ЛИНЕЙНАЯ АЛГЕБРА

Выпуск I

Учебно-методическое пособие для студентов I *курса*

Линейная алгебра возникла как наука о решении систем линейных алгебраических уравнений. Впоследствии предмет линейной алгебры расширился, и сейчас она в существенном представляет собой теорию линейных преобразований (операторов) в конечномерных векторных пространствах. Точный смысл сказанного станет ясен по мере прохождения курса, и мы не будем входить в дальнейшие пояснения.

Предлагаемое пособие представляет собой первый выпуск по курсу линейной алгебры для студентов физического факультета СПбГУ. Этот выпуск содержит часть главы 1 курса — "Матрицы и определители". Второй выпуск будет содержать оставшуюся часть главы 1 и главу 2 курса — "Системы линейных алгебраических уравнений". Авторы надеются впоследствии издать дальнейшие выпуски пособия с тем, чтобы охватить весь лекционный курс линейной алгебры.

Предполагается, что настоящее пособие будет служить лишь материалом для повторения. Работа над ним не может заменить систематического слушания лекционного курса. Поэтому различные пояснения и мотивировки сведены в тексте до минимума. Авторы считают, что строгий отбор материала и сравнительно небольшой объем пособия создаст удобства для слушателей курса и для читателей.

Ниже обозначение $a := b$ означает, что "a по определению равно b". Значок • означает конец доказательства. Дополнительный (необязательный) материал помечен верхним значком $*$. Как правило, этот материал не входит в лекционный курс. При ссылках на формулы, теоремы и пункты из другого параграфа применяется двойная нумерация, а из другой главы — тройная.

Глава 1. Матрицы и определители

В этой главе мы познакомимся с формальным аппаратом, используемым в линейной алгебре, — с алгеброй матриц. При таком "предварительном" введении понятий матричной алгебры определения могут выглядеть недостаточно мотивированными. Однако их смысл проясняется в дальнейшем изложении курса.

§ 1. Действия над матрицами

1. Определение матрицы. Матрицей называют прямоугольную таблицу чисел (вещественных или комплексных). Эти числа[1] называют *элементами* матрицы. Матрицу будем записывать следующим способом

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{1}$$

Элементы a_{ij} нумеруются двумя индексами; первый из них есть номер строки и меняется вдоль столбца, второй — номер столбца, который меняется вдоль строки. Для матрицы (1) употребляется также краткое обозначение, которое явно указывает на ее размеры:

$$A = \{a_{ij}\}_{i=1,\ldots,m}^{j=1,\ldots,n}. \tag{2}$$

Матрица, составленная из m строк и n столбцов, называется $(m \times n)$-матрицей. Такая матрица возникнет, например,

[1] Иногда рассматривают матрицы, составленные не из чисел, а из элементов другой природы.

при последовательном выписывании коэффициентов при неизвестных в системе из m линейных алгебраических уравнений с n неизвестными. Множество всех $(m \times n)$-матриц будем обозначать через $M^{m,n}$. Наконец, условимся для элемента a_{ij} матрицы A в некоторых случаях использовать обозначение $[A]_{ij}$:

$$a_{ij} = [A]_{ij}, \quad i = 1, \ldots, m; \; j = 1, \ldots, n.$$

2. Линейные действия над матрицами. Введем линейные действия над матрицами — сложение матриц и умножение матрицы на число.

Сложение матриц определяется только для матриц совпадающих размеров: если $A \in M^{m,n}$, $B \in M^{m,n}$, то $A + B \in M^{m,n}$, где

$$[A + B]_{ij} = [A]_{ij} + [B]_{ij}, \quad i = 1, \ldots, m; \; j = 1, \ldots, n. \tag{3}$$

Таким образом, сложение матриц состоит в поэлементном сложении.

Умножение матрицы на число состоит в умножении на это число каждого элемента матрицы. Для произведения матрицы A на число α используется как обозначение αA, так и обозначение $A\alpha$. Таким образом,

$$[\alpha A]_{ij} = [A\alpha]_{ij} = \alpha [A]_{ij}, \quad i = 1, \ldots, m; \; j = 1, \ldots, n. \tag{4}$$

Ясно, что $\alpha A \in M^{m,n}$, если $A \in M^{m,n}$. В подробных обозначениях

$$\alpha A = A\alpha = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}.$$

При этом предполагается, что в случае вещественных матриц их можно умножать на вещественные множители. В классе комплексных матриц подразумевается умножение на комплексные числа.

Перечислим *свойства линейных операций* в классе матриц $M^{m,n}$.

1. Коммутативность сложения:

$$A + B = B + A.$$

2. Ассоциативность сложения:

$$(A + B) + C = A + (B + C).$$

3. Существование нуля. Существует матрица $\mathbb{O} \in M^{m,n}$, называемая *нулевой матрицей*, такая, что

$$A + \mathbb{O} = A, \quad \forall A \in M^{m,n}.$$

Это матрица, все элементы которой равны нулю $[\mathbb{O}]_{ij} = 0$.

4. Существование *противоположной* матрицы. Для любой матрицы $A \in M^{m,n}$ существует матрица $\tilde{A} \in M^{m,n}$ такая, что

$$A + \tilde{A} = \mathbb{O}.$$

Матрица $\tilde{A}$, называемая противоположной к матрице A, обозначается $-A$. При этом $[\tilde{A}]_{ij} = [-A]_{ij} = -[A]_{ij}$.

5. Дистрибутивность умножения на число относительно сложения матриц:

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

6. Дистрибутивность умножения на матрицу относительно сложения чисел:

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

7. Ассоциативность относительно числовых множителей:

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

8. Особая роль числового множителя 1:

$$1 \cdot A = A, \quad \forall A \in M^{m,n}.$$

Проверка этих свойств не представляет труда, если использовать аналогичные свойства для отдельных элементов матрицы, то есть для чисел. Отметим, что именно эти восемь свойств линейных операций в дальнейшем кладутся в основу абстрактного понятия линейного пространства.

Разность матриц A и B определяется как сумма A и $-B$:

$$A - B := A + (-B).$$

3. Действия транспонирования и сопряжения. *Транспонированная матрица* получается из данной матрицы A, если строки и столбцы поменять ролями. Транспонированную матрицу обозначают через A^t. Если $A \in M^{m,n}$, то $A^t \in M^{n,m}$, причем

$$[A^t]_{ij} = [A]_{ji}, \quad i = 1, \ldots, n; \; j = 1, \ldots, m. \tag{5}$$

Отметим *свойство линейности операции транспонирования*: для любых матриц $A \in M^{m,n}$, $B \in M^{m,n}$ и любых чисел α, β,

$$(\alpha A + \beta B)^t = \alpha A^t + \beta B^t. \tag{6}$$

Ясно также, что $(A^t)^t = A$.

Для матриц с комплексными членами вводится понятие *сопряженной матрицы*. Сопряженная матрица A^* для данной матрицы A получается из транспонированной матрицы A^t, если все ее элементы заменить комплексно сопряженными числами. Таким образом, если $A \in M^{m,n}$, то $A^* \in M^{n,m}$, причем

$$[A^*]_{ij} = \overline{a_{ji}}, \quad i = 1, \ldots, n; \quad j = 1, \ldots, m. \tag{7}$$

Очевидно, для вещественных матриц транспонированная и сопряженная матрицы совпадают. Операция сопряжения обладает *свойством антилинейности*: для любых матриц $A \in M^{m,n}$, $B \in M^{m,n}$ и любых чисел $\alpha \in \mathbb{C}$, $\beta \in \mathbb{C}$,

$$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*. \tag{8}$$

Отметим также, что $A^{**} := (A^*)^* = A$.

4. Умножение матриц. Умножение матриц вводится лишь для матриц *согласованных порядков*: число столбцов первой матрицы должно совпадать с числом строк второй матрицы. По определению, для $A \in M^{m,n}$ и $B \in M^{n,p}$, их произведение $C = AB \in M^{m,p}$ есть матрица с элементами

$$c_{ij} = [AB]_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad i = 1, \ldots, m; \ j = 1, \ldots, p. \tag{9}$$

Суммирование[2] ведется по второму значку для элементов A и по первому для элементов B; согласованность порядков сказывается в том, что оба эти значка меняются от 1 до n. У матрицы C m строк (столько, сколько у A) и p столбцов (столько, сколько у B). Формулы (9) называют "*правилом умножения строки на столбец*". Именно, для получения элемента c_{ij}, стоящего в i-ой строке и j-ом столбце матрицы C, умножают k-ый элемент i-ой строки матрицы A на k-ый элемент j-го столбца матрицы B, и эти произведения суммируют [3] по k от 1 до n.

Пример: $m = 3, \ n = 2, \ p = 3,$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}.$$

---

[2] Важно помнить, что сумма не зависит от обозначения индекса (значка) суммирования. Значки суммирования иногда называют поэтому "немыми". В нашем случае немым значком является k.

[3] Суммы (9) образуются по правилу, которое при $n = 3$ формально совпадает с выражением скалярного произведения двух векторов через их декартовы координаты.

$$C = AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \end{pmatrix}.$$

Произведение $C = AB$ есть (3×3)-матрица.

Упражнение. Составьте (2×2)-матрицу $D = BA$.

Умножение матриц не коммутативно даже для (2×2)-матриц. Например, при

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathbb{O}.$$

Этот же пример показывает существование "*делителей нуля*" для матриц. Действительно, $A \neq \mathbb{O}$, $B \neq \mathbb{O}$, но $BA = \mathbb{O}$.

