Казанский государственный университет
им. В.И. Ульянова-Ленина

С.Н. Ильин

ЭЛЕМЕНТЫ АЛГЕБРЫ: КОМПЛЕКСНЫЕ ЧИСЛА, СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ, МНОГОЧЛЕНЫ

УЧЕБНОЕ ПОСОБИЕ

Казань
2006

УДК 512

Печатается по решению
Учебно-методической комиссии
механико-математического факультета КГУ

Научный редактор
кандидат физико-математических наук, доцент Корешков Н.А.

Ильин С.Н.

Элементы алгебры: комплексные числа, системы линейных уравнений, многочлены. Учебное пособие. — Казань: Казанский государственный университет им. В.И. Ульянова-Ленина, 2006. — 68 с.

Учебное пособие предназначено для студентов I курса механико-математического факультета КГУ.

§0. Начальные определения и понятия

Как известно, любая наука опирается на ряд основных определений и понятий. Для алгебры таковыми являются понятия множества, отображения и алгебраической операции.

Множеством называется произвольная совокупность объектов, называемых *элементами* множества. Обычно множества обозначают заглавными латинскими буквами $A, B, C, \ldots$, а их элементы — прописными буквами $a, b, c, \ldots$ Если элемент a лежит в множестве A, то пишут $a \in A$, в противном случае — $a \notin A$. Говорят, что A есть *подмножество* множества B (обозначение: $A \subseteq B$), если любой элемент множества A лежит в B. Множества A и B *равны*, если $A \subseteq B$ и $B \subseteq A$. Если $A \subseteq B$, но $B \nsubseteq A$, то подмножество A множества B называется *собственным* (в этом случае пишут: $A \subset B$). Множество, не содержащее ни одного элемента, называется *пустым* и обозначается символом $\emptyset$. Ясно, что $\emptyset \subseteq A$ для любого множества A.

Важными примерами множеств являются известные из курса школьной математики множества $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, соответственно, натуральных, целых, рациональных и вещественных чисел. Очевидно, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Объединением множеств A и B называется множество $A \cup B$, состоящее из всех элементов, каждый из которых лежит хотя бы в одном из множеств A и B, то есть, $A \cup B = \{c : c \in A \text{ или } c \in B\}$. Двойственным образом определяется *пересечение* множеств: $A \cap B = \{c : c \in A \text{ и } c \in B\}$. *Разностью* множеств A и B называется множество $A \backslash B = \{c : c \in A, c \notin B\}$.

Упражнение 0.1 Докажите справедливость соотношений:

1) $A \cap B \subseteq A \subseteq A \cup B$;

2) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;

3) $A \cap (B \backslash A) = \emptyset$;

4) $(A \backslash B) \backslash C = A \backslash (B \cup C)$.

Декартовым произведением множеств A и B называется множество $A \times B$, состоящее из упорядоченных пар, первый элемент которых лежит в A, а второй — в B, то есть, $A \times B = \{(a, b) : a \in A, b \in B\}$. Данное определение легко распространяется на любое конечное число сомножителей $A_1, \ldots, A_k$, что позволяет

определить декартово произведение $A_1 \times \cdots \times A_k$. В случае, когда $A_1 = \cdots = A_k = A$, соответствующее произведение называется *декартовой степенью* множества A и обозначается A^k.

Пусть A и B — множества. *Отображением* из A в B называется соответствие φ, которое каждому элементу $a \in A$ сопоставляет некоторый элемент $\varphi(a) \in B$. Отображение $*$ множества $A \times A$ в A называется *бинарной алгебраической операцией* на множестве A. Образ пары $(a, b) \in A \times A$ при этом записывают обычно в виде $a * b$. Естественными примерами алгебраических операций на числовых множествах являются обычные операции сложения "$+$" и умножения "$\cdot$" чисел. Легко видеть, что вычитание является алгебраической операцией на множествах $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, но не является таковой на множестве $\mathbb{N}$. На каких числовых множествах будет алгебраической операцией деление?

Основными объектами изучения в алгебре являются *алгебраические системы* — множества с заданными на них алгебраическими операциями, при этом главное значение имеет не природа самих множеств, а свойства алгебраических операций. Наиболее важными примерами алгебраических систем являются группы, кольца и поля.

Система $(G, *)$ называется *группой*, если выполняются следующие условия:

1) для всех $a, b, c \in G$ верно $(a * b) * c = a * (b * c)$ (ассоциативность);

2) существует элемент $e \in G$ такой, что $e * a = a * e = a$ для всех $a \in G$ (существование нейтрального элемента);

3) для любого $a \in G$ существует $a' \in G$ такой, что $a * a' = a' * a = e$ (существование обратного элемента).

Если дополнительно выполняется условие

4) для всех $a, b \in G$ верно $a * b = b * a$ (коммутативность),

то группа G называется *абелевой*.

В зависимости от выбора знака алгебраической операции различают мультипликативную и аддитивную терминологии. Различия между ними приведены в таблице 1. Аддитивную терминологию применяют, как правило, для абелевых групп.

Легко убедиться в том, что множества $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ образуют абелевы группы относительно сложения, а множества $\mathbb{Q}^* = \mathbb{Q}\backslash\{0\}$ и $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$ — относительно умножения.

терминология	знак	e	a'
мультипликативная	$\cdot$	единичный элемент, единица, e, 1	обратный элемент, a^{-1}
аддитивная	$+$	нейтральный элемент, нуль, 0	противоположный элемент, $-a$

Таб. 1

Система $(R, +, \cdot)$ называется *кольцом*, если

1) $(R, +)$ — абелева группа;

2) для всех $a, b, c \in R$ верно $(a + b)c = ac + bc$ и $a(b + c) = ab + ac$ (дистрибутивность).

Если умножение в R обладает дополнительными свойствами, например, ассоциативностью, коммутативностью или в кольце существует единица, то говорят, что кольцо R ассоциативно, коммутативно или, соответственно, обладает единицей. Если $(R\backslash\{0\}, \cdot)$ — абелева группа, то кольцо R называется *полем*.

