

Первое утверждение следует из того, что при умножении линейных преобразований перемножаются их матрицы. Второе вытекает из теоремы 4 § 16.

Обратимся теперь к вопросу о построении ортогональных матриц.

Лемма. *Пусть мы имеем k нормированных и попарно ортогональных столбцов $T_1, T_2, \dots, T_k$ высоты n (с вещественными элементами). Если $k < n$, то можно построить нормированный столбец T, ортогональный ко всем столбцам $T_1, T_2, \dots, T_k$.*

Доказательство. Пусть

$$T_1 = \begin{Vmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{Vmatrix}, \quad T_2 = \begin{Vmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{Vmatrix}, \quad \dots, \quad T_k = \begin{Vmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{nk} \end{Vmatrix}, \quad T = \begin{Vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Vmatrix}.$$

Ортогональность столбца T к столбцам $T_1, T_2, \dots, T_k$ равносильна выполнению следующих равенств:

$$\left.\begin{array}{l} c_{11}x_1 + c_{21}x_2 + \dots + c_{n1}x_n = 0, \\ c_{12}x_1 + c_{22}x_2 + \dots + c_{n2}x_n = 0, \\ \dots\dots\dots\dots\dots\dots \\ c_{1k}x_1 + c_{2k}x_2 + \dots + c_{nk}x_n = 0. \end{array}\right\}$$

Для определения столбца T мы получили однородную систему k линейных уравнений с n неизвестными. Так как $k < n$, то, по теореме 3 из § 12, система имеет ненулевые решения. Пусть $a_1, a_2, \dots, a_n$ — какое-нибудь ненулевое вещественное решение системы (коэффициенты c_{ij} по условию вещественны). Положим

$$c_i = \frac{a_i}{\sqrt{a_1^2 + \dots + a_n^2}} \qquad (i = 1, 2, \dots, n).$$

Столбец с элементами $c_1, c_2, \ldots, c_n$ будет, очевидно, нормированным и ортогональным ко всем столбцам $T_1, \ldots, T_l$.

Повторное применение доказанной леммы (пока не получим систему из n нормированных попарно ортогональных столбцов) позволяет сформулировать следующую теорему.

Теорема 4. *Для любой системы k нормированных попарно ортогональных столбцов $T_1, \ldots, T_k$ существует ортогональная матрица C, первые k столбцов которой совпадают с заданными столбцами $T_1, \ldots, T_k$.*

Следствие. *Для любого нормированного столбца T существует ортогональная матрица C, первый столбец которой совпадает с T.*

Упражнения

1. Понятия нормированности и ортогональности могут быть перенесены и на строчки. Показать, что в ортогональной матрице все строчки нормированы и попарно ортогональны.

2. Построить ортогональную матрицу, две первые строчки которой совпадают со строчками $\left\| \frac{2}{3} \; -\frac{2}{3} \; \frac{1}{3} \right\|$ и $\left\| \frac{2}{3} \; \frac{1}{3} \; -\frac{2}{3} \right\|$.

Отв. Надо добавить одну из строчек

$$\left\| \frac{1}{3} \quad \frac{2}{3} \quad \frac{2}{3} \right\| \text{ или } \left\| -\frac{1}{3} \quad -\frac{2}{3} \quad -\frac{2}{3} \right\|.$$

3. Ясно, что целочисленная матрица является ортогональной тогда и только тогда, когда в каждой строчке и в каждом столбце имеется только один отличный от нуля элемент, равный ± 1. Доказать, что всего имеется $2^n n!$ целочисленных ортогональных матриц порядка n.

4. Доказать, что всякая ортогональная матрица 2-го порядка с определителем $+1$ имеет вид

$$\begin{Vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{Vmatrix}.$$

§ 21. Приведение квадратичной формы к диагональному виду ортогональным преобразованием

В предыдущем параграфе мы видели, что для решения задачи приведения общего уравнения кривой 2-го порядка на плоскости или поверхности 2-го порядка в пространстве важно уметь приводить вещественные квадратичные формы к диагональному виду ортогональным преобразованием переменных (а не произвольным неособенным преобразованием, как это мы делали в § 18 и 19).

В дальнейшем столбец из переменных $x_1, x_2, \ldots, x_n$ мы будем обозначать через X, столбец из $y_1, y_2, \ldots, y_n$ — через Y и т. д.

Теорема 1. *Каждая вещественная квадратичная форма $f(x_1, x_2, \ldots, x_n)$ с матрицей A при помощи некоторого ортогонального преобразования переменных $X=CY$ может быть приведена к диагональному виду*

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2. \tag{1}$$

Коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_n$ при квадратах новых переменных с точностью до порядка определены формой $f(x_1, \ldots, x_n)$ однозначно: они совпадают с корнями характеристического многочлена $\varphi(t)=|A-tE|$ матрицы A. Столбцы $T_1, T_2, \ldots, T_n$ ортогональной матрицы C являются собственными столбцами матрицы A, соответствующими собственным числам $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Доказательство. 1°. Предположим сначала, что для формы f существует ортогональное преобразование $X=CY$, приводящее ее к виду (1). Докажем, что при этом предположении справедливы оба последних утверждения теоремы.