Исходя из определений (3)–(5), (7), (9) можно проверить справедливость следующих свойств, в формулировке которых порядки матриц предполагаются согласованными.

1. Дистрибутивность умножения относительно сложения:

$$(A + B)C = AC + BC, \quad A(C + D) = AC + AD. \tag{10}$$

2. Ассоциативность относительно умножения на число:

$$(\alpha A)C = \alpha(AC), \quad A(\beta C) = \beta(AC). \tag{11}$$

Свойства (10) и (11) вместе означают линейность произведения матриц по каждому из сомножителей:

$$(\alpha A + \beta B)C = \alpha(AC) + \beta(BC), \quad A(\alpha C + \beta D) = \alpha(AC) + \beta(AD). \tag{12}$$

3. Ассоциативность умножения матриц:

$$A(BC) = (AB)C. \tag{13}$$

4. Правило транспонирования произведения матриц:

$$(AB)^t = B^t A^t. \tag{14}$$

5. Правило сопряжения произведения матриц:

$$(AB)^* = B^* A^*. \tag{15}$$

Докажем свойства (13) и (14). Свойства (10), (11) и (15) докажите самостоятельно в качестве **упражнения**.

Доказательство ассоциативности умножения матриц. Пусть $A \in M^{m,n}$, $B \in M^{n,p}$, $C \in M^{p,r}$. Ясно, что тогда $A(BC) \in M^{m,r}$, $(AB)C \in M^{m,r}$. Вычислим сначала матрицу $A(BC)$. Имеем:

$$[BC]_{ik} = \sum_{j=1}^{p} b_{ij} c_{jk}, \quad i = 1, \ldots, n; \quad k = 1, \ldots, r,$$

$$[A(BC)]_{lk} = \sum_{i=1}^{n} a_{li} [BC]_{ik} = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{li} b_{ij} c_{jk}, \quad \left\{ \begin{array}{c} l = 1, \ldots, m; \\ k = 1, \ldots, r. \end{array} \right\} \tag{16}$$

Теперь вычислим $(AB)C$:

$$[AB]_{lj} = \sum_{i=1}^{n} a_{li} b_{ij}, \quad l = 1, \ldots, m; \quad j = 1, \ldots, p,$$

$$[(AB)C]_{lk} = \sum_{j=1}^{p} [AB]_{lj} c_{jk} = \sum_{j=1}^{p} \sum_{i=1}^{n} a_{li} b_{ij} c_{jk}, \quad \left\{ \begin{array}{c} l = 1, \ldots, m; \\ k = 1, \ldots, r. \end{array} \right\} \tag{17}$$

Мы видим, что в (16), (17) правые части отличаются лишь порядком суммирования. Однако, поскольку значки суммирования i, j меняются независимо друг от друга, порядок суммирования безразличен. Итак, матрицы $A(BC)$ и $(AB)C$ поэлементно совпадают. •

Доказательство правила транспонирования произведения матриц. При $A \in M^{m,n}$ и $B \in M^{n,p}$ будет $C = AB \in M^{m,p}$, $C^t \in M^{p,m}$. Имеем:

$$[C^t]_{ij} = [C]_{ji} = \sum_{k=1}^{n} [A]_{jk} [B]_{ki} = \sum_{k=1}^{n} [B^t]_{ik} [A^t]_{kj} = [B^t A^t]_{ij}, \quad i = 1, \ldots, p; \ j = 1, \ldots, m.$$

Это и означает справедливость (14). •

§ 2. Квадратные матрицы. След матрицы

1. Класс квадратных матриц. Рассмотрим матрицы класса $M^{n,n}$ (этот класс будем более коротко обозначать через M^n), то есть квадратные $(n \times n)$-матрицы. К таким матрицам применимы все введенные выше действия над матрицами, причем в результате снова получаются матрицы класса M^n. Особую роль в классе M^n играют нулевая матрица $\mathbb{O}$ и *единичная матрица*

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Иногда для I будем пользоваться обозначением $I = I_n$, указывая тем самым, что $I \in M^n$. На *главной диагонали* единичной матрицы стоят единицы, а все остальные элементы равны нулю. Элементы единичной матрицы совпадают с *символом Кронекера*, который определяется формулами

$$\delta_{ik} = \begin{cases} 1 & \text{при } i = k \\ 0 & \text{при } i \neq k \end{cases}.$$

Имеем: $[I]_{ik} = \delta_{ik}$. При любых $A \in M^n$ выполнено

$$\mathbb{O}A = A\mathbb{O} = \mathbb{O},$$

$$IA = AI = A.$$

Например,

$$[IA]_{ik} = \sum_{j=1}^{n} \delta_{ij} a_{jk} = a_{ik} = [A]_{ik}.$$

Матрица

$$\alpha I = \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha \end{pmatrix}$$

класса M^n называется *кратной единичной.*

Упражнения.

1. Пусть I — единичная матрица в M^n. Проверить, что при любом m выполнено равенство $IA = A$ для всех $A \in M^{n,m}$ и $BI = B$ для всех $B \in M^{m,n}$.

2^*. Пусть $A \in M^2$ — произвольная $(2 \times 2)-$матрица, $A^2 := AA$, $\alpha = a_{11} + a_{22}$, $\beta = a_{11}a_{22} - a_{12}a_{21}$. Проверьте справедливость тождества[4]

$$A^2 - \alpha A + \beta I = \mathbb{O}.$$

2. Коммутатор и антикоммутатор. Отметим, что матрица вида αI (в том числе, единичная и нулевая матрицы) коммутирует с любой матрицей $A \in M^n$. Как уже отмечалось, две произвольные матрицы $A \in M^n$, $B \in M^n$, вообще говоря, не коммутируют. *Коммутатором* матриц A и B называется матрица

$$[A, B] := AB - BA.$$

Равенство $[A, B] = \mathbb{O}$ означает, что матрицы A и B коммутируют.

Упражнение*. Если матрица $A \in M^n$ коммутирует с любой матрицей $B \in M^n$, то матрица A кратна единичной. Проверьте.

Вводят также понятие *антикоммутатора* матриц A и B — это матрица $\{A, B\} := AB + BA$. Если $\{A, B\} = \mathbb{O}$, то говорят, что матрицы A и B антикоммутируют. Отметим, что коммутатор и антикоммутатор линейны по каждому аргументу:

$$[\alpha A + \gamma C, B] = \alpha[A, B] + \gamma[C, B],$$

$$[A, \beta B + \gamma C] = \beta[A, B] + \gamma[A, C];$$

$$\{\alpha A + \gamma C, B\} = \alpha\{A, B\} + \gamma\{C, B\},$$

$$\{A, \beta B + \gamma C\} = \beta\{A, B\} + \gamma\{A, C\}.$$

---

[4] В выпуске 2 (см. п. 8.4) будет установлено более общее *тождество Кэли*, справедливое для матриц класса M^n, $n \geq 2$.

3. Треугольные и диагональные матрицы. Эрмитовы матрицы. В классе квадратных матриц M^n выделяют треугольные матрицы и диагональные матрицы. Матрицу $A \in M^n$ называют *верхнетреугольной*, если $a_{ij} = 0$ при $i > j$. Ниже главной диагонали в верхнетреугольной матрице стоят нули. В подробной записи такая матрица имеет вид

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Матрицу $A \in M^n$ называют *нижнетреугольной*, если $a_{ij} = 0$ при $i < j$. Если $A \in M^n$ и $B \in M^n$ — две верхнетреугольных матрицы, то $\alpha A + \beta B$ и AB также являютя верхнетреугольными матрицами. Докажите это самостоятельно в качестве **упражнения**. Аналогичный факт верен и для нижнетреугольных матриц. Операции транспонирования и сопряжения переводят верхнетреугольные матрицы в нижнетреугольные и наоборот.

Матрица A, которая одновременно является и верхнетреугольной и нижнетреугольной, называется *диагональной*. У такой матрицы все элементы вне главной диагонали равны нулю:

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Для диагональной матрицы используют краткое обозначение

$$A = \operatorname{diag}\{a_{11}, a_{22}, \cdots, a_{nn}\}.$$

Матрица αI является диагональной. Если $A \in M^n$, $B \in M^n$ — диагональные матрицы, то $\alpha A + \beta B$, AB, A^t, A^* — также диагональные матрицы. Более того, $A^t = A$. Проверьте, что при умножении диагональных матриц соответствующие элементы на главной диагонали перемножаются:

$$[AB]_{ii} = a_{ii} b_{ii}.$$

Отсюда сразу следует, что две диагональные матрицы всегда коммутируют.

Еще один важный класс квадратных матриц — *эрмитовы*[5] матрицы. Матрица $A \in M^n$ называется эрмитовой или *самосопряженной*, если $A^* = A$. Из определения A^* следует, что элементы эрмитовой матрицы удовлетворяют соотношениям $a_{ij} = \overline{a_{ji}}$, $i, j = 1, \ldots, n$. Это означает, что элементы на главной диагонали — вещественные, а элементы, расположенные симметрично относительно главной диагонали, комплексно сопряжены. Для (2×2)-матриц общий вид эрмитовой матрицы таков

$$A = \begin{pmatrix} \alpha & \overline{\beta} \\ \beta & \gamma \end{pmatrix}, \quad \alpha, \gamma \in \mathbb{R}, \quad \beta \in \mathbb{C}. \tag{1}$$

4. След. Каждой матрице $A \in M^n$ сопоставляют ее *след* — число, равное сумме элементов $[A]_{ii}$ главной диагонали матрицы A. Обозначение:

$$\operatorname{Tr} A = \sum_{i=1}^{n} [A]_{ii} = \sum_{i=1}^{n} a_{ii}, \quad A \in M^n. \tag{2}$$

(Символ Tr происходит от английского слова trace — след; иногда используют обозначение Sp — от немецкого Spur).