Непосредственно проверяется, что относительно обычных операций сложения и умножения множество $\mathbb{Z}$ образует коммутативное ассоциативное кольцо с единицей, множество четных чисел — коммутативное ассоциативное кольцо без единицы, а $\mathbb{Q}$ и $\mathbb{R}$ являются полями.

§1. Поле комплексных чисел

В уже упоминавшейся ранее цепочке $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ каждое последующее множество обладает более “хорошими” алгебраическими свойствами по сравнению с предыдущим: натуральные числа можно только складывать и умножать, целые — еще и вычитать, рациональные — делить (если делитель отличен от 0), из неотрицательных вещественных чисел можно извлекать арифметические корни. Нельзя ли еще расширить поле вещественных чисел так, чтобы получившееся множество по прежнему являлось полем, но при этом корни извлекались бы из всех чисел? Оказывается, можно, и соответствующее множество, которое сейчас будет построено, называется полем комплексных чисел.

1.1 Построение поля комплексных чисел. Алгебраическая форма комплексного числа

Пусть $\mathbb{C} = \mathbb{R} \times \mathbb{R}$. Зададим на $\mathbb{C}$ операции сложения и умножения, положив

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),$$
$$(a_1, b_1)(a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + b_1a_2)$$

для всех $(a_1, b_1), (a_2, b_2) \in \mathbb{C}$. (Для обозначения операций сложения и умножения пар используются те же знаки "$+$" и "$\cdot$", что и для сложения и умножения вещественных чисел. Это не совсем корректно, но удобно.)

Теорема 1.1 $(\mathbb{C}, +, \cdot)$ — *поле.*

Доказательство. Ассоциативность и коммутативность сложения в $\mathbb{C}$ немедленно вытекают из аналогичных свойств сложения вещественных чисел. (Например, $(a_1, b_1)+(a_2, b_2) = (a_1+a_2, b_1+b_2) = (a_2+a_1, b_2+b_1) = (a_2, b_2)+(a_1, b_1)$.) Нулем будет элемент (0,0), а противоположным к (a, b) элементом — элемент $(-a, -b)$. Вполне очевидна и коммутативность умножения. Для проверки ассоциативности умножения вычислим выражение $\big((a_1, b_1)(a_2, b_2)\big)(a_3, b_3)$:

$$\big((a_1, b_1)(a_2, b_2)\big)(a_3, b_3) = (a_1a_2 - b_1b_2, a_1b_2 + b_1a_2)(a_3, b_3) =$$
$$(a_1a_2a_3 - b_1b_2a_3 - a_1b_2b_3 - b_1a_2b_3, a_1a_2b_3 - b_1b_2b_3 + a_1b_2a_3 + b_1a_2a_3).$$

Непосредственно проверяется, что вычисление выражения $(a_1, b_1)\big((a_2, b_2)(a_3, b_3)\big)$ дает тот же результат. Проверим дистрибутивность:

$$\begin{aligned}\big((a_1, b_1) + (a_2, b_2)\big)(a_3, b_3) = \\ (a_1a_3 + a_2a_3 - b_1b_3 - b_2b_3, a_1b_3 + a_2b_3 + b_1a_3 + b_2a_3) = \\ (a_1a_3 - b_1b_3, a_1b_3 + b_1a_3) + (a_2a_3 - b_2b_3, a_2b_3 + b_2a_3) = \\ (a_1, b_1)(a_3, b_3) + (a_2, b_2)(a_3, b_3),\end{aligned}$$

следовательно,

$$\big((a_1, b_1) + (a_2, b_2)\big)(a_3, b_3) = (a_1, b_1)(a_3, b_3) + (a_2, b_2)(a_3, b_3).$$

С учетом коммутативности умножения получаем второй закон дистрибутивности

$$(a_1, b_1)\big((a_2, b_2) + (a_3, b_3)\big) = (a_1, b_1)(a_2, b_2) + (a_1, b_1)(a_3, b_3).$$

Легко видеть, что роль единицы в $\mathbb{C}$ играет элемент (1,0). Наконец, покажем, что любой отличный от (0,0) элемент (a, b) обладает обратным элементом. В самом деле,

$$(a,b)\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = \left(\frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2}, -\frac{ab}{a^2+b^2} + \frac{ab}{a^2+b^2}\right) = (1,0),$$

откуда ввиду коммутативности умножения

$$\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)(a,b) = (1,0).$$

Теорема доказана. $\lhd$

Итак, поле $\mathbb{C}$ комплексных чисел построено. Заметим, что числа вида $(a, 0)$, где $a \in \mathbb{R}$, складываются и перемножаются так же, как и вещественные числа: $(a,0) + (b,0) = (a+b,0)$, $(a,0)(b,0) = (ab,0)$. Поэтому для упрощения записи можно отождествить такие комплексные числа с вещественными и вместо $(a, 0)$ писать просто a. В этом смысле можно считать, что $\mathbb{R} \subseteq \mathbb{C}$. Отметим также, что $a(b,c) = (a,0)(b,c) = (ab, ac)$. Следовательно,

$$(a,b) = (a,0) + (0,b) = a + b(0,1).$$

Обозначив $(0,1)$ через i, получаем представление комплексного числа $z = (a,b)$ в *алгебраической форме*:

$$z = a + bi.$$

Число i называют *мнимой единицей*, его замечательное свойство состоит в том, что $i^2 = -1$. В самом деле,

$$i^2 = (0,1)(0,1) = (-1,0) = -1.$$

Последнее равенство, в частности, доказывает, что полученное ранее включение $\mathbb{R} \subseteq \mathbb{C}$ является строгим: $\mathbb{R} \subset \mathbb{C}$.

Вещественные числа a и b в представлении комплексного числа $z = a + bi$ называются его *вещественной* и *мнимой* частями и обозначаются $\operatorname{Re} z$ и $\operatorname{Im} z$, соответственно. Таким образом, число z является вещественным тогда и только тогда, когда его мнимая часть равна 0.

Если же у ненулевого числа z равна нулю вещественная часть, то такое число называют *чисто мнимым*.

Легко видеть, что введенные выше операции сложения и умножения комплексных чисел при переходе к алгебраической форме принимают вид

$$(a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i,$$
$$(a_1 + b_1 i)(a_2 + b_2 i) = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2)i,$$

что соответствует обычным правилам преобразований буквенных выражений с учетом равенства $i^2 = -1$.