Согласно теореме 1 из § 18 мы имеем равенство

$$C'AC = \begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{Vmatrix}. \tag{2}$$

Характеристический многочлен матрицы справа (см. § 17), очевидно, равен $(\lambda_1-t)(\lambda_2-t)\ldots(\lambda_n-t)$. Найдем характеристический многочлен матрицы слева. Мы имеем

$$\begin{aligned} |C'AC-tE| &= |C'AC-C'(tE)C| = \\ &= |C'(A-tE)C| = |C'|\cdot|A-tE|\cdot|C| = \\ &= |C'|\cdot|C|\cdot\varphi(t) = |C'C|\cdot\varphi(t) = |E|\cdot\varphi(t) = \varphi(t). \end{aligned}$$

(При выкладках мы воспользовались равенством $C'C=E$, в силу которого $C'(tE)C=t(C'EC)=t(C'C)=tE$, затем теоремой 1 из § 14 и теоремой из § 15.) Таким образом,

$$\varphi(t) = (\lambda_1 - t)(\lambda_2 - t) \dots (\lambda_n - t),$$

а значит, числа $\lambda_1, \lambda_2, \dots, \lambda_n$ являются корнями характеристического многочлена $\varphi(t)$ матрицы A. Этим доказано второе утверждение теоремы.

Обратимся теперь к равенству (2). Так как для ортогональной матрицы $C' = C^{-1}$, то, умножая (2) на C слева, мы получим

$$AC = C \begin{Vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \lambda_n \end{Vmatrix}.$$

Выполним умножения слева и справа, выразив матрицу C через ее столбцы: $C = \| T_1\ T_2 \dots T_n \|$. Согласно изложенному в § 13 произведение слева может быть представлено в виде совокупности столбцов $\| AT_1\ AT_2 \dots AT_n \|$. Что касается произведения справа, то оно, как легко видеть, представляется в виде совокупности столбцов следующим образом: $\| \lambda_1 T_1\ \lambda_2 T_2 \dots \lambda_n T_n \|$. Из равенства

$$\| AT_1\ AT_2 \dots AT_n \| = \| \lambda_1 T_1\ \lambda_2 T_2 \dots \lambda_n T_n \|$$

получаем теперь, приравнивая соответствующие столбцы, что

$$AT_1 = \lambda_1 T_1,\ AT_2 = \lambda_2 T_2,\ \dots,\ AT_n = \lambda_n T_n. \tag{3}$$

Мы доказали, таким образом, последнее утверждение теоремы 1.

2°. Доказательство существования ортогонального преобразования, приводящего квадратичную форму к диагональному виду, мы проведем индукцией по числу переменных n. При $n = 1$ доказывать нечего, так как всякая форма от одной переменной диагональна. Пусть $n \geqslant 2$, и пусть для форм от $n - 1$ переменных теорема уже доказана. Обозначим через λ какое-нибудь собственное число матрицы A, т. е. матрицы нашей квадратичной формы $f(x_1, \dots, x_n)$ от n переменных. Так как матрица A вещественна и симметрична, то, по теореме 2 из § 17, число λ вещественно. Найдем какой-нибудь ненулевой вещественный собственный столбец T матрицы A, соответствующий собственному числу λ, т. е. такой столбец, что

$$AT = \lambda T. \tag{4}$$

О существовании T свидетельствуют результаты § 17. Домножив T на надлежащий множитель, мы можем превратить его в нормированный столбец, не нарушая при этом равенства (4). Можно поэтому считать, что столбец T в (4) нормированный, т. е. что

$$T'T = 1. \tag{5}$$

Согласно следствию теоремы 4 из § 20 мы можем построить ортогональную матрицу $C_1 = \|T \ T_2 \ \dots \ T_n\|$, первый столбец которой совпадает с T. Выполним ортогональное преобразование переменных $X = C_1 Y$ (с ортогональной матрицей C_1). После преобразования форма $f(x_1, \dots, x_n)$ перейдет в форму, матрица которой равна $C_1' A C_1$ (теорема 1 из § 18). Найдем в этой матрице первый столбец (в силу симметричности матрицы мы тем самым будем знать и первую строчку):

$$C_1'AC_1 = C_1'A\|T \ T_2 \dots T_n\| = C_1'\|AT \ AT_2 \dots AT_n\| =$$

$$= \begin{Vmatrix} T' \\ T_2' \\ \cdot \\ \cdot \\ \cdot \\ T_n' \end{Vmatrix} \|\lambda T \ \dots\| = \begin{Vmatrix} \lambda T'T & \dots \\ \lambda T_2'T & \dots \\ \cdot & \\ \cdot & \\ \cdot & \\ \lambda T_n'T & \dots \end{Vmatrix} =$$

$$= \begin{Vmatrix} \lambda & \dots \\ 0 & \dots \\ \cdot & \\ \cdot & \\ \cdot & \\ 0 & \dots \end{Vmatrix} = \left\Vert \begin{array}{c|ccc} \lambda & 0 & \dots & 0 \\ \hline 0 & & & \\ \cdot & & & \\ \cdot & & * & \\ \cdot & & & \\ 0 & & & \end{array} \right\Vert .$$