След обладает следующими свойствами.

1. Линейность следа:

$$\operatorname{Tr}(\alpha A + \beta B) = \alpha \operatorname{Tr} A + \beta \operatorname{Tr} B. \tag{3}$$

---

[5] Шарль Эрмит — знаменитый французский математик 19 века.

2. При транспонировании след не меняется:

$$\operatorname{Tr} A^t = \operatorname{Tr} A. \tag{4}$$

Эти свойства очевидны. Ясно также, что $\operatorname{Tr} A^* = \overline{\operatorname{Tr} A}$. Менее очевидным является следующее важное свойство.

3. *След произведения не зависит от порядка сомножителей*:

$$\operatorname{Tr} AB = \operatorname{Tr} BA, \quad A \in M^{m,n}, \quad B \in M^{n,m}. \tag{5}$$

Доказательство формулы (5). Вычислим сначала след матрицы $AB \in M^m$.

$$[AB]_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad i, j = 1, \ldots, m,$$

$$\operatorname{Tr} AB = \sum_{i=1}^{m} [AB]_{ii} = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{ki}. \tag{6}$$

Аналогично для матрицы $BA \in M^n$ имеем:

$$[BA]_{ij} = \sum_{k=1}^{m} b_{ik} a_{kj}, \quad i, j = 1, \ldots, n,$$

$$\operatorname{Tr} BA = \sum_{i=1}^{n} \sum_{k=1}^{m} b_{ik} a_{ki} = \sum_{k=1}^{m} \sum_{i=1}^{n} a_{ki} b_{ik}. \tag{7}$$

Индексы i, k в (7) — немые, поэтому их можно заменить любыми другими (но разными) буквами и, в частности, поменять индексы местами: $\operatorname{Tr} BA = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} b_{ki}$, что совпадает с (6). •

Упражнения. 1) Пусть $A \in M^{m,n}$. Проверьте, что след матрицы $A^t A$ (и след матрицы AA^t) есть сумма квадратов всех элементов матрицы A:

$$\operatorname{Tr} A^t A = \operatorname{Tr} AA^t = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik}^2.$$

Аналогично,

$$\operatorname{Tr} A^*A = \operatorname{Tr} AA^* = \sum_{i=1}^{m} \sum_{k=1}^{n} |a_{ik}|^2.$$

2) Пусть $A, B \in M^n$ и $[A, B] = \alpha I$. Докажите, что тогда $\alpha = 0$. *Указание*: вычислите след от обеих частей равенства.

Понятие следа будет неоднократно встречаться в дальнейших разделах курса.

5*. Матрицы Паули. Рассмотрим комплексные (2×2)-матрицы

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8}$$

Эти матрицы носят название матриц Паули[6] . Условимся нумеровать матрицы (8) индексом s "по модулю 3". Таким образом, $\sigma_4 = \sigma_1$, $\sigma_5 = \sigma_2$ и т. п. по определению. Непосредственно проверяются свойства

$$\sigma_s = \sigma_s^*, \quad s = 1, 2, 3, \tag{9}$$

то есть матрицы Паули эрмитовы;

$$\sigma_s^2 = I, \quad s = 1, 2, 3; \tag{10}$$

$$\sigma_s \sigma_{s+1} = i\sigma_{s+2}, \quad s = 1, 2, 3. \tag{11}$$

Из (9)–(11) видно, что

$$\{\sigma_s, \sigma_t\} = 2\delta_{st} I, \quad s, t = 1, 2, 3. \tag{12}$$

В частности, любые две различные матрицы (8) антикоммутируют. Из (10), (11) следует также, что

$$\sigma_1 \sigma_2 \sigma_3 = iI. \tag{13}$$

---

[6] Вольфганг Паули — знаменитый швейцарский физик-теоретик. Матрицы (8) были введены им для нужд квантовой механики. С их помощью Паули описал взаимодействие внешнего магнитного поля со спином электрона.

Матрицы (8) образуют *полный* набор эрмитовых матриц со свойством (12). Более точно, справедливо следующее

Предложение. *Не существует матрицы* $\sigma \in M^2$ *со свойствами*

$$\sigma = \sigma^*, \tag{14}$$

$$\sigma^2 = I, \tag{15}$$

$$\{\sigma, \sigma_s\} = \mathbb{O}, \quad s = 1, 2, 3. \tag{16}$$

Доказательство. В силу (14) (см. также (1)),

$$\sigma = \begin{pmatrix} \alpha & \overline{\beta} \\ \beta & \gamma \end{pmatrix}, \quad \beta = \beta_1 + i\beta_2; \quad \alpha, \beta_1, \beta_2, \gamma \in \mathbb{R}.$$

Тогда

$$\sigma^2 = \begin{pmatrix} \alpha^2 + |\beta|^2 & (\alpha+\gamma)\overline{\beta} \\ (\alpha+\gamma)\beta & \gamma^2 + |\beta|^2 \end{pmatrix}.$$

Это согласуется с (15) только в двух следующих случаях: $\sigma = \pm I$ и

$$\sigma = \begin{pmatrix} \alpha & \overline{\beta} \\ \beta & -\alpha \end{pmatrix}, \quad \alpha^2 + |\beta|^2 = 1. \tag{17}$$

Ясно, что при $\sigma = \pm I$ будет $\{\sigma, \sigma_1\} = \pm 2\sigma_1$, что противоречит (16). Пусть теперь σ имеет вид (17). Тогда, очевидно, $\sigma = \beta_1\sigma_1 + \beta_2\sigma_2 + \alpha\sigma_3$ и, в силу (12), $\{\sigma, \sigma_1\} = 2\beta_1 I$. Вместе с (16) (при $s = 1$) это означает, что $\beta_1 = 0$. Аналогично получаем $\beta_2 = \alpha = 0$. Но тогда $\sigma = \mathbb{O}$, что противоречит (15). •

§ 3. Одностолбцовые матрицы. Координатные пространства. Линейные отображения

1. *Одностолбцовую матрицу* $\mathbf{x} \in M^{n,1}$ можно записать в виде

$$\mathbf{x} = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}.$$

Поскольку второй индекс в этом случае не несет информации, его принято опускать, и матрицу $\mathbf{x}$ записывают так:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{1}$$

Сложение и умножение на число одностолбцовых матриц производится покомпонентно:

$$\alpha\mathbf{x} + \beta\mathbf{y} = \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \vdots \\ \alpha x_n + \beta y_n \end{pmatrix}, \tag{2}$$

что соответствует общим определениям (1.3) и (1.4). При $n = 3$ эти операции формально совпадают с правилами сложения и умножения на число векторов (в трехмерном пространстве), заданных координатами в каком-либо базисе. В силу этого одностолбцовые матрицы тоже называют *векторами* (n-мерными), а множество всех векторов вида (1) — *координатным пространством размерности* n. Вместо обозначения $M^{n,1}$ будем пользоваться для координатного пространства стандартным обозначением $\mathbb{R}^n$ в вещественном случае и $\mathbb{C}^n$ — в комплексном. Формула (2) задает в $\mathbb{R}^n$ (или в $\mathbb{C}^n$) действия сложения векторов и умножения их на число. При этом для $\mathbf{x} \in \mathbb{R}^n$ допускается умножение лишь на вещественные числа, а для $\mathbf{x} \in \mathbb{C}^n$ — на комплексные. Пространство $\mathbb{R}^1$ естественно отождествляется с множеством $\mathbb{R}$ всех вещественных чисел, пространство $\mathbb{C}^1$ — с множеством $\mathbb{C}$ всех комплексных чисел.

Вектора

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, \; \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{3}$$

назовем *стандартными ортами* в $\mathbb{R}^n$. Набор (3) стандартных ортов образует (стандартный) *базис* в $\mathbb{R}^n$ в том смысле, что для любого вектора (1) справедливо представление

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n. \tag{4}$$

Ясно, что коэффициенты $x_1, \ldots, x_n$ в сумме (4) определяются единственным образом и совпадают с координатами столбца (1).

2. Линейные отображения. Теперь мы свяжем понятие матрицы с понятием линейного отображения[7] . Если каждому элементу $\mathbf{x} \in \mathbb{R}^n$ поставлен в соответствие некоторый элемент $\mathbf{y} \in \mathbb{R}^m$, то говорят, что задано *отображение* (пребразование) $\mathbb{R}^n$ в $\mathbb{R}^m$. Если $\mathcal{A}$ — такое отображение, то пишут $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$ и $\mathbf{y} = \mathcal{A}\mathbf{x}$. Отображение $\mathcal{A}$ называют *линейным*, если для любых $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n$ и любых вещественных чисел α, β справедливо равенство

$$\mathcal{A}(\alpha\mathbf{x} + \beta\mathbf{z}) = \alpha\mathcal{A}\mathbf{x} + \beta\mathcal{A}\mathbf{z}. \tag{5}$$

Пусть задана вещественная матрица $A \in M^{m,n}$ и $\mathbf{x} \in \mathbb{R}^n$. Тогда имеет смысл произведение $A\mathbf{x} \in M^{m,1} = \mathbb{R}^m$. Другими словами, матрица A класса $M^{m,n}$ задает отображение координатного пространства $\mathbb{R}^n$ в пространство $\mathbb{R}^m$. При этом мы считаем матрицу A фиксированной, а столбец $\mathbf{x} \in \mathbb{R}^n$ — аргументом, пробегающим все пространство $\mathbb{R}^n$. Ясно, что действие матрицы A на векторы удовлетворяет условию (5). Это прямо следует из формулы (1.12). Таким образом, всякая вещественная матрица $A \in M^{m,n}$ задает *линейное* отображение $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$. В свою очередь, верно следующее утверждение.