Каждому комплексному числу $z = a + bi$ можно сопоставить *комплексно-сопряженное* число $\bar{z} = a - bi$. В качестве упражнения докажите следующие свойства:

1°. $\bar{\bar{z}} = z$;

2°. $\bar{z} = z \;\Leftrightarrow\; \operatorname{Im} z = 0$;

3°. $\bar{z} = -z \;\Leftrightarrow\; \operatorname{Re} z = 0$;

4°. $z + \bar{z} = 2 \operatorname{Re} z \in \mathbb{R}$;

5°. $z\bar{z} = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 \geq 0$, причем $z\bar{z} = 0 \;\Leftrightarrow\; z = 0$.

Последнее свойство объясняет вид числа z^{-1} в доказательстве теоремы 1.1:

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2} i.$$

Приведем еще несколько свойств операции комплексного сопряжения:

6°. $\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$;

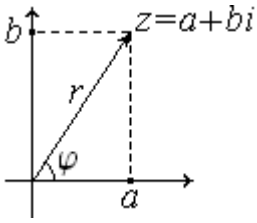
7°. $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$;

8°. $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$.

1.2 Тригонометрическая форма комплексных чисел

Как известно, вещественные числа принято изображать точками на вещественной прямой. Для комплексных чисел естественно дополнить вещественную ось абсцисс мнимой осью ординат и изображать число $z = a + bi$ точкой на плоскости с координатами (a, b).

Каждое комплексное число $z = a + bi$ можно также отождествить с вектором, выходящим из начала координат и заканчивающимся в точке (a, b). Легко видеть, что сложение/вычитание комплексных чисел согласуется со сложением/вычитанием соответствующих векторов.



Кроме декартовой системы координат на плоскости существует также полярная система координат, в которой положение точки $z = (a, b)$ характеризуется расстоянием r от начала координат и (в случае $z \neq 0$) углом φ между положительной полуосью и соответствующим числу z вектором. Используя методы школьной геометрии, нетрудно вывести равенства

$$r = \sqrt{a^2 + b^2}, \quad \cos\varphi = \frac{a}{r}, \quad \sin\varphi = \frac{b}{r}.$$

Тогда число z можно записать в виде

$$z = a + bi = r\left(\frac{a}{r} + \frac{b}{r}\,i\right) = r(\cos\varphi + i\sin\varphi),$$

который называется *тригонометрической формой* числа z. Число r называется *модулем* числа z и обозначается $|z|$, а угол φ — *аргументом* ($\arg z$). Следует подчеркнуть, что аргумент существует только у ненулевых чисел и находится из условий

$$\cos\varphi = \frac{a}{r},\ \sin\varphi = \frac{b}{r}$$

с точностью до угла, кратного 2π.

Приведем очевидные свойства модуля:

1°. $|z| = \sqrt{z\bar{z}}$;

2°. $|z| = 0 \ \Leftrightarrow \ z = 0$;

3°. $|z| = |-z|$;

4°. $|z| = |\bar{z}|$.

Пусть z_1, z_2 — ненулевые комплексные числа. Запишем их в тригонометрической форме: $z_k = r_k(\cos\varphi_k + i\sin\varphi_k)$, $k = 1, 2$. Тогда

$$z_1 z_2 = r_1 r_2(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)) = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)).$$

Ясно, что

$$|z_1 z_2| = r_1 r_2 = |z_1|\,|z_2|, \quad \arg(z_1 z_2) = \varphi_1 + \varphi_2 = \arg\, z_1 + \arg\, z_2,$$

другими словами, *при умножении комплексных чилел их модули перемножаются, а аргументы складываются.* Применяя данное правило к произведению одинаковых сомножителей, получаем **формулу Муавра**:

$$\big(r(\cos\varphi + i\sin\varphi)\big)^n = r^n(\cos(n\varphi) + i\sin(n\varphi)).$$

Если $z = r(\cos\varphi + i\sin\varphi)$, то

$$z^{-1} = \frac{\bar z}{z\bar z} = \frac{r(\cos\varphi - i\sin\varphi)}{r^2} = \frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)),$$

значит, $|z^{-1}| = |z|^{-1}$, $\arg\, z^{-1} = -\arg\, z$, так что

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \quad \arg\Big(\frac{z_1}{z_2}\Big) = \arg\, z_1 - \arg\, z_2,$$

то есть, *при делении комплексных чисел модуль первого числа делится на модуль второго и из аргумента первого числа вычитается аргумент второго.*

Напомним, что $|z|$ — это длина вектора, соответствующего числу z, поэтому приведенные выше свойства модуля можно дополнить **неравенством треугольника**:

$$\Big||z_1| - |z_2|\Big| \le |z_1 \pm z_2| \le |z_1| + |z_2|.$$

В самом деле, неравенство $|z_1 + z_2| \le |z_1| + |z_2|$ есть переформулировка известного геометрического факта, что длина любой стороны треугольника не превосходит суммы длин двух других его сторон. Воспользовавшись этим неравенством, получаем

$$|z_1| = |(z_1 + z_2) - z_2| \le |z_1 + z_2| + |-z_2| = |z_1 + z_2| + |z_2|,$$

откуда

$$|z_1| - |z_2| \le |z_1 + z_2|.$$

Аналогично доказывается неравенство $|z_2| - |z_1| \le |z_1 + z_2|$, так что

$$\Big||z_1| - |z_2|\Big| \le |z_1 + z_2| \le |z_1| + |z_2|.$$

Для завершения доказательства осталось заменить в последней цепочке неравенств z_2 на $-z_2$.$\lhd$

1.3 Извлечение корней из комплексных чисел

Как было показано выше, тригонометрическая форма комплексных чисел удобна для их умножения, деления и возведения в степень. То же относится и к извлечению корней.

Пусть $n \in \mathbb{N}$, $z \in \mathbb{C}$. Обозначим через $\sqrt[n]{z}$ множество $\{w \in \mathbb{C} : w^n = z\}$ *корней* степени n из числа z. Выясним, как оно устроено.