При выкладках, наряду с правилом умножения матриц (см. § 13), мы воспользовались равенствами (4) и (5), а также тем, что $T_2'T = \dots = T_n'T = 0$ (последнее имеет место в силу ортогональности столбцов матрицы C_1). После преобразования $X = C_1 Y$ форма f примет, следовательно, вид

$$f = \lambda y_1^2 + g(y_2, \dots, y_n), \tag{6}$$

где g — некоторая (вещественная) квадратичная форма от $n-1$ переменных. К форме g мы можем применить индукционное предположение. Значит, существует такая ортогональная матрица D порядка $n-1$, что после преобразования

$$\begin{Vmatrix} y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{Vmatrix} = D \begin{Vmatrix} z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{Vmatrix}$$

получим

$$g(y_2, \ldots, y_n) = \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2.$$

Рассмотрим матрицу n-го порядка

$$C_2 = \begin{Vmatrix} 1 & 0 \ldots 0 \\ 0 & \\ \cdot & \\ \cdot & D \\ \cdot & \\ 0 & \end{Vmatrix}.$$

Так как все ее столбцы, очевидно, нормированы и попарно ортогональны, то она является ортогональной матрицей. После ортогонального преобразования $Y = C_2 Z$ форма (6) перейдет, как легко видеть, в форму

$$\lambda z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2. \tag{7}$$

Итак, после двух последовательно выполненных ортогональных преобразований переменных форма f приняла диагональный вид (7). Остается сделать ссылку на следствие теоремы 3 из § 20, и теорема 1 доказана.

З а м е ч а н и е 1. Из пункта 1° доказательства теоремы 1 следует, что если мы построили ортогональную матрицу C, все столбцы которой являются собственными столбцами матрицы A, то преобразование $X = CY$ и будет тем ортогональным преобразованием, которое приводит форму f к диагональному виду. В самом деле, из равенств (3) при обратном рассуждении легко по лучить равенство (2).

З а м е ч а н и е 2. При построении столбцов ортогональной матрицы C мы должны наряду с условием нормированности столбцов соблюсти и условие их попарной ортогональности. Однако когда мы строим матрицу C из собственных столбцов вещественной симметрической матрицы A, то за условием ортогональности надо следить только для собственных столбцов, отвечающих одному и тому же собственному числу λ. Действительно, согласно теореме 3 из § 17 собственные столбцы вещественной симметрической матрицы, соответствующие различным собственным числам, автоматически ортогональны.

Пример. Квадратичную форму

$$f = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 4x_1x_2 + 4x_1x_3 + 4x_1x_4 + 4x_2x_3 + \\ + 4x_2x_4 + 4x_3x_4$$

привести к диагональному виду ортогональным преобразованием переменных.

Составляем матрицу A квадратичной формы f и находим для нее характеристический многочлен. В данном случае он равен

$$\varphi(t) = (7 - t)(-1 - t)^3.$$

Следовательно, после надлежащего ортогонального преобразования $X = CY$ форма примет вид

$$7y_1^2 - y_2^2 - y_3^2 - y_4^2. \tag{8}$$

Для каждого собственного числа находим теперь все собственные столбцы матрицы A.

$$\text{Для } \lambda = 7 \quad \begin{Vmatrix} \alpha \\ \alpha \\ \alpha \\ \alpha \end{Vmatrix}; \quad \text{для } \lambda = -1 \quad \begin{Vmatrix} \alpha \\ \beta \\ \gamma \\ -\alpha - \beta - \gamma \end{Vmatrix},$$

где α, β, γ произвольны. Из первой серии нам надо выбрать один столбец, из второй — три. Выбирая ненулевые столбцы, вначале следим только за ортогональностью. Согласно замечанию 2 ортогональность столбцов из различных серий автоматически имеет место. Из второй серии первый столбец мы можем взять произвольно (только бы он не был нулевым). Второй столбец выбираем так, чтобы он был ортогонален первому. Наконец, третий надо выбрать так, чтобы он был ортогонален первым двум. В качестве искомых четырех столбцов можно взять, например,

$$\begin{Vmatrix}1\\1\\1\\1\end{Vmatrix},\quad \begin{Vmatrix}1\\1\\-1\\-1\end{Vmatrix},\quad \begin{Vmatrix}1\\-1\\1\\-1\end{Vmatrix},\quad \begin{Vmatrix}1\\-1\\-1\\1\end{Vmatrix}.$$

Умножая каждый из найденных столбцов на подходящий множитель, мы превращаем их в нормированные столбцы (в данном случае все столбцы надо умножить на $\frac{1}{2}$). Расположив затем эти нормированные и попарно ортогональные собственные столбцы матрицы A в соответствии с расположением отвечающих им собственных чисел в диагональной форме (8), мы получим ортогональную матрицу C, для которой преобразование $X=CY$ и осуществляет переход от заданной формы f к форме (8).