Предложение 1. *Всякое линейное отображение $\mathbb{R}^n$ в $\mathbb{R}^m$ задается некоторой матрицей класса $M^{m,n}$, и при этом только одной.*

---

[7] Слова "функция", "отображение", "преобразование" и т.п. обозначают одно и то же понятие. При использовании их в различных разделах математики учитывают, что слово "функция" имеет скорее аналитический, а "преобразование" — геометрический оттенок.

Доказательство. Действительно, пусть $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$ — линейное отображение. Рассмотрим в $\mathbb{R}^m$ векторы $\mathbf{a}_1 = \mathcal{A}\mathbf{e}_1, \ldots, \mathbf{a}_n = \mathcal{A}\mathbf{e}_n$, где орты $\mathbf{e}_1, \ldots, \mathbf{e}_n$ определены в (3), и запишем их в виде столбцов:

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ \ldots, \ \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}. \qquad (6)$$

Введем в рассмотрение матрицу $A = \{a_{ij}\} \in M^{m,n}$. Тогда, в соответствии с (4),

$$\mathcal{A}x = \sum_{k=1}^{n} x_k \mathcal{A}\mathbf{e}_k = \sum_{k=1}^{n} x_k \mathbf{a}_k,$$

или, в координатной записи,

$$\mathcal{A}\mathbf{x} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ a_{21}x_1 + \ldots + a_{2n}x_n \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}.$$

Мы видим, таким образом, что $\mathcal{A}\mathbf{x} = A\mathbf{x}$. Единственность матрицы A следует из того, что векторы $A\mathbf{e}_1, \ldots, A\mathbf{e}_n$ очевидно совпадают со столбцами матрицы A:

$$A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n). \qquad (7)$$

•

Подобным же образом комплексная матрица класса $M^{m,n}$ определяет линейное отображение $\mathbb{C}^n$ в $\mathbb{C}^m$. При этом в (5) α и β — любые комплексные числа.

С точки зрения линейных отображений определения действий над матрицами оказываются совершенно естественными. Ниже (обычно без оговорок) будем рассматривать только линейные отображения. Пусть $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^m$ — два отображения. Их *суммой* называется отображение $\mathcal{C} : \mathbb{R}^n \to \mathbb{R}^m$, действующее по формуле

$$\mathbf{y} = \mathcal{C}\mathbf{x} = \mathcal{A}\mathbf{x} + \mathcal{B}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n. \tag{8}$$

Если теперь отображение $\mathcal{A}$ задано матрицей $A \in M^{m,n}$, а отображение $\mathcal{B}$ задано матрицей $B \in M^{m,n}$, то, в соответствии с формулой (1.10), можно записать (8) в виде

$$\mathbf{y} = \mathcal{C}\mathbf{x} = (A + B)\mathbf{x}. \tag{9}$$

Сопоставление (8) и (9) показывает, что сумма матриц $C = A + B$ определяет сумму $\mathcal{C}$ соответствующих отображений.

Пусть $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$ - отображение, $\alpha \in \mathbb{R}$. *Произведением отображения $\mathcal{A}$ на число α* называется отображение $\mathcal{D} : \mathbb{R}^n \to \mathbb{R}^m$, действующее по формуле

$$\mathbf{y} = \mathcal{D}\mathbf{x} = \alpha(\mathcal{A}\mathbf{x}), \ \mathbf{x} \in \mathbb{R}^n. \tag{10}$$

Если отображение $\mathcal{A}$ задано матрицей $A \in M^{m,n}$, то, в соответствии с формулой (1.11), можно записать (10) в виде

$$\mathbf{y} = \mathcal{D}\mathbf{x} = (\alpha A)\mathbf{x}. \tag{11}$$

Сопоставление (10) и (11) показывает, что матрица $D = \alpha A$ определяет произведение отображения $\mathcal{A}$ на число α.

Пусть теперь $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$, $\mathcal{B} : \mathbb{R}^m \to \mathbb{R}^p$ — два отображения. Их *композицией* называется отображение $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^p$, заданное формулой

$$\mathbf{y} = \mathcal{F}\mathbf{x} = \mathcal{B}(\mathcal{A}\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n. \tag{12}$$

Если отображение $\mathcal{A}$ задано матрицей $A \in M^{m,n}$, а отображение $\mathcal{B}$ — матрицей $B \in M^{p,m}$, то, в соответствии с формулой (1.13), можно записать (12) в виде

$$\mathbf{y} = \mathcal{F}\mathbf{x} = (BA)\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n. \tag{13}$$

Таким образом, *произведение матриц $F = BA \in M^{p,n}$ определяет композицию соответствующих отображений.*

Мы видим, что действия сложения, умножения на число и умножения матриц введены так, чтобы они порождали сумму соответствующих линейных отображений, произведение отображения на число и композицию отображений. Этим объясняется целесообразность формально введенных в § 1 действий над матрицами.

Упражнения.

1. Если $\mathbf{x} \in \mathbb{R}^n$, то транспонированная матрица $\mathbf{x}^t$ есть *матрица-строка* $\mathbf{x}^t = (x_1, \cdots, x_n)$. Пусть $B \in M^{n,p}$. Проверьте, что $\mathbf{x}^t B$ есть также матрица-строка; какова длина этой строки?

2. Если $\mathbf{x}$ и $\mathbf{y}$ — векторы из $\mathbb{R}^n$, то $\mathbf{x}^t\mathbf{y} \in \mathbb{R}^1 = \mathbb{R}$, $\mathbf{y}\mathbf{x}^t \in M^n$. Выпишите явные формулы для числа $\mathbf{x}^t\mathbf{y}$ и для элементов матрицы $\mathbf{y}\mathbf{x}^t$.

3. Если $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, то $\mathbf{w} = (\mathbf{y}\mathbf{x}^t)\mathbf{z}$ — также вектор из $\mathbb{R}^n$. Покажите, что $\mathbf{w} = \alpha\mathbf{y}$ и найдите число α.

3. Матричная запись системы линейных алгебраических уравнений. Рассмотрим систему m уравнений с n неизвестными

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = f_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = f_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = f_m \end{array}\right. . \tag{14}$$

Будем рассматривать $x_1, \ldots, x_n$ как координаты неизвестного вектора $\mathbf{x} \in \mathbb{R}^n$, $f_1, \ldots, f_m$ — как координаты известного вектора $\mathbf{f} \in \mathbb{R}^m$. Пусть $A = \{a_{ik}\}_{i=1,\ldots,m}^{k=1,\ldots,n}$ — матрица коэффициентов; ясно, что $A \in M^{m,n}$. Тогда систему (14) можно (и удобно) записать в виде одного матричного уравнения

$$A\mathbf{x} = \mathbf{f}. \tag{15}$$

Матрица A задает линейное отображение $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$. Решить систему — значит для заданного вектора $\mathbf{f} \in \mathbb{R}^m$ найти вектор $\mathbf{x} \in \mathbb{R}^n$, удовлетворяющий (15). Иными словами, необходимо найти *прообраз* вектора $\mathbf{f}$ при отображении $\mathcal{A}$. Система (15) разрешима тогда и только тогда, когда такой прообраз существует (хотя бы один). В полезности и содержательности такой точки зрения на систему (14) мы неоднократно убедимся в дальнейшем.

§ 4. Перестановки и подстановки

Здесь собран вспомогательный материал, который понадобится при изучении определителей.

1. Перестановки. Перестановкой порядка n называется набор чисел $1, 2, \ldots, n$, расположенных (переставленных) в определенном порядке. Будем использовать обозначение

$$\mathcal{I} = (i_1, i_2, \ldots, i_n);$$

здесь числа $i_1, i_2, \ldots, i_n$, принимают одно из значений $1, 2, \ldots, n$, причем среди них нет повторяющихся. Если для какой-либо пары (i_k, i_l) при $k < l$ окажется $i_k > i_l$, то скажем, что этой паре в перестановке $\mathcal{I}$ соответствует *беспорядок* (*инверсия*). Например, в перестановке $\mathcal{I} = (3, 5, 2, 1, 4)$ инверсии соответствуют парам $(3, 2)$, $(3, 1)$, $(5, 2)$, $(5, 1)$, $(5, 4)$, $(2, 1)$. В зависимости от четности числа беспорядков перестановку называют *четной* или *нечетной*. *Тождественная перестановка* $I = (1, 2, \ldots, n)$ — четная. Перестановка в рассмотренном примере также четная (число инверсий равно 6).

Поменяем в $\mathcal{I}$ местами какие-нибудь два элемента; такая операция называется *транспозицией*.