Пусть $w \in \sqrt[n]{z}$. Представим z и w в тригонометрической форме:

$$z = r(\cos\varphi + i\sin\varphi), \quad w = \rho(\cos\psi + i\sin\psi).$$

Тогда по формуле Муавра имеем $z = w^n = \rho^n(\cos(n\psi) + i\sin(n\psi))$, следовательно,

$$r = \rho^n, \quad n\psi = \varphi + 2\pi k,\ k \in \mathbb{Z},$$

откуда

$$\rho = \sqrt[n]{r}\ (-\text{ арифметический корень(!)}), \qquad \psi = \frac{\varphi + 2\pi k}{n}.$$

Таким образом, каждое входящее в $\sqrt[n]{z}$ число имеет вид

$$w_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right),\ k \in \mathbb{Z}. \tag{1.1}$$

Заметим, что

$$w_{n+k} = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi(k+n)}{n} + i\sin\frac{\varphi + 2\pi(k+n)}{n}\right) =$$

$$\sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right) = w_k,$$

при всех $k \in \mathbb{Z}$, так что числа w_k циклически повторяются через каждые n шагов, следовательно, окончательно имеем

$$\sqrt[n]{z} = \{w_0, w_1, \ldots, w_{n-1}\}. \tag{1.2}$$

В частности, при $z = 1$ получаем формулу для нахождения комплексных корней из единицы: $\sqrt[n]{1} = \{\epsilon_0, \epsilon_1, \ldots, \epsilon_{n-1}\}$, где

$$\epsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n} \ \text{ для всех } k.$$

Множество $\sqrt[n]{1}$ обычно обозначают через $\mathbf{U}_n$. Используя правила арифметических действий с комплексными числами в тригонометрической форме и учитывая доказанную выше циклическую повторяемость

корней ($\epsilon_k = \epsilon_l$, если $k-l$ кратно n), нетрудно проверить справедливость следующих свойств:

1°. $\epsilon_k = \epsilon_1^k$.

2°. $\epsilon_k \epsilon_l = \epsilon_{k+l}$.

3°. $\epsilon_k^{-1} = \epsilon_{-k}$.

4°. $(\mathbf{U}_n, \cdot)$ — абелева группа.

§2. Матрицы

Пусть K — поле. *Матрицей* над K называется составленная из элементов поля прямоугольная таблица, содержащая m строк и n столбцов:

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

Числа m и n называются размерами матрицы A. Элементы матрицы нумеруются парами индексов: a_{ij} — элемент матрицы A, находящийся в i-й строке и j-м столбце. Если $m = n$, то матрица A называется *квадратной*, или еще говорят, что A — матрица *порядка* n. Набор $(a_{11}, a_{22}, \dots, a_{nn})$ элементов такой матрицы называется ее *главной диагональю*, набор $(a_{1n}, a_{2,n-1}, \dots, a_{n1})$ — *побочной диагональю*. Квадратная матрица, в которой все элементы, находящиеся вне главной диагонали, равны 0, называется *диагональной*. В случае, когда равны 0 все элементы, расположенные ниже (выше) главной диагонали, матрица называется *верхнетреугольной* (*нижнетреугольной*).

В дальнейшем наиболее часто будут использоваться матрицы, составленные из вещественных или комплексных чисел. Их обычно, для краткости, называют *вещественными* или, соответственно, *комплексными* матрицами.

2.1 Действия с матрицами

Перечислим некоторые операции с матрицами.

1. *Сложение матриц. Суммой* $m \times n$*-матриц* $A = (a_{ij})$ и $B = (b_{ij})$ называется матрица $C = (c_{ij})$ тех же размеров, где $c_{ij} = a_{ij} + b_{ij}$ для всех i, j, другими словами, сложение матриц выполняется

поэлементно. Поскольку сложение в K коммутативно и ассоциативно, теми же свойствами обладает сложение матриц:

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

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

Обозначив через 0 матрицу, все элементы которой равны 0, а через $-A$ — матрицу, составленную из элементов, противоположных к элементам матрицы A, получаем еще два свойства:

1.3. Существует матрица 0 такая, что $A+0 = 0+A = A$ для любой матрицы A.

1.4. Для всякой матрицы A существует противоположная матрица $-A$ такая, что $A + (-A) = -A + A = 0$.

Свойства 1.1−1.4 означают, что множество матриц фиксированных размеров над K образует относительно сложения абелеву группу.

2. *Умножение матриц на элементы поля.* Пусть $A = (a_{ij})$ — $m \times n$-матрица, $\lambda \in K$. Под λA понимается матрица $B = (b_{ij})$ тех же размеров, где $b_{ij} = \lambda a_{ij}$ при всех i, j. Таким образом, чтобы умножить матрицу A на λ, нужно умножить на λ каждый ее элемент. Вполне очевидны свойства:

2.1. $(\lambda\mu)A = \lambda(\mu A)$.

2.2. $\lambda(A + B) = \lambda A + \lambda B$.

2.3. $(\lambda + \mu)A = \lambda A + \mu A$.

2.4. $1 \cdot A = A$ $(1 \in K)$.

В терминах линейной алгебры свойства 1.1–1.4, 2.1–2.4 означают, что множество матриц фиксированных размеров с элементами из K является векторным пространством над K.

3. *Умножение матриц.* Пусть $A = (a_{ij})$ — $m \times n$-матрица, $B = (b_{ij})$ — $n \times k$-матрица. *Произведением* матриц A и B называется $m \times k$-матрица $C = (c_{ij})$, элементы которой при всех i, j вычисляются по правилу:

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}.$$

Умножение матриц обладает следующими свойствами:

3.1. $(AB)C = A(BC)$.

3.2. $(A + B)C = AC + BC$.

3.3. $A(B + C) = AB + AC$.

3.4. Не всегда $AB = BA$.