Ортогональным преобразованием, переводящим форму f в диагональную форму (8), таким образом, является преобразование

$$\left.\begin{aligned} x_1&=\frac{1}{2}y_1+\frac{1}{2}y_2+\frac{1}{2}y_3+\frac{1}{2}y_4,\\ x_2&=\frac{1}{2}y_1+\frac{1}{2}y_2-\frac{1}{2}y_3-\frac{1}{2}y_4,\\ x_3&=\frac{1}{2}y_1-\frac{1}{2}y_2+\frac{1}{2}y_3-\frac{1}{2}y_4,\\ x_4&=\frac{1}{2}y_1-\frac{1}{2}y_2-\frac{1}{2}y_3+\frac{1}{2}y_4.\end{aligned}\right\}\tag{9}$$

До сих пор мы занимались вопросом приведения к диагональному виду только одной квадратичной формы. Естественно поставить вопрос, нельзя ли две квадратичные формы одновременно одним и тем же неособенным преобразованием переменных привести к диагональному виду. В общем случае ответ оказывается отрицательным. Однако в некоторых частных случаях такое одновременное приведение возможно. Оно возможно, например, если одна из форм положительно определенная.

Теорема 2. *Пусть даны две вещественные квадратичные формы $f(x_1, \ldots, x_n)$ и $g(x_1, \ldots, x_n)$, из которых* *одна положительно определенная. Тогда существует такое вещественное неособенное линейное преобразование переменных, при котором обе формы одновременно приводятся к диагональному виду.*

Доказательство. Предположим, что форма f положительно определенная. По теореме 3 из § 19 при помощи некоторого неособенного вещественного преобразования $X = BY$ форма f приводится к чистой сумме квадратов:

$$f = y_1^2 + y_2^2 + \ldots + y_n^2. \tag{10}$$

Подвергая тому же преобразованию $X = BY$ форму g, мы получим некоторую форму h от новых переменных:

$$g(x_1, \ldots, x_n) = h(y_1, \ldots, y_n).$$

По теореме 1 при помощи надлежащего ортогонального преобразования переменных $Y = CZ$ форма h может быть приведена к диагональному виду:

$$h = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2.$$

При том же преобразовании $Y = CZ$ форма (10), согласно теореме 2 из § 20, перейдет в чистую сумму квадратов.

Таким образом, после вещественного неособенного преобразования $X = (BC) Z$ обе формы f и g одновременно приняли диагональный вид:

$$f = z_1^2 + z_2^2 + \ldots + z_n^2,$$
$$g = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \ldots + \lambda_n z_n^2.$$

Теорема 2 доказана.

Упражнения

1. При помощи ортогонального преобразования переменных привести к диагональному виду квадратичные формы:

a) $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 - 2x_1x_4 - 2x_2x_3 + 2x_3x_4$;

b) $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 + 2x_3x_4$;

c) $4x_1x_2 - 2x_1x_4 - 2x_2x_3 + 4x_3x_4$;

d) $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4$.

Отв. Преобразованием (9) каждая из форм приводится к виду:

a) $y_1^2 + 3y_2^2 + y_3^2 - y_4^2$; b) $2y_1^2 + 2y_2^2$;

c) $y_1^2 + 3y_2^2 - y_3^2 - 3y_4^2$; d) $2y_1^2 + 2y_2^2 + 2y_3^2 - 2y_4^2$.

2) Квадратичную форму

$$x_1^2 + 2x_3^2 - 4x_1x_2 + 4x_1x_3$$

привести к диагональному виду ортогональным преобразованием переменных.

Отв. $y_1^2 + 4y_2^2 - 2y_3^2$; $\left.\begin{aligned} x_1 &= -\frac{1}{3}y_1 + \frac{2}{3}y_2 + \frac{2}{3}y_3, \\ x_2 &= \frac{2}{3}y_1 - \frac{1}{3}y_2 + \frac{2}{3}y_3, \\ x_3 &= \frac{2}{3}y_1 + \frac{2}{3}y_2 - \frac{1}{3}y_3. \end{aligned}\right\}$

3. Пусть все собственные числа вещественной симметричной матрицы A больше вещественного μ. Доказать, что квадратичная форма с матрицей $A - \mu E$ положительно определенная.