Предложение 1. *Всякая транспозиция меняет четность перестановки.*

Доказательство. Сначала предположим, что поменяли местами два соседних элемента i_k, i_{k+1}. Тогда либо добавится один новый беспорядок (если $i_k < i_{k+1}$), либо один беспорядок исчезнет (если $i_k > i_{k+1}$). В любом случае число беспорядков изменится на единицу, и четность перестановки поменяется. Рассмотрим теперь общий случай. Пусть переставлены местами i_k и i_{k+s+1}, $s > 0$. Эту транспозицию можно осуществить так. Переставим i_k последовательно с i_{k+1}, $i_{k+2}, \ldots, i_{k+s+1}$, на что потребуется $(s + 1)$ "соседних" транспозиций. Затем i_{k+s+1} последовательно переставим с

$i_{k+s}, \ldots, i_{k+2}, i_{k+1}$, что потребует еще s "соседних" транспозиций. Таким образом, четность изменится $2s+1$ раз, т.е. окажется противоположной исходной. •

Легко понять, что любую перестановку $\mathcal{I}$ можно перевести в любую другую перестановку $\mathcal{J}$ некоторым числом транспозиций. Этот результат может быть достигнут различными способами. Однако, если четность $\mathcal{I}$ и $\mathcal{J}$ одинакова, то число транспозиций обязательно четно, ибо каждая транспозиция меняет четность. Напротив, если четность $\mathcal{I}$ и $\mathcal{J}$ различна, то число потребных транспозиций — нечетное. В частности, любая четная перестановка получается из тождественной четным, а нечетная перестановка — нечетным числом транспозиций.

Всего различных перестановок порядка n имеется $n!$. Количество четных и нечетных перестановок одинаково и равно $n!/2$. Это следует из того, что если мы *все* $n!$ перестановок одновременно подвергнем какой-нибудь *одной и той же* транспозиции, то четные перестановки перейдут в нечетные и наоборот.

Перестановкам сопоставляют *знак перестановки* $\varepsilon(\mathcal{I})$: $\varepsilon(\mathcal{I}) = +1$, если $\mathcal{I}$ — четная; $\varepsilon(\mathcal{I}) = -1$, если $\mathcal{I}$ — нечетная.

2. Подстановки. Подстановкой порядка n называется взаимно однозначное отображение множества, состоящего из n различных предметов, на себя. Пронумеруем эти предметы. Пусть первый предмет отображается в предмет с номером i_1, второй — в предмет с номером i_2 и т.д. Числа $\mathcal{I} = (i_1, \ldots, i_n)$ образуют перестановку, которая однозначно определяет нашу подстановку. Символически это можно записать так:

$$\sigma = \begin{pmatrix} 1, & 2, & \ldots, & n \\ i_1, & i_2, & \ldots, & i_n \end{pmatrix} = (I, \mathcal{I}), \tag{1}$$

где σ — рассматриваемая подстановка. Ту же подстановку можно записать и иначе. Пусть в первый предмет переходит предмет с номером j_1, во второй — с номером j_2, и т.д. Тогда σ запишется в виде

$$\sigma = \begin{pmatrix} j_1, & j_2, & \ldots, & j_n \\ 1, & 2, & \ldots, & n \end{pmatrix} = (\mathcal{J}, I). \tag{2}$$

Например,

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

Вообще, ясно, что σ можно записать в виде

$$\sigma = \begin{pmatrix} k_1, & k_2, & \ldots, & k_n \\ l_1, & l_2, & \ldots, & l_n \end{pmatrix} = (\mathcal{K}, \mathcal{L}), \tag{3}$$

если (3) получено из (1) применением *одной и той же* перестановки к элементам обеих строк формулы (1). Таким образом, подстановка σ определяется *парой* перестановок $(\mathcal{K}, \mathcal{L})$, определенных с точностью до одной и той же перестановки над элементами $\mathcal{K}$ и $\mathcal{L}$. Вместо (1)–(3) удобно коротко писать

$$\sigma = (I, \mathcal{I}) = (\mathcal{J}, I) = (\mathcal{K}, \mathcal{L}).$$

Тождественную подстановку будем обозначать через $\mathbb{1}$:

$$\mathbb{1} = (I, I) = (\mathcal{K}, \mathcal{K}),$$

где $\mathcal{K}$ — любая перестановка.

Совокупность всех подстановок обозначим через $\mathcal{P}_n$. Так как всякая подстановка единственным образом записывается в виде (1), то множество $\mathcal{P}_n$ содержит $n!$ элементов.

3. Произведение подстановок. Выполняя последовательно подстановки (отображения) τ и σ, мы снова получим подстановку, отвечающую композиции этих отображений. Эта подстановка обозначается через $\sigma\tau$ и называется *произведением* подстановок τ и σ. Это произведение при $n > 2$ не коммутативно: $\sigma\tau$ и $\tau\sigma$, вообще говоря, различны. При образовании произведения $\sigma\tau$ удобно представить τ и σ в виде $\tau = (I, \mathcal{K})$, $\sigma = (\mathcal{K}, \mathcal{L})$. Ясно, что тогда $\sigma\tau = (I, \mathcal{L})$. Например, пусть

$$\tau = \begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 3, & 5, & 2, & 1, & 4 \end{pmatrix},$$

$$\sigma = \begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 5, & 3, & 4, & 1, & 2 \end{pmatrix} = \begin{pmatrix} 3, & 5, & 2, & 1, & 4 \\ 4, & 2, & 3, & 5, & 1 \end{pmatrix}.$$

Тогда

$$\sigma\tau = \begin{pmatrix} 1, & 2, & 3, & 4, & 5 \\ 4, & 2, & 3, & 5, & 1 \end{pmatrix}.$$

Упражнение. Найдите для этого примера подстановку $\tau\sigma$.

Приведем *свойства произведения подстановок.*

1. Ассоциативность:

$$\mu(\sigma\tau) = (\mu\sigma)\tau.$$

Для **доказательства** представим подстановки в виде: $\tau = (I, \mathcal{K})$, $\sigma = (\mathcal{K}, \mathcal{L})$, $\mu = (\mathcal{L}, \mathcal{J})$. Тогда $\sigma\tau = (I, \mathcal{L})$, $\mu(\sigma\tau) = (I, \mathcal{J})$. С другой стороны, $\mu\sigma = (\mathcal{K}, \mathcal{J})$, $(\mu\sigma)\tau = (I, \mathcal{J})$. •

2. Существование "единицы". Для тождественной подстановки $\mathbb{1}$ и любой подстановки σ, очевидно,

$$\mathbb{1}\sigma = \sigma\mathbb{1} = \sigma.$$

3. Существование обратной подстановки. Для любой подстановки σ существует подстановка ρ, такая что

$$\rho\sigma = \sigma\rho = \mathbb{1}.$$

Подстановка ρ называется *обратной* к подстановке σ и обозначается через σ^{-1}. Легко понять, что для $\sigma = (I, \mathcal{I})$ обратной является подстановка $\sigma^{-1} = (\mathcal{I}, I)$. В частности, $\mathbb{1}^{-1} = \mathbb{1}$.

Свойства 1, 2, 3 означают, что множество подстановок $\mathcal{P}_n$ с операцией произведения образует *группу*, то есть множество с ассоциативной внутренней операцией, в котором существует единица и каждый элемент имеет обратный.

4. Знак подстановки. Знаком подстановки $\sigma = (\mathcal{K}, \mathcal{L})$ называется число $\varepsilon(\sigma) = \varepsilon(\mathcal{K})\varepsilon(\mathcal{L})$. Проверим корректность этого определения. Знак $\varepsilon(\sigma)$ зависит только от σ, но не от того, какой именно парой перестановок задается σ. В самом деле, если элементы $\mathcal{K}$ и $\mathcal{L}$ подвергнуть одной и той же перестановке, то четность и $\mathcal{K}$, и $\mathcal{L}$ либо сохранится (если выполняемая перестановка четная), либо изменится. В обоих случаях произведение $\varepsilon(\mathcal{K})\varepsilon(\mathcal{L})$ останется прежним.

Очевидно, $\varepsilon(\mathbb{1}) = 1$. Покажем теперь, что

$$\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau). \tag{4}$$

Для **доказательства** представим подстановки в виде $\tau = (I, \mathcal{K})$, $\sigma = (\mathcal{K}, \mathcal{L})$. Тогда $\sigma\tau = (I, \mathcal{L})$ и $\varepsilon(\tau) = \varepsilon(I)\varepsilon(\mathcal{K}) = \varepsilon(\mathcal{K})$, $\varepsilon(\sigma) = \varepsilon(\mathcal{K})\varepsilon(\mathcal{L})$, $\varepsilon(\sigma\tau) = \varepsilon(\mathcal{L})$. Поэтому, $\varepsilon(\sigma)\varepsilon(\tau) = (\varepsilon(\mathcal{K}))^2\varepsilon(\mathcal{L}) = \varepsilon(\mathcal{L}) = \varepsilon(\sigma\tau)$. •

Из (4), в частности, следует, что $\varepsilon(\sigma\tau) = \varepsilon(\tau\sigma)$. Кроме того, $\varepsilon(\sigma^{-1})\varepsilon(\sigma) = \varepsilon(\sigma^{-1}\sigma) = \varepsilon(\mathbb{1}) = 1$, а потому

$$\varepsilon(\sigma^{-1}) = \varepsilon(\sigma).$$

Подстановка называется *четной*, если $\varepsilon(\sigma) = 1$, и *нечетной*, если $\varepsilon(\sigma) = -1$. Множество четных подстановок образует группу (подгруппу группы $\mathcal{P}_n$). Множество нечетных подстановок не образует группу, поскольку произведение двух нечетных подстановок является четной подстановкой.