Доказательство. Проверим свойство 3.1. Пусть A имеет размеры $m \times n$, B — $n \times k$, C — $k \times l$. Тогда и $(AB)C$, и $A(BC)$ имеют размеры $m \times l$. Проверим равенство соответствующих элементов:

$$\big((AB)C\big)_{ij} = \sum_s (AB)_{is} c_{sj} = \sum_s \Big(\sum_t a_{it} b_{ts}\Big) c_{sj} = \sum_{s,\,t} a_{it} b_{ts} c_{sj}.$$

Аналогично,

$$\big(A(BC)\big)_{ij} = \sum_{s,\,t} a_{it} b_{ts} c_{sj},$$

следовательно, $(AB)C = A(BC)$. Схожим образом устанавливается справедливость свойств 3.2 и 3.3. Наконец, равенства $AB = BA$ в 3.4 может не быть, например, из-за несоответствия размеров, а именно, если A — $m \times n$-матрица, B — $n \times k$-матрица и $m \neq k$, то произведение AB определено, а BA — нет. Если же $m = k$, но $m \neq n$, то определены оба произведения AB и BA, но они имеют разные размеры — $m \times m$ и $n \times n$. И даже если $m = n = k$ (то есть, A, B, AB, BA — квадратные матрицы одного порядка), то AB необязательно совпадает с BA, в чем легко убедиться, перемножив, например, вещественные матрицы $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ и $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. В самом деле, непосредственные вычисления показывают, что $AB = B$, а $BA = 0$.$\lhd$

Обозначим через $E_n = (\delta_{ij})$ диагональную матрицу порядка n, все диагональные элементы которой равны 1. Такая матрица называется *единичной*. (В случае, когда порядок фиксирован, нижний индекс у единичной матрицы обычно опускают и вместо E_n пишут просто E.) Числа $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ называют *символами Кронекера*. Вполне очевидно свойство

3.5. Если A — $m \times n$-матрица, то $E_m A = A E_n = A$.

Из приведенных выше рассуждений ясно, что с точки зрения выполнимости всевозможных операций наиболее удачным является множество $M_n(K)$ матриц порядка n над K. Свойства 1.1–1.4, 3.1–3.5 показывают, что $M_n(K)$ является ассоциативным некоммутативным кольцом с единицей E.

В дальнейшем нам понадобятся *матричные единицы* — матрицы E_{ij}, устроенные следующим образом: (i,j)-элемент матрицы E_{ij} равен 1, а все остальные элементы — нулевые. Легко проверяется правило умножения матричных единиц:

$$E_{ij}E_{kl} = \begin{cases} E_{il}, & \text{если } j = k, \\ 0, & \text{в противном случае.} \end{cases}$$

4. *Транспонирование. Транспонированной* к $m \times n$-матрице $A = (a_{ij})$ называется $n \times m$-матрица $A^t = (a'_{ij})$, где $a'_{ij} = a_{ji}$ для всех i, j. Таким образом, строки матрицы A являются столбцами матрицы A^t и наоборот. Свойства операции транспонирования:

4.1. $(A^t)^t = A$.

4.2. $(A + B)^t = A^t + B^t$.

4.3. $(\lambda A)^t = \lambda A^t$.

4.4. $(AB)^t = B^t A^t$.

Доказательство. Свойства 4.1–4.3 вполне очевидны. Докажем 4.4. Для всех i, j имеем:

$$\big((AB)^t\big)_{ij} = (AB)_{ji} = \sum_k a_{jk}b_{ki} = \sum_k (B^t)_{ik}(A^t)_{kj} = \big(B^tA^t\big)_{ij},$$

что и требовалось. $\lhd$

2.2 Элементарные преобразования и элементарные матрицы

Каждую матрицу можно рассматривать как упорядоченный набор строк и/или столбцов. Эти наборы можно изменять по определенным правилам, особо выделяют так называемые *элементарные преобразования*. Ниже будут рассматриваться преимущественно элементарные преобразования строк, проведение аналогичных рассуждений о преобразованиях столбцов оставляется в качестве упражнения.

1. Пусть $s \neq t$. Обозначим через $\mathcal{F}_{st}$ преобразование, меняющее местами s-ю и t-ю строки матрицы. Такое преобразование называют элементарным преобразованием I-го рода.

2. Пусть $s \neq t$, $\lambda \in K$. Прибавим к t-й строке матрицы s-ю строку, предварительно умноженную на λ. Такое преобразование обозначается через $\mathcal{F}_{st}(\lambda)$ и называется элементарным преобразованием II-го рода.

3. Умножим все элементы s-й строки на $\lambda \neq 0$. Назовем такое преобразование $\mathcal{F}_s(\lambda)$ элементарным преобразованием III-го рода.

Применим перечисленные выше преобразования к матрице A, действуя по следующему алгоритму.

1) Двигаясь сверху вниз, ищем в первом столбце отличный от нуля элемент. Если его нет, то повторяем алгоритм для матрицы, полученной из исходной вычеркиванием первого столбца. Если просмотренный нулевой столбец оказался последним, то алгоритм завершен. Теперь предположим, что найден ненулевой элемент a_{i1}. Поменяем местами 1-ю и i-ю строки (то есть, применим $\mathcal{F}_{1i}$). Получим матрицу $\tilde{A} = (\tilde{a}_{ij})$, где $\tilde{a}_{11} = a_{i1} \neq 0$.

2) Для каждого $i > 1$ последовательно применим преобразование $\mathcal{F}_{1i}\,(-\tilde{a}_{i1}/\tilde{a}_{11})$. В результате получим матрицу, в которой все элементы первого столбца (кроме, разумеется, первого элемента) равны 0. Теперь мысленно вычеркиваем из матрицы первую строку и первый столбец и, если еще остались строки и столбцы, повторяем алгоритм с шага 1) для оставшейся части матрицы; если же строк и столбцов не осталось, то работа алгоритма закончена.