4. Пусть даны две вещественные квадратичные формы. Доказать, что одна из них может быть преобразована в другую ортогональным преобразованием переменных тогда и только тогда, когда характеристические многочлены их матриц совпадают.

§ 22. Приведение к каноническому виду общего уравнения линии и поверхности 2-го порядка

Общим уравнением второй степени называется уравнение вида

$$F(x_1, x_2, \ldots, x_n) = 0, \tag{1}$$

где $F(x_1, \ldots, x_n)$ — многочлен второй степени от n переменных $x_1, x_2, \ldots, x_n$. Многочлен F мы можем, очевидно, представить в виде суммы

$$F(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) + 2\sum_{i=1}^{n} b_i x_i + c,$$

т. е. в виде суммы квадратичной формы f, линейной формы (однородного многочлена первой степени) и свободного члена c. Если через A мы обозначим матрицу квадратичной формы f, через X — столбец из переменных $x_1, x_2, \ldots, x_n$ и через T — столбец из коэффициентов $b_1, b_2, \ldots, b_n$, то, пользуясь матричной записью, мы сможем многочлен F записать в виде

$$F(x_1, \ldots, x_n) = X'AX + 2X'T + c. \tag{2}$$

Имея в виду приложения к аналитической геометрии, мы будем считать коэффициенты многочлена F вещественными. Заметим, что матрица A ненулевая, так как в противном случае многочлен (2) имел бы степень < 2.

Если $n=2$, то уравнение (1) является уравнением линии 2-го порядка на плоскости в декартовой прямоугольной системе координат. При $n=3$ уравнение (1) есть уравнение поверхности 2-го порядка в пространстве. Для общего уравнения линии на плоскости или поверхности в пространстве в аналитической геометрии ставится задача отыскания новой прямоугольной системы координат (с одним и тем же масштабом) так, чтобы уравнение данной линии или поверхности приняло наиболее простой вид.

Переход от одной прямоугольной системы координат к другой (как на плоскости, так и в пространстве) может быть осуществлен в два приема: 1) путем параллельного переноса начала координат в новую точку с сохранением направления осей и 2) путем поворотов осей при сохранении начала.

При параллельном переносе начала формулы преобразования координат, выражающие старые координаты произвольной точки через ее новые координаты, имеют вид

$$x_i = v_i + y_i \qquad (1 \leqslant i \leqslant n), \tag{3}$$

где v_i — координаты нового начала относительно старой системы координат (здесь n равно 2 или 3). Если через X мы обозначим столбец из старых координат x_i, через Y — столбец из новых координат y_i и через V — столбец из координат v_i нового начала, то формулы (3) можно будет записать в виде одного равенства:

$$X = V + Y. \tag{4}$$

При поворотах осей вокруг неподвижного начала формулы преобразования координат, как мы видели в § 20, в матричной форме могут быть записаны в виде

$$X = CY, \tag{5}$$

где C — ортогональная матрица (2-го или 3-го порядка).

Приведение уравнения (1) к простому (так называемому каноническому) виду с алгебраической точки зрения сводится, таким образом, к упрощению многочлена (2) при помощи преобразований переменных вида

(4) и (5). Мы сейчас решим эту задачу при произвольном n.

Прежде всего возникает вопрос, нельзя ли избавиться от членов, содержащих переменные в первой степени. При помощи преобразования (5) этого сделать нельзя, так как для ненулевого столбца T ввиду неособенности матрицы C' столбец $C'T$ всегда ненулевой, а значит,

$$X'T=(CY)'T=Y'C'T\neq 0$$

(форма $X'AX$ при преобразовании (5) переходит, как мы знаем, опять в квадратичную форму).

Подвергнем многочлен (2) преобразованию (4) и посмотрим, нельзя ли столбец V подобрать так, чтобы исчезла линейная часть:

$$F=(V'+Y')A(V+Y)+2(V'+Y')T+c=$$
$$=Y'AY+Y'AV+V'AY+V'AV+2V'T+2Y'T+c.$$

Так как $V'AY=(V'AY)'=Y'A'V''=Y'AV$, то

$$F=Y'AY+2Y'(AV+T)+c_1, \qquad (6)$$

где $c_1=V'AV+2V'T+c$. Отсюда видим, что первые степени новых переменных будут отсутствовать, если столбец V удовлетворяет условию

$$AV+T=O \qquad (7)$$

(справа стоит нулевой столбец). Определение столбца V сводится, как видим, к решению системы n линейных уравнений с n неизвестными. Поскольку вещественная симметричная матрица A и столбец T могут быть произвольными, то система (7) может оказаться как совместной, так и несовместной. (Совместность системы (7) при любом столбце T мы можем гарантировать лишь при неособенной матрице A.) В нашей задаче приведения многочлена F к наиболее простому виду мы должны поэтому различать два случая в зависимости от того, будет ли матричное уравнение (7) разрешимо относительно столбца V или нет.