§ 5. Определители

1. Определение. Мы уже встречались с определителями второго и третьего порядка в курсе аналитической геометрии. Аналогичным образом определители вводятся для квадратных матриц любого порядка. Они играют существенную роль при исследовании систем линейных уравнений, при изучении свойств линейных преобразований координатных пространств и в ряде других вопросов алгебры и анализа.

Пусть $A \in M^n$, т.е.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Мономом называется произведение n элементов матрицы A, в которое входит по одному элементу из каждой строки и из каждого столбца. Моном имеет вид

$$a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_n j_n} = a_{1k_1} a_{2k_2} \dots a_{nk_n} = a_{l_1 1} a_{l_2 2} \dots a_{l_n n}. \tag{1}$$

Среди первых индексов $i_1, \dots, i_n$ нет одинаковых, т.е. эти индексы образуют перестановку; то же относится ко вторым индексам $j_1, \dots, j_n$. Далее, мы воспользовались тем, что сомножители в произведении можно переставлять местами, причем можно добиться упорядочения или по первым значкам, или по вторым. Впрочем, можно и не стремиться к такому упорядочению (см. первое выражение в (1)). Ясно, что каждый моном определяется некоторой подстановкой $\sigma \in \mathcal{P}_n$:

$$\sigma = \begin{pmatrix} i_1, & \dots, & i_n \\ j_1, & \dots, & j_n \end{pmatrix} = \begin{pmatrix} 1, & \dots, & n \\ k_1, & \dots, & k_n \end{pmatrix} = \begin{pmatrix} l_1, & \dots, & l_n \\ 1, & \dots, & n \end{pmatrix}.$$

Различных мономов столько, сколько различных подстановок, т.е. $n!$. Условимся обозначать моном коротко через $(a)_\sigma$:

$$(a)_\sigma = a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_n j_n}, \quad \sigma = \begin{pmatrix} i_1, & \dots, & i_n \\ j_1, & \dots, & j_n \end{pmatrix} = (\mathcal{I}, \mathcal{J}). \qquad (2)$$

Каждому моному (2) припишем знак $\varepsilon(\sigma)$ и составим сумму всех полученных $n!$ слагаемых вида $\varepsilon(\sigma)(a)_\sigma$. Получившаяся величина (число) называется *определителем* (или *детерминантом*) матрицы A и обозначается через $\det A$:

$$\det A = \sum_{\sigma \in \mathcal{P}_n} \varepsilon(\sigma)(a)_\sigma. \qquad (3)$$

Например, при $n = 2$ $\det A = a_{11}a_{22} - a_{12}a_{21}$; при $n = 3$

$$\begin{aligned} \det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}, \end{aligned} \qquad (4)$$

что совпадает с уже известными нам определениями. (При выписывании членов мы упорядочили произведения по первому значку). Для обозначения определителя пользуются также более подробным символом

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

2. Свойства определителей. Прямое вычисление и исследование определителей на основании определения (3) не вполне удобно, так как с ростом n число слагаемых быстро растет. Поэтому мы обсудим свойства определителей, позволяющие упростить их исследование и вычисление.

1. *При замене строк и столбцов ролями определитель матрицы не меняется.* Иначе говоря,

$$\det A^t = \det A. \tag{5}$$

Доказательство. Действительно, $\det A$ и $\det A^t$, очевидно, составлены из одних и тех же мономов. При этом, если в $\det A$ моном (2) входит со знаком подстановки

$$\sigma = (\mathcal{I}, \mathcal{J}) = \begin{pmatrix} i_1, & \ldots, & i_n \\ j_1, & \ldots, & j_n \end{pmatrix},$$

то в $\det A^t$ этот же моном входит со знаком подстановки

$$\sigma^{-1} = (\mathcal{J}, \mathcal{I}) = \begin{pmatrix} j_1, & \ldots, & j_n \\ i_1, & \ldots, & i_n \end{pmatrix},$$

что соответствует перемене ролей номеров строк и столбцов. Так как $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$, то суммы (3) для $\det A$ и $\det A^t$ состоят из одних и тех же слагаемых. •

Отметим равенство, которое прямо вытекает из (5):

$$\det A^* = \overline{\det A}.$$

В дальнейшем можно все свойства формулировать и доказывать только для строк (или только для столбцов).

2. *Определитель* $\det A$ *есть линейная однородная функция элементов фиксированной строки (столбца).*

Доказательство. В каждый моном входит ровно один сомножитель из i-ой строки. Соберем все члены суммы (3), содержащие элемент a_{ij}; вынесем a_{ij} за скобки, коэффициент при нем обозначим через A_{ij}. Тогда сумма выделенных членов имеет вид $a_{ij}A_{ij}$. Если мы (при фиксированном i) будем менять

j от 1 до n, то в итоге переберем в точности все слагаемые, входящие в (3). Тот же результат получится, если мы при фиксированном j будем менять i от 1 до n. В результате мы приходим к равенствам

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij}, \quad i = 1, \dots, n, \tag{6_1}$$

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij}, \quad j = 1, \dots, n. \tag{6_2}$$

Формулы (6_1) (соответственно, (6_2)) и означают, что $\det A$ есть линейная однородная функция элементов i-ой строки (соответственно, j-ого столбца). •

Формулы (6) представляют собой $2n$ различных формул для вычисления определителя. Формула (6_1) дает разложение по элементам i-ой строки, формула (6_2) — по элементам j-го столбца. Формулы (6) приобретут практический смысл, если мы найдем удобный способ вычисления коэффициентов A_{ij}. Это будет сделано в п. **3**. Коэффициент A_{ij} называют *алгебраическим дополнением* элемента a_{ij} матрицы A.

Пример. Разложить определитель третьего порядка по элементам второй строки. Группируя члены в (4), получаем:

$$\det A = \\ a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) + a_{23}(a_{12}a_{31} - a_{11}a_{32}). \tag{7}$$

3. *Определитель матрицы, имеющей строку (столбец), состоящую сплошь из нулей, равен нулю.*

Это сразу следует из (6).

4. Следующее свойство (правило сложения), выражаемое формулой

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

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{i1} & \tilde{a}_{i2} & \cdots & \tilde{a}_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \qquad (8)$$

также облегчает вычисление определителей. В формуле (8) все строки, кроме i-ой, одинаковы для всех трех определителей.

Для **доказательства** надо разложить определители по i-ой строке (см. (6_1)). Коэффициенты A_{ij} при элементах i-ой строки не содержат в себе элементов этой строки. Поэтому коэффициенты A_{ij} во всех трех определителях в (8) одни и те же. Имеем:

$$\sum_{j=1}^{n}(a_{ij} + \tilde{a}_{ij})A_{ij} = \sum_{j=1}^{n} a_{ij}A_{ij} + \sum_{j=1}^{n} \tilde{a}_{ij}A_{ij}. \bullet$$

Упражнение. Сформулируйте и проверьте аналогичное правило сложения для выделенного столбца определителя.

5. *Из элементов строки (столбца) определителя можно выносить общий множитель:*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \qquad (9)$$

Эта формула также прямо следует из (6_1). Из (9), в свою очередь, получаем

$$\det(\alpha A) = \begin{vmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{i1} & \alpha a_{i2} & \cdots & \alpha a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{n1} & \alpha a_{n2} & \cdots & \alpha a_{nn} \end{vmatrix} = \alpha^n \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

то есть

$$\det(\alpha A) = \alpha^n \det A. \tag{10}$$

6. *При перестановке местами двух строк (столбцов) определитель меняет знак.*

Доказательство. Пусть для определенности переставлены местами k-ый и l-ый столбцы матрицы A. Полученную матрицу обозначим через $\tilde{A}$. Произведения (мономы) (2), входящие в $\det \tilde{A}$, при этом останутся теми же. Однако знак при каждом мономе будет уже определяться не прежней подстановкой

$$\sigma = \begin{pmatrix} i_1, & \dots, & i_n \\ j_1, & \dots, & j_n \end{pmatrix},$$

а подстановкой $\tilde{\sigma}$, которая получается из σ одной транспозицией в нижней строке: номера j_k и j_l должны поменяться местами. Так как при этом знак $\tilde{\sigma}$, очевидно, противоположен знаку σ, т.е. $\varepsilon(\tilde{\sigma}) = -\varepsilon(\sigma)$, то все члены в сумме вида (3) для $\det \tilde{A}$ отличаются знаком от соответствующих членов для $\det A$. •

7. *Определитель матрицы с двумя одинаковыми строками (столбцами) равен нулю.*

Доказательство. Если мы переставим эти строки местами, то определитель должен изменить знак. Однако матрица (и определитель) не изменится. Поэтому $\det A = -\det A$, а следовательно, $\det A = 0$. •

8. *Если две строки (два столбца) пропорциональны, то определитель равен нулю.*

Это сразу же следует из свойств 5 и 7.