Нетрудно понять, что после таких преобразований матрица примет вид

$$\begin{pmatrix} 0 & \dots & \underline{|\,\bar{a}_{1j_1}} & \dots & \bar{a}_{1j_2} & \dots & \bar{a}_{1j_r} & \dots \\ 0 & \dots & 0 & \dots & \underline{|\,\bar{a}_{2j_2}} & \dots & \bar{a}_{2j_r} & \dots \\ & & \dots & & & \dots & & \\ 0 & \dots & 0 & \dots & 0 & \dots & \underline{|\,\bar{a}_{rj_r}} & \dots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots \\ & & \dots & & & \dots & & \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots \end{pmatrix}, \tag{2.1}$$

называемый *ступенчатым*, а сам процесс преобразования матрицы называют *приведением к ступенчатому виду*. Отметим, что в матрице ступенчатого вида нулевых столбцов слева и нулевых строк внизу может не быть. Подчеркнем также, что элементы $\bar{a}_{1j_1}, \bar{a}_{2j_2}, \dots, \bar{a}_{rj_r}$, находящиеся в вершинах "ступенек", отличны от нуля. Итак, доказана

Лемма 2.1 *Любая матрица элементарными преобразованиями строк I-го и II-го рода может быть приведена к ступенчатому виду.* $\lhd$

Каждому элементарному преобразованию сопоставим *элементарную*

матрицу:

1. $F_{st} = E - E_{ss} - E_{tt} + E_{st} + E_{ts}$ — I-го рода,
2. $F_{st}(\lambda) = E + \lambda E_{ts}$ — II-го рода,
3. $F_s(\lambda) = E + (\lambda - 1)E_{ss}$ — III-го рода.

Непосредственно проверяется

Лемма 2.2 *Выполнение элементарного преобразования строк матрицы равносильно ее домножению слева на соответствующую элементарную матрицу.* $\lhd$

2.3 Обратимые матрицы

Элементарные преобразования и элементарные матрицы имеют многочисленные применения. Одно из них — проверка обратимости матрицы и нахождение обратной матрицы.

Матрица A называется *обратимой справа* (*слева*), если существует такая матрица B, что $AB = E$ ($BA = E$), при этом B называют *обратной правой* (*левой*) к A матрицей. Матрица *обратима*, если она одновременно обратима справа и слева. Отметим, что в последнем случае обратная правая и обратная левая матрицы совпадают. В самом деле, если $AB = E$ и $CA = E$, то $B = EB = CAB = CE = C$. Поэтому матрицу B называют просто *обратной* к A матрицей и обозначают A^{-1}.

Упражнение 2.1

1. Если A обратима, то $(A^t)^{-1} = (A^{-1})^t$.
2. Если A и B обратимы, то $(AB)^{-1} = B^{-1}A^{-1}$.

Лемма 2.3 *Все элементарные матрицы обратимы.*

Доказательство. Непосредственно проверяется, что $F_{st}^{-1} = F_{st}$, $(F_{st}(\lambda))^{-1} = F_{st}(-\lambda)$, $(F_s(\lambda))^{-1} = F_s(\lambda^{-1})$. $\lhd$

Предложение 2.4 *Если матрица A обратима справа, то в ней элементарными преобразованиями строк невозможно получить нулевую строку.*

Доказательство. Если A обратима справа, то для некоторой матрицы B верно равенство $AB = E$. Следовательно, в A нет нулевых строк, поскольку в противном случае произведение AB также содержало бы нулевую строку. Предположим теперь, что некоторыми элементарными

преобразованиями строк A приведена к матрице $\tilde{A}$, содержащей нулевую строку. В силу лемм 2.2 и 2.3 найдется обратимая матрица F такая, что $\tilde{A} = FA$. Тогда $\tilde{A}(BF^{-1}) = FABF^{-1} = FEF^{-1} = FF^{-1} = E$, то есть, содержащая нулевую строку матрица $\tilde{A}$ обратима справа, но выше было показано, что это невозможно. Следовательно, в A нулевую строку получить нельзя. $\lhd$

Следствие 2.5 *Если $m \times n$-матрица A обратима, то $m = n$.*

Доказательство. Матрица A обратима, следовательно, обратима справа. Воспользовавшись леммой 2.1, приведем A к ступенчатому виду $\tilde{A}$. В силу предложения 2.4 матрица $\tilde{A}$ не содержит нулевых строк, но с учетом строения матрицы ступенчатого вида (см. (2.1)) это возможно только при $m \leq n$. С помощью аналогичных рассуждений об обратимой (см. упражнение 2.1) матрице A^t получаем неравенство $n \leq m$.$\lhd$

Теорема 2.6 *Квадратная матрица A обратима справа $\Leftrightarrow$ она обратима слева.*

Доказательство. ($\Rightarrow$): Пусть A обратима справа, то есть, $AB = E$ для некоторой матрицы B. Как и в предыдущем доказательстве заметим, что матрица A после приведения к ступенчатому виду $\tilde{A}$ не содержит нулевых строк. Но для квадратной матрицы это возможно только тогда, когда вершины всех ступенек находятся на главной диагонали. Таким образом, все диагональные элементы $\tilde{a}_{11}, \ldots, \tilde{a}_{nn}$ отличны от нуля. Следовательно, элементарными преобразованиями строк II-го рода матрицу $\tilde{A}$ можно привести к диагональной матрице $\mathrm{diag}[\tilde{a}_{11}, \ldots, \tilde{a}_{nn}]$. (Сначала с помощью преобразований $\mathcal{F}_{ni}(-\tilde{a}_{in}/\tilde{a}_{nn})$ обнуляем первые $n-1$ элементов последнего столбца, затем, используя предпоследнюю строку, обнуляем первые $n-2$ элементов предпоследнего столбца и так далее.) Наконец, с помощью элементарных преобразований III-го рода диагональная матрица превращается в единичную.

Итак, обратимую справа квадратную матрицу A элементарными преобразованиями строк можно привести к единичной матрице E. Ввиду леммы 2.2 это означает, что существует такая матрица F, что $FA = E$, следовательно, A обратима слева.