Предположим сначала, что уравнение (7) разрешимо. Тогда после преобразования переноса $X=V+Y$ многочлен F ввиду (6) примет вид

$$F=Y'AY+c_1.$$

Сделаем теперь вслед за переносом ортогональное преобразование переменных $Y=CZ$, приводящее квадратичную форму $Y'AY$ к диагональному виду. Наши два последовательно выполненных преобразования $X=V+Y$ и $Y=CZ$ равносильны одному преобразованию

$$X=V+CZ. \tag{8}$$

В результате этого преобразования многочлен F принимает вид

$$F=\lambda_1 z_1^2+\lambda_2 z_2^2+\ldots+\lambda_n z_n^2+c_1. \tag{9}$$

(Некоторые из коэффициентов λ_i, но не все, могут, конечно, случайно оказаться равными нулю.) Сформулируем полученный нами результат.

Теорема 1. *Пусть дан многочлен второй степени от n переменных, записанный в виде* (2). *Если уравнение* (7) *разрешимо относительно столбца V, то при помощи преобразования переменных вида* (8) *с ортогональной матрицей C многочлен F приводится к каноническому виду* (9).

Рассмотренный нами случай (разрешимость уравнения (7)) соответствует в аналитической геометрии так называемым *центральным* линиям и поверхностям 2-го порядка. Начало новой координатной системы находится в центре линии или поверхности, так как для всякой точки $(z_1, \ldots, z_n)$, удовлетворяющей уравнению $F=0$, симметричная с ней относительно начала точка $(-z_1, \ldots, -z_n)$ также удовлетворяет этому уравнению.

Обратимся теперь ко второму случаю, когда уравнение (7) неразрешимо. Как уже было отмечено, это может случиться только при особенной матрице A. Для особенной матрицы A существует ненулевой столбец T_0 такой, что $AT_0=O$ (теорема 2 из § 12). Переписав последнее равенство в виде $AT_0=0\cdot T_0$, убеждаемся, что нуль является собственным числом матрицы A.

Подвергнем переменные многочлена (2) ортогональному преобразованию $X=CY$ так, чтобы квадратичная форма $X'AX$ приняла диагональный вид. Так как среди собственных чисел матрицы A имеются нулевые, часть коэффициентов при квадратах новых переменных будет равна нулю. Опустим эти квадраты с нулевыми коэффициентами. Мы получим для многочлена (2) в новых переменных выражение

$$F=\lambda_1 y_1^2+\ldots+\lambda_r y_r^2+2\sum_{i=1}^{n} c_i y_i+c,$$

где $r<n$ и $\lambda_1\neq 0, \ldots, \lambda_r\neq 0$ (столбец из c_i совпадает со столбцом $C'T$). При помощи параллельного переноса здесь можно избавиться теперь от первых степеней переменных $y_1, \ldots, y_r$. Действительно,

$$\lambda_i y_i^2+2c_i y_i=\lambda_i\left(y_i+\frac{c_i}{\lambda_i}\right)^2-\frac{c_i^2}{\lambda_i} \qquad (1\leqslant i\leqslant r),$$

поэтому после параллельного переноса

$$\left.\begin{aligned} y_i&=-\frac{c_i}{\lambda_i}+z_i \quad (1\leqslant i\leqslant r),\\ y_j&=z_j \qquad\quad (r+1\leqslant j\leqslant n) \end{aligned}\right\} \tag{10}$$

многочлен F примет вид

$$F=\lambda_1 z_1^2+\ldots+\lambda_r z_r^2+2\sum_{j=r+1}^{n} c_j z_j+d, \tag{11}$$

где

$$d=c-\sum_{i=1}^{r}\frac{c_i^2}{\lambda_i}.$$

Коэффициенты $c_{r+1}, \ldots, c_n$ не могут быть все равными нулю. В самом деле, если бы $c_j=0$ при $r+1\leqslant j\leqslant n$, то это означало бы по доказанному выше, что разрешимо уравнение $C'ACV+C'T=O$, а тогда и уравнение (7), как легко видеть, было бы разрешимым вопреки нашему предположению.

Положим $\mu=\sqrt{c_{r+1}^2+\ldots+c_n^2}$. Так как сумма квадратов элементов

$$\frac{c_{r+1}}{\mu}, \ldots, \frac{c_n}{\mu} \tag{12}$$

равна 1, то, по следствию теоремы 4 из § 20, существует ортогональная матрица D_0 порядка $n-r$, первый столбец которой состоит из чисел (12). Составим клеточно-диагональную матрицу

$$D=\left\|\begin{array}{c|c} E & \\ \hline & D_0 \end{array}\right\|,$$

где E — единичная матрица порядка r. Легко видеть, что D — ортогональная матрица (ее столбцы нормированы и попарно ортогональны). Сделаем в выражении (11) ортогональное преобразование переменных $Z = DU$. Так как обратное преобразование $U = D^{-1}Z = D'Z$ имеет вид

$$\left.\begin{array}{l} u_1 = z_1, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ u_r = \quad z_r, \\ u_{r+1} = \quad \frac{c_{r+1}}{\mu} z_{r+1} + \ldots + \frac{c_n}{\mu} z_n, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ , \end{array}\right\}$$