9. *Определитель не изменится, если к какой-нибудь его строке прибавить какую-либо линейную комбинацию остальных строк. (Аналогично и для столбцов.)*

Доказательство. Пусть, например, к i-ой строке прибавлены остальные строки, умноженные соответственно на числа $\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n$. (При этом сами эти строки на своих местах остаются без изменений.) Тогда в i-ой строке на j-ом месте будет стоять элемент

$$a_{ij} + \alpha_1 a_{1j} + \ldots + \alpha_{i-1} a_{i-1,j} + \alpha_{i+1} a_{i+1,j} + \ldots + \alpha_n a_{nj}.$$

Применяя формулу (8) несколько раз, представим этот определитель в виде суммы n определителей, причем первое слагаемое есть $\det A$, а остальные слагаемые представляют собой определители с двумя пропорциональными строками. Такие определители равны нулю. •

3. Миноры и алгебраические дополнения. Пусть в матрице $A \in M^n$ вычеркнуты i-ая строка и j-ый столбец. Определитель оставшейся матрицы порядка $n-1$ обозначим через M_{ij}:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

Определитель M_{ij} называют *минором*, отвечающим элементу a_{ij}. Мы покажем, что коэффициенты A_{ij} (алгебраические дополнения) в разложениях (6) определителя $\det A$ по элементам i-ой строки или j-ого столбца выражаются через миноры по формуле

$$A_{ij} = (-1)^{i+j} M_{ij}. \tag{11}$$

Тем самым, формулы (6) сводят вычисление $\det A$ к вычислению определителей меньшего порядка.

Доказательство формулы (11). Установим сначала, что $A_{nn} = M_{nn}$. Все мономы, содержащие a_{nn}, имеют вид

$$a_{1i_1} a_{2i_2} \ldots a_{n-1,i_{n-1}} a_{nn}, \tag{12}$$

где $(i_1, i_2, \ldots, i_{n-1})$ — произвольная перестановка индексов $1, 2, \ldots, n-1$. Знак при этом мономе определяется знаком $\varepsilon(\sigma)$

подстановки

$$\sigma = \begin{pmatrix} 1, & 2, & \ldots, & n-1, & n \\ i_1, & i_2, & \ldots, & i_{n-1}, & n \end{pmatrix} \in \mathcal{P}_n.$$

Ясно, что $\varepsilon(\sigma) = \varepsilon(\tau)$, где

$$\tau = \begin{pmatrix} 1, & 2, & \ldots, & n-1 \\ i_1, & i_2, & \ldots, & i_{n-1} \end{pmatrix} \in \mathcal{P}_{n-1}.$$

Действительно, число инверсий в нижних строках подстановок σ и τ одинаково. Сумма всех мономов вида (12) есть $a_{nn}A_{nn}$, где алгебраическое дополнение A_{nn} имеет вид

$$A_{nn} = \sum_{\tau \in \mathcal{P}_{n-1}} \varepsilon(\tau) a_{1i_1} a_{2i_2} \ldots a_{n-1,i_{n-1}}.$$

Правая часть по определению есть детерминант матрицы порядка $n-1$, стоящей в верхнем левом углу матрицы A. Итак, $A_{nn} = M_{nn}$.

Пусть теперь выделен элемент a_{ij}. Переставим строки матрицы A *последовательно* так, чтобы i-ая строка стала последней, а остальные шли в "правильном" порядке. На это уйдет $n - i$ перестановок. Затем j-ый столбец также переставим на последнее место (еще $n - j$ перестановок). В новом определителе элемент a_{ij} займет место в правом нижнем углу, а *его минор останется неизменным*, т.к. остальные строки и столбцы стоят в прежнем порядке. Нужно учесть, что новый определитель отличается от старого множителем $(-1)^{2n-i-j} = (-1)^{i+j}$. Одновременно и алгебраическое дополнение выделенного элемента в новом определителе отличается от старого тем же множителем. Тогда (11) следует из уже полученного соотношения $A_{nn} = M_{nn}$. •

Рассмотрим теперь сумму

$$\sum_{j=1}^{n} a_{ij} A_{kj}, \quad k \neq i, \tag{13}$$

т.е. будем умножать элементы строки на алгебраические дополнения "чужой" строки. Такую сумму мы получим, если будем вычислять определитель, у которого на место k-ой строки подставлена i-ая строка (остальные строки — те же, что в исходной матрице). Тем самым, получаем определитель с двумя одинаковыми строками, он равен нулю. Итак, сумма (13) равна нулю. Аналогично,

$$\sum_{i=1}^{n} a_{ij} A_{ik} = 0, \quad k \neq j.$$

Объединим этот результат с формулами (6), используя символ Кронекера. Тогда получим

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \delta_{ik} \det A, \qquad \sum_{i=1}^{n} a_{ij} A_{ik} = \delta_{jk} \det A. \tag{14}$$

4. Примеры вычисления определителей. Вычисление определителя упрощается, когда матрица содержит много нулей. Пусть, например, матрица верхнетреугольная:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Тогда, раскрывая $\det A$ по элементам первого столбца, находим:

$$\det A = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

Продолжая этот процесс, получим, что $\det A = a_{11} a_{22} \ldots a_{nn}$. Такая же формула получится для нижнетреугольной матрицы. Таким образом, *определитель треугольной матрицы равен* *произведению элементов главной диагонали.* В частности, $\det I = 1$.

Вычислим теперь *определитель Вандермонда* порядка $n+1$:

$$\Delta_{n+1} = \Delta_{n+1}(x_1, x_2, \ldots, x_{n+1}) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_{n+1} \\ x_1^2 & x_2^2 & \cdots & x_{n+1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & \cdots & x_{n+1}^n \end{vmatrix}, \qquad (15)$$

где $x_1, x_2, \ldots, x_{n+1}$ — какие-либо комплексные числа. Докажем,что

$$\Delta_{n+1} = \Delta_{n+1}(x_1, x_2, \ldots, x_{n+1}) = \prod_{j>k}(x_j - x_k). \qquad (16)$$

Если среди $x_1, x_2, \ldots, x_{n+1}$ хотя бы два числа совпадают, то определитель (15) содержит два одинаковых столбца и, следовательно, равен нулю. Формула (16) в этом случае очевидна. Поэтому можно считать, что среди чисел $x_1, \ldots, x_n$ (без x_{n+1}) нет одинаковых. Рассмотрим Δ_{n+1} как функцию от переменной $x = x_{n+1}$ при фиксированных $x_1, x_2, \ldots, x_n$. Раскладывая определитель (15) по последнему столбцу, видим, что Δ_{n+1} есть полином степени n от x:

$$\Delta_{n+1}(x_1, \ldots, x_n, x) = a_n x^n + \ldots + a_1 x + a_0 := P_n(x). \qquad (17)$$

Коэффициент a_n при x^n есть алгебраическое дополнение $A_{n+1,n+1}$ элемента x^n, стоящего в правом нижнем углу определителя. Очевидно,

$$a_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \Delta_n \qquad (18)$$

является определителем Вандермонда порядка n. Как уже отмечалось, при $x = x_j$, $j = 1, \ldots, n$, определитель равен нулю, то есть числа $x_1, \ldots, x_n$ суть корни многочлена P_n. Выпишем разложение P_n на множители,

$$P_n(x) = a_n(x - x_1)(x - x_2)\ldots(x - x_n). \tag{19}$$

Учитывая (17), (18), получаем рекуррентную формулу

$$\Delta_{n+1} = (x_{n+1} - x_1)(x_{n+1} - x_2)\ldots(x_{n+1} - x_n)\Delta_n. \tag{20}$$

Применяя формулу (20) последовательно n раз и учитывая, что $\Delta_1 = 1$, получаем (16).

§ 6. Теорема об определителе произведения матриц

Теорема 1. *Определитель произведения матриц равен произведению их определителей:*

$$\det AB = \det A \det B, \quad A \in M^n, \quad B \in M^n. \tag{1}$$

Мы дадим два различных[8] доказательства этой фундаментальной теоремы.

Первый способ. *Прямое доказательство.* Пусть $C = AB$. Тогда $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$. Для $\det C$ запишем выражение в виде суммы $n!$ мономов, упорядоченных по первому значку:

$$\det C = \sum_{\mathcal{J}} \varepsilon(\mathcal{J}) c_{1j_1} \ldots c_{nj_n}.$$

Здесь суммирование ведется по всем перестановкам $\mathcal{J} = (j_1, \ldots, j_n)$ чисел $1, 2, \ldots, n$. Подставляя вместо элементов c_{ij}

[8] При изучении курса достаточно разобрать какое-либо одно доказательство. В руководствах по линейной алгебре можно найти и другие доказательства соотношения (1).