($\Leftarrow$): Если A обратима слева, то $BA = E$ для некоторой матрицы B. Ясно, что B обратима справа. Тогда в силу первой части доказа-

тельства матрица B обратима слева, а значит, обратима, так что $A = B^{-1}$ и, в частности, $AB = E$. Таким образом, A обратима справа. $\lhd$

Подчеркнем важное обстоятельство: в процессе доказательства теоремы 2.6 было установлено, что обратная к A матрица является произведением элементарных матриц, соответствующих элементарным преобразованиям строк, приводящих A к единичной матрице. А именно, если $\mathcal{F}_1, \ldots, \mathcal{F}_k$ — соответствующая последовательность элементарных преобразований и $F_1, \ldots, F_k$ — последовательность отвечающих этим преобразованиям элементарных матриц, то $A^{-1} = F_k \ldots F_1$. На этой формуле основан следующий

Способ нахождения обратной матрицы.
Приписав справа к матрице A единичную матрицу того же порядка, составим $n \times 2n$-матрицу $(A|E)$. С помощью элементарных преобразований строк приведем ее к ступенчатому виду. Если при этом в матрице A появляется нулевая строка, то ввиду предложения 2.4 матрица A не имеет обратной. Если же нулевых строк нет, то согласно первой части доказательства теоремы 2.6 матрицу $(A|E)$ можно привести к такому виду, чтобы в первых ее n столбцах получилась матрица E. Тогда последние n столбцов образуют матрицу A^{-1}.

Обоснование. Элементы n первых и n последних столбцов $n \times 2n$-матрицы $(A|E)$ меняются по одним и тем же правилам, поэтому если $F_k \ldots F_1 A = E$, то $F_k \ldots F_1 E = F_k \ldots F_1 = A^{-1}$. $\lhd$

В заключение отметим, что обратные матрицы можно использовать для решения матричных уравнений. В самом деле, пусть требуется решить матричное уравнение $AX = B$, где A и B — заданные матрицы, а X — неизвестная матрица. Очевидно, что если матрица A обратима, то единственным решением будет матрица $X = A^{-1}B$. Аналогично, единственным решением уравнения $YA = B$ в случае обратимости матрицы A является матрица $Y = BA^{-1}$.

Полученные формулы можно использовать и для решения систем линейных уравнений при некоторых дополнительных условиях. Пусть

дана система

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \quad \cdots \qquad \cdots \qquad \cdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{cases}$$

Легко видеть, что она может быть переписана в матричном виде $Ax = b$, где

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Если A обратима, то система имеет единственное решение $x = A^{-1}b$.

2.4 Ранг матрицы

Для дальнейшего изложения нам понадобится ряд понятий и результатов из курса линейной алгебры. (Более подробно о перечисленных ниже фактах и их доказательствах см., напр., [2], [7].)

Система $\{a_1, \dots, a_k\}$ векторов линейного пространства V над полем K называется *линейно зависимой*, если существуют такие $\lambda_1, \dots, \lambda_k \in K$, не все равные нулю, что $\sum\limits_i \lambda_i a_i = 0$. В противном случае система называется *линейно независимой*.

Система $\{a_1, \dots, a_n\}$ называется *максимально линейно независимой*, если 1) она линейно независима, 2) добавление к ней любого вектора превращает ее в линейно зависимую систему.

Система $\{a_1, \dots, a_n\}$ называется *базисом*, если 1) она линейно независима, 2) любой вектор пространства через нее линейно выражается.

Каждая максимально линейно независимая система является базисом, и наоборот. Количество векторов в базисе пространства V называется его *размерностью* и обозначается $\dim V$. Размерность пространства не зависит от выбора базиса.

Аналогично определяется понятие базиса конечной системы векторов. Вместо слова "размерность" в этом случае употребляют термин *ранг*.

Непустое подмножество векторного пространства называется *подпространством*, если оно замкнуто относительно сложения и умножения на элементы поля.

Множество $\{\sum\limits_i \lambda_i a_i \ : \ \lambda_i \in K$ для всех $i\}$ всевозможных линейных комбинаций векторов системы $\{a_1, \ldots, a_k\}$ называется их *линейной оболочкой* и обозначается $\langle a_1, \ldots, a_k\rangle$. Линейная оболочка всегда является подпространством. Размерность линейной оболочки равна рангу образующей системы, то есть, $\dim\langle a_1, \ldots, a_k\rangle = \mathrm{rk}\{a_1, \ldots, a_k\}$. Имеет место

Лемма 2.7 *Если каждый вектор системы* $\{a_1, \ldots, a_k\}$ *линейно выражается через векторы системы* $\{b_1, \ldots, b_m\}$, *то* $\mathrm{rk}\{a_1, \ldots, a_k\} \leq \mathrm{rk}\{b_1, \ldots, b_m\}$.

Вернемся к матрицам. Каждую $m \times n$-матрицу A можно рассматривать как систему $\{A_{(1)}, \ldots, A_{(m)}\}$ строк — векторов n-мерного пространства. Ранг этой системы называется *рангом* матрицы A *по строкам* и обозначается $\mathrm{rk}_{\text{с}}(A)$. Ранг $\mathrm{rk}_{\text{в}}(A)$ системы столбцов $\{A^{(1)}, \ldots, A^{(n)}\}$ матрицы A называется ее *рангом по столбцам.*

Предложение 2.8 *Величины* $\mathrm{rk}_{\text{с}}(A)$ *и* $\mathrm{rk}_{\text{в}}(A)$ *не меняются при элементарных преобразованиях строк матрицы* A.

Доказательство. Достаточно ограничиться случаем, когда матрица $\tilde{A}$ получена из A с помощью одного элементарного преобразования.

1. Сначала докажем равенство $\mathrm{rk}_{\text{с}}(A) = \mathrm{rk}_{\text{с}}(\tilde{A})$. Возможны три случая.

1) К A было применено преобразование $\mathcal{F}_{st}$ I-го рода. В этом случае изменился лишь порядок векторов системы, но не сама система, следовательно, не изменился и ее ранг.