то после преобразования получим

$$F = \lambda_1 u_1^2 + \ldots + \lambda_r u_r^2 + 2\mu u_{r+1} + d.$$

Нашим последним преобразованием будет параллельный перенос

$$\left.\begin{array}{l} u_{r+1} = -\frac{d}{2\mu} + w_{r+1}, \\ u_i = w_i \quad (i \neq r+1). \end{array}\right\} \tag{13}$$

После его выполнения получим

$$F = \lambda_1 w_1^2 + \ldots + \lambda_r w_r^2 + 2\mu w_{r+1}. \tag{14}$$

Чтобы перейти от (2) к (14), мы выполнили вперемежку два ортогональных преобразования $X = CY$ и $Z = DU$ и два переноса (10) и (13). Запишем эти переносы в виде $Y = V_1 + Z$ и $U = V_2 + W$ и найдем вид окончательного преобразования, выражающего x_i через w_j. Мы имеем

$$X = CY = C(V_1 + Z) = CV_1 + CDU = CV_1 + \\ + CD(V_2 + W) = CV_1 + CDV_2 + CDW = V_0 + CDW,$$

где V_0 — некоторый столбец. Мы убедились, таким образом, что результирующее преобразование $X = V_0 + CDW$ в рассматриваемом случае также имеет вид (8), т. е. может быть разбито только на два преобразования — параллельный перенос с последующим ортогональным преобразованием.

Нами получена следующая теорема.

Т е о р е м а 2. *Пусть для многочлена* (2) *уравнение* (7) *неразрешимо. Тогда при помощи преобразования переменных вида* (8) *с ортогональной матрицей* C *этот многочлен может быть приведен к виду*

$$F = \lambda_1 z_1^2 + \ldots + \lambda_r z_r^2 + 2\mu z_{r+1}$$
$$(\lambda_1 \neq 0, \ \ldots, \ \lambda_r \neq 0; \ \mu \neq 0; \ 1 \leqslant r < n).$$

Случай теоремы 2 соответствует так называемым *нецентральным* линиям и поверхностям 2-го порядка.

Применим доказанные теоремы к изучению общего уравнения 2-го порядка на плоскости. Согласно теоремам 1 и 2 каждое уравнение линии 2-го порядка в надлежащей системе координат имеет один из следующих видов:

$$\lambda_1 x^2 + \lambda_2 y^2 + c = 0 \qquad (\lambda_1 \neq 0),$$
$$\lambda x^2 = -2\mu y \quad (\lambda \neq 0, \quad \mu \neq 0).$$

Перебирая возможные нулевые значения для λ_2 и c и учитывая знаки встречающихся коэффициентов, мы приходим к следующим геометрическим образам (α и β положительны):

1) $\alpha x^2 = 0$ — одна прямая;
2) $\alpha x^2 + \beta y^2 = 0$ — одна точка;
3) $\alpha x^2 - \beta y^2 = 0$ — две пересекающиеся прямые;
4) $\alpha x^2 = 1$ — две параллельные прямые;
5) $-\alpha x^2 = 1$ — пустое множество;
6) $\alpha x^2 + \beta y^2 = 1$ — эллипс;
7) $\alpha x^2 - \beta y^2 = 1$ — гипербола;
8) $-\alpha x^2 - \beta y^2 = 1$ — пустое множество;
9) $\alpha x^2 = y$ — парабола

(в случае 9 надо, быть может, дополнительно изменить направления осей).

Аналогично общее уравнение 2-го порядка в пространстве определяет один из следующих геометрических образов (α, β, γ положительны):

1) $\alpha x^2 = 0$ — одна плоскость;
2) $\alpha x^2 + \beta y^2 = 0$ — одна прямая;
3) $\alpha x^2 - \beta y^2 = 0$ — две пересекающиеся плоскости;
4) $\alpha x^2 + \beta y^2 + \gamma z^2 = 0$ — одна точка;
5) $\alpha x^2 + \beta y^2 - \gamma z^2 = 0$ — конус;
6) $\alpha x^2 = 1$ — две параллельные плоскости;
7) $-\alpha x^2 = 1$ — пустое множество;

8) $\alpha x^2+\beta y^2=1$ — эллиптический цилиндр;
9) $\alpha x^2-\beta y^2=1$ — гиперболический цилиндр;
10) $-\alpha x^2-\beta y^2=1$ — пустое множество;
11) $\alpha x^2+\beta y^2+\gamma z^2=1$ — эллипсоид;
12) $\alpha x^2+\beta y^2-\gamma z^2=1$ — однополостный гиперболоид;
13) $\alpha x^2-\beta y^2-\gamma z^2=1$ — двуполостный гиперболоид;
14) $-\alpha x^2-\beta y^2-\gamma z^2=1$ — пустое множество;
15) $\alpha x^2=y$ — параболический цилиндр;
16) $\alpha x^2+\beta y^2=z$ — эллиптический параболоид;
17) $\alpha x^2-\beta y^2=z$ — гиперболический параболоид.