их выражения, получим

$$\det C =$$
$$= \sum_{\mathcal{J}} \varepsilon(\mathcal{J}) \left(\sum_{k_1=1}^{n} a_{1k_1} b_{k_1 j_1} \right) \left(\sum_{k_2=1}^{n} a_{2k_2} b_{k_2 j_2} \right) \dots \left(\sum_{k_n=1}^{n} a_{nk_n} b_{k_n j_n} \right)$$
$$= \sum_{\mathcal{J}=(j_1,\dots,j_n)} \varepsilon(\mathcal{J}) \sum_{k_1,\dots,k_n=1}^{n} a_{1k_1} a_{2k_2} \dots a_{nk_n} b_{k_1 j_1} b_{k_2 j_2} \dots b_{k_n j_n}.$$

Во внутренней сумме индексы $k_1, \dots, k_n$ могут принимать и совпадающие значения, в то время как во внешней сумме значения всех индексов $j_1, j_2, \dots, j_n$ различны. Меняя порядок суммирования, найдем:

$$\begin{aligned} \det C &= \sum_{k_1,\dots,k_n=1}^{n} a_{1k_1} a_{2k_2} \dots a_{nk_n} \sum_{\mathcal{J}=(j_1,\dots,j_n)} \varepsilon(\mathcal{J}) b_{k_1 j_1} b_{k_2 j_2} \dots b_{k_n j_n} \\ &= \sum_{k_1,\dots,k_n=1}^{n} a_{1k_1} a_{2k_2} \dots a_{nk_n} S_{k_1,k_2,\dots,k_n}, \end{aligned} \tag{2}$$

где

$$S_{k_1,k_2,\dots,k_n} = \sum_{\mathcal{J}=(j_1,\dots,j_n)} \varepsilon(\mathcal{J}) b_{k_1 j_1} b_{k_2 j_2} \dots b_{k_n j_n}. \tag{3}$$

Поменяем в (3) какие-либо два из индексов $k_1, k_2, \dots, k_n$ местами. В выражении (3) это соответствует перестановке двух соответствующих сомножителей вида $b_{k_s j_s}$ местами. Иначе говоря, над перестановкой $\mathcal{J} = (j_1, \dots, j_n)$ совершается одна транспозиция. В результате каждое слагаемое в (3) изменит знак. Таким образом, получаем, что *перестановка местами каких-либо двух из индексов $k_1, \dots, k_n$ приводит к изменению знака у суммы* (3). Отсюда следует, что $S_{k_1,k_2,\dots,k_n} = 0$, *если среди индексов $k_1, \dots, k_n$ найдутся одинаковые.* Все это позволяет в выражении (2) заменить суммирование по независимо меняющимся индексам $k_1, \dots, k_n$ на *суммирование* *по перестановкам* $\mathcal{K} = (k_1, \ldots, k_n)$:

$$\det C =$$

$$= \sum_{\mathcal{K}=(k_1,\ldots,k_n)} a_{1k_1} a_{2k_2} \ldots a_{nk_n} \sum_{\mathcal{J}=(j_1,\ldots,j_n)} \varepsilon(\mathcal{J}) b_{k_1 j_1} b_{k_2 j_2} \ldots b_{k_n j_n} \tag{4}$$

$$= \sum_{\mathcal{K}=(k_1,\ldots,k_n)} \varepsilon(\mathcal{K}) a_{1k_1} a_{2k_2} \ldots a_{nk_n} \sum_{\mathcal{J}=(j_1,\ldots,j_n)} \varepsilon(\mathcal{K}, \mathcal{J}) b_{k_1 j_1} b_{k_2 j_2} \ldots b_{k_n j_n}.$$

В последней выкладке использовано, что $\varepsilon(\mathcal{K})\varepsilon(\mathcal{K}, \mathcal{J}) = \varepsilon^2(\mathcal{K})\varepsilon(\mathcal{J}) = \varepsilon(\mathcal{J})$. Внутренняя сумма в правой части (4) есть $\det B$. Таким образом,

$$\det C = \left(\sum_{\mathcal{K}=(k_1,\ldots,k_n)} \varepsilon(\mathcal{K}) a_{1k_1} a_{2k_2} \ldots a_{nk_n} \right) \det B = \det A \det B,$$

что совпадает с (1). •

2. Линейные и полилинейные формы. В этом пункте мы подготовим материал для второго способа доказательства теоремы 1. Этот материал представляет и самостоятельный интерес.

Линейной формой от n комплексных переменных $x_1, \ldots, x_n$ называют линейную однородную функцию этих переменных. Произвольная линейная форма Φ переменных $x_1, \ldots, x_n$ имеет вид

$$\Phi(\mathbf{x}) = \varphi_1 x_1 + \varphi_2 x_2 + \ldots + \varphi_n x_n. \tag{5}$$

Здесь набор $x_1, \ldots, x_n$ отождествлен с вектором $\mathbf{x} \in \mathbb{C}^n$ вида (3.1), а числа $\varphi_1, \ldots, \varphi_n$ (коэффициенты формы) фиксированы. Подставляя в (5) $\mathbf{x} = \mathbf{e}_k$, где $\mathbf{e}_k$ — орт (3.3), получим:

$$\Phi(\mathbf{e}_k) = \varphi_k, \quad k = 1, \ldots, n. \tag{6}$$

Таким образом, форма (5) полностью определяется своими значениями на ортах. Если заданы две формы Φ и Ψ, то для их равенства на всех $\mathbf{x} \in \mathbb{C}^n$, необходимо и достаточно выполнения n равенств

$$\Phi(\mathbf{e}_k) = \Psi(\mathbf{e}_k), \quad k = 1, \ldots, n. \tag{7}$$

Пусть теперь $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^n$ и форма $\Phi = \Phi(\mathbf{x}_1, \mathbf{x}_2)$ линейна по $\mathbf{x}_1$ и $\mathbf{x}_2$ порознь. Форма Φ тогда называется 2-формой. Используя разложение (3.4) для $\mathbf{x}_1$ и $\mathbf{x}_2$, находим, что форма $\Phi(\mathbf{x}_1, \mathbf{x}_2)$ полностью определяется значениями n^2 чисел $\Phi(\mathbf{e}_{k_1}, \mathbf{e}_{k_2})$, $k_1, k_2 = 1, \ldots, n$. Аналогично определяются формы более высоких порядков (*полилинейные* формы). Особое значение для нас имеют n-формы $\Phi(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Очевидно, две n-формы Φ, Ψ совпадают тогда и только тогда, когда выполнено n^n равенств

$$\begin{aligned} \Phi(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n}) = \Psi(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \ldots, \mathbf{e}_{k_n}), \\ k_1, k_2, \ldots, k_n = 1, \ldots, n. \end{aligned} \tag{8}$$

Совокупность равенств (8) сводится к всего лишь одному равенству, если n-формы Φ, Ψ *полностью антисимметричны*, т.е. меняют знак при перестановке любой пары аргументов $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Справедлива

Теорема 2. *Две полностью антисимметричные n-формы Φ и Ψ совпадают тогда и только тогда, когда*

$$\Phi(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \Psi(\mathbf{e}_1, \ldots, \mathbf{e}_n). \tag{9}$$

Доказательство. Если среди индексов $k_1, \ldots, k_n$ в (8) есть совпадающие, то, очевидно, обе части равны нулю. Таким образом, достаточно считать, что индексы $k_1, \ldots, k_n$ в (8) образуют *перестановку* $\mathcal{K} = (k_1, k_2, \ldots, k_n)$. Переводя ее транспозициями в тождественную перестановку $(1, 2, \ldots, n)$, мы приведем (8) к (9). •

Исходя из теоремы 2, можно дать явное описание всех полностью антисимметричных n-форм. При этом нам сейчас удобно записывать аргументы в виде столбцов $\mathbf{a}_1, \ldots, \mathbf{a}_n$ (см. (3.6)), объединяя их в матрицу A (см. (3.7)). Таким образом, станем писать $\Phi(A) = \Phi(\mathbf{a}_1, \ldots, \mathbf{a}_n)$.

Теорема 3. *Пусть $\Phi(A)$ — полностью антисимметричная n-форма относительно столбцов матрицы A. Тогда*

$$\Phi(A) = (\det A)\Phi(I). \tag{10}$$

Доказательство. Величина $\det A$ линейно зависит от каждого из столбцов $\mathbf{a}_1, \ldots, \mathbf{a}_n$ (см. формулы (5.6_2)), т.е. является n-формой. Она полностью антисимметрична (см. свойство 6 в § 5). В силу теоремы 2 достаточно проверить (10) при $A = I$. Но тогда (10), очевидно, выполнено в силу равенства $\det I = 1$. •

3. Второе доказательство теоремы 1 (*метод n-форм*). Чтобы не менять обозначений, установим (1) в эквивалентной форме:

$$\det A \det B = \det BA. \tag{11}$$

Будем считать матрицу B фиксированной. Очевидно, левая часть в (11) — полностью антисимметричная n-форма относительно столбцов матрицы A. Проверим, что то же верно и для правой части. Обозначим через $\mathbf{b}_1^t, \ldots, \mathbf{b}_n^t$ последовательные строки матрицы B. Ясно, что тогда

$$\det BA = \begin{vmatrix} \mathbf{b}_1^t\mathbf{a}_1 & \mathbf{b}_1^t\mathbf{a}_2 & \ldots & \mathbf{b}_1^t\mathbf{a}_n \\ \mathbf{b}_2^t\mathbf{a}_1 & \mathbf{b}_2^t\mathbf{a}_2 & \ldots & \mathbf{b}_2^t\mathbf{a}_n \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{b}_n^t\mathbf{a}_1 & \mathbf{b}_n^t\mathbf{a}_2 & \ldots & \mathbf{b}_n^t\mathbf{a}_n \end{vmatrix}. \tag{12}$$

Из представления (12) непосредственно ясно, что $\det BA$ есть n-форма относительно векторов $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Ясно также, что перемена местами любых двух векторов из набора $\mathbf{a}_1, \ldots, \mathbf{a}_n$ приводит к транспозиции двух соответствующих столбцов в определителе (12). Таким образом, n-форма $\det BA$ полностью антисимметрична. В силу теоремы 2, достаточно проверить (11) при $A = I$. Но при $A = I$ соотношение (11) очевидно. •

С о д е р ж а н и е