Пример 1. Привести к каноническому виду уравнение поверхности

$$2x_1^2+x_2^2-4x_1x_2-4x_2x_3+12x_1-8x_2+8x_3+6=0.$$

Уравнение (7) здесь разрешимо: $v_1=-1$, $v_2=2$, $v_3=0$. Делаем параллельный перенос по формулам (4): $x_1=-1+y_1$, $x_2=2+y_2$, $x_3=y_3$. Уравнение принимает вид

$$2y_1^2+y_2^2-4y_1y_2-4y_2y_3-8=0.$$

При помощи ортогонального преобразования

$$\left.\begin{aligned} y_1 &= \frac{2}{3}z_1+\frac{2}{3}z_2+\frac{1}{3}z_3, \\ y_2 &= -\frac{2}{3}z_1+\frac{1}{3}z_2+\frac{2}{3}z_3, \\ y_3 &= \frac{1}{3}z_1-\frac{2}{3}z_2+\frac{2}{3}z_3 \end{aligned}\right\}$$

квадратичная форма в левой части уравнения приводится к виду $4z_1^2+z_2^2-2z_3^2$. Следовательно, в новой координатной системе уравнением поверхности будет

$$\frac{z_1^2}{2}+\frac{z_2^2}{8}-\frac{z_3^2}{4}=1$$

(однополостный гиперболоид).

Пример 2. Привести к каноническому виду уравнение поверхности

$$2x_1^2+2x_2^2+3x_3^2+4x_1x_2+2x_1x_3+2x_2x_3+10x_2+1=0.$$

Уравнение (7) в данном случае не имеет решения — значит, поверхность нецентральная. Выполним ортогональное преобразование, приводящее квадратичную форму к диагональному виду:

$$\left.\begin{aligned} x_1 &= \frac{1}{\sqrt{3}}y_1 + \frac{1}{\sqrt{6}}y_2 + \frac{1}{\sqrt{2}}y_3, \\ x_2 &= \frac{1}{\sqrt{3}}y_1 + \frac{1}{\sqrt{6}}y_2 - \frac{1}{\sqrt{2}}y_3, \\ x_3 &= \frac{1}{\sqrt{3}}y_1 - \frac{2}{\sqrt{6}}y_2. \end{aligned}\right\}$$

После этого преобразования уравнение принимает вид

$$5y_1^2 + 2y_2^2 + 10\left(\frac{1}{\sqrt{3}}y_1 + \frac{1}{\sqrt{6}}y_2 - \frac{1}{\sqrt{2}}y_3\right) + 1 = 0.$$

Делаем теперь параллельный перенос, чтобы избавиться от членов с y_1 и y_2 и свободного члена:

$$y_1 = -\frac{1}{\sqrt{3}} + z_1, \quad y_2 = -\frac{5}{2\sqrt{6}} + z_2, \quad y_3 = -\frac{11\sqrt{2}}{40} + z_3.$$

Окончательно уравнение принимает вид

$$5z_1^2 + 2z_2^2 = 5\sqrt{2}\,z_3$$

(эллиптический параболоид).

Упражнения

1. Привести к каноническому виду уравнение линии

$$x^2 + y^2 + 3xy + x + 4y = 0.$$

Отв. $-\frac{5}{2}x'^2 + \frac{1}{2}y'^2 = 1;\quad x = -2 + \frac{x'+y'}{\sqrt{2}},\ y = 1 + \frac{x'-y'}{\sqrt{2}}.$

2. Убедиться, что линия, определяемая уравнением

$$9x^2 - 6xy + y^2 - x + 2y - 14 = 0,$$

является параболой.

Зенон Иванович Боревич

Определители и матрицы

М., 1970 г., 200 стр. с илл.

Редактор *А. Ф. Лапко*
Техн. редактор *В. Н. Кондакова*
Корректор *Т. С. Плетнева*

Сдано в набор 13/VI 1969 г. Подписано к печати 3/XII 1969 г. Бумага 84×108$^{1}/_{32}$. Физ. печ. л. 6,25. Условн. печ. л. 10,5. Уч.-изд. л. 9,11. Тираж 58 000 экз. Т-16035. Цена книги 32 коп. Заказ № 616.

Издательство «Наука»
Главная редакция
физико-математической литературы
Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени Ленинградская типография № 1 «Печатный Двор» имени А. М. Горького Главполиграфпрома Комитета по печати при Совете Министров СССР, Ленинград, Гатчинская ул., 26.

Цена 3~~2 к.~~

30к