2) К A было применено преобразование $\mathcal{F}_{st}(\lambda)$ II-го рода. Системы строк матриц A и $\tilde{A}$ линейно выражаются друг через друга:

$$\begin{array}{cc} \tilde{A}_{(1)} = A_{(1)} & A_{(1)} = \tilde{A}_{(1)} \\ \ldots & \ldots \\ \tilde{A}_{(s)} = A_{(s)} & A_{(s)} = \tilde{A}_{(s)} \\ \ldots & \ldots \end{array}$$

$$\begin{array}{ll} \tilde{A}_{(t)} = A_{(t)} + \lambda A_{(s)} & A_{(t)} = \tilde{A}_{(t)} - \lambda \tilde{A}_{(s)} \\ \dots & \dots \\ \tilde{A}_{(n)} = A_{(n)} & A_{(n)} = \tilde{A}_{(n)} \end{array}$$

следовательно, по лемме 2.7 имеем

$$\mathrm{rk}_{\varepsilon}(A) = \mathrm{rk}\{A_{(1)}, \dots, A_{(n)}\} = \mathrm{rk}\{\tilde{A}_{(1)}, \dots, \tilde{A}_{(n)}\} = \mathrm{rk}_{\varepsilon}(\tilde{A}).$$

3) К A было применено преобразование $\mathcal{F}_s(\lambda)$ III-го рода. Как и в случае 2), системы строк матриц A и $\tilde{A}$ линейно выражаются друг через друга:

$$\begin{array}{ll} \tilde{A}_{(1)} = A_{(1)} & A_{(1)} = \tilde{A}_{(1)} \\ \dots & \dots \\ \tilde{A}_{(s)} = \lambda A_{(s)} & A_{(s)} = \lambda^{-1} \tilde{A}_{(s)} \\ \dots & \dots \\ \tilde{A}_{(n)} = A_{(n)} & A_{(n)} = \tilde{A}_{(n)} \end{array}$$

следовательно, снова $\mathrm{rk}_{\varepsilon}(A) = \mathrm{rk}\{A_{(1)}, \dots, A_{(n)}\} = \mathrm{rk}\{\tilde{A}_{(1)}, \dots, \tilde{A}_{(n)}\} = \mathrm{rk}_{\varepsilon}(\tilde{A})$.

2. Для совпадения рангов по столбцам матриц A и $\tilde{A}$ достаточно, чтобы равенство нулю произвольной линейной комбинации столбцов матрицы A выполнялось тогда и только тогда, когда равна нулю точно такая же линейная комбинация столбцов матрицы $\tilde{A}$ с теми же номерами, то есть

$$\sum_i \lambda_i A^{(i)} = 0 \;\Leftrightarrow\; \sum_i \lambda_i \tilde{A}^{(i)} = 0. \tag{2.2}$$

(Докажите, что если верно (2.2), то каждой максимально линейно независимой системе столбцов матрицы A отвечает максимально линейно независимая система столбцов матрицы $\tilde{A}$ с теми же номерами.) Составим из коэффициентов $\lambda_1, \dots, \lambda_n$ столбец $\bar{\lambda}$. Тогда условие (2.2) примет более простой вид:

$$A\bar{\lambda} = 0 \;\Leftrightarrow\; \tilde{A}\bar{\lambda} = 0. \tag{2.3}$$

Матрица $\tilde{A}$ получена из A элементарным преобразованием строк, следовательно, $\tilde{A} = FA$ для некоторой элементарной матрицы F. Теперь с учетом леммы 2.3 выполнение условия (2.3) почти очевидно: если $A\bar{\lambda} = 0$, то $\tilde{A}\bar{\lambda} = FA\bar{\lambda} = F0 = 0$, и обратно, если $\tilde{A}\bar{\lambda} = 0$, то $A\bar{\lambda} = F^{-1}FA\bar{\lambda} = F^{-1}\tilde{A}\bar{\lambda} = F^{-1}0 = 0$, что и требовалось. $\lhd$

Теорема 2.9 $\mathrm{rk}_{\text{г}}(A) = \mathrm{rk}_{\text{в}}(A)$.

Доказательство. Ввиду леммы 2.1 и предложения 2.8 можно считать, что матрица A имеет ступенчатый вид (2.1). Более того, с помощью элементарных преобразований строк II-го и III-го рода можно привести матрицу к такому виду, чтобы ее столбцы с номерами $j_1, \ldots, j_r$, проходящие через вершины "ступенек", имели вид $(1, 0, \ldots, 0)^t$, $(0, 1, \ldots, 0)^t$ и так далее, то есть, чтобы они совпадали с первыми r столбцами единичной матрицы. Очевидно, такие столбцы образуют базис системы столбцов матрицы A, так что $\mathrm{rk}_{\text{в}}(A) = r$. С другой стороны, не менее очевидно, что в матрице указанного вида первые r строк также образуют базис системы всех строк матрицы, поэтому $\mathrm{rk}_{\text{г}}(A) = r$. Следовательно, $\mathrm{rk}_{\text{г}}(A) = \mathrm{rk}_{\text{в}}(A)$. $\lhd$

Итак, ранг по строкам любой матрицы A равен ее рангу по столбцам, поэтому данную величину естественно называть просто *рангом*. Обозначение ранга матрицы: $\mathrm{rk}\, A$.

Фактически, доказательство теоремы 2.9 дает способ нахождения ранга матрицы: *ранг равен количеству ненулевых строк, остающихся в матрице после приведения ее к ступенчатому виду.*

Ранг является одной из наиболее важных характеристик матрицы. В частности, зная ранг квадратной матрицы, легко решить вопрос о ее обратимости.

Теорема 2.10 (критерий обратимости матрицы в терминах ранга) *Матрица A порядка n обратима* $\Leftrightarrow \mathrm{rk}\, A = n$.

Доказательство. Если матрица A обратима, то после приведения к ступенчатому виду в ней не должно быть нулевых строк, так что $\mathrm{rk}\, A = n$. Обратно, приведение к ступенчатому виду квадратной матрицы A ранга n дает верхнетреугольную матрицу с ненулевыми элементами по главной диагонали, а из такой матрицы с помощью элементарных преобразований строк можно получить единичную матрицу. Следовательно, матрица A обратима. $\lhd$

Теорема 2.11

$$\mathrm{rk}\,(AB) \le \min\{\mathrm{rk}\, A, \mathrm{rk}\, B\}. \tag{2.4}$$

Доказательство. Рассмотрим строки матрицы $C = AB$. Непосредственно

из правила умножения матриц получаем:

$$C_{(i)} = A_{(i)}B = \sum_j a_{ij}B_{(j)},$$

то есть, каждая строка матрицы C является линейной комбинацией строк матрицы B, следовательно,

$$\operatorname{rk} C \leq \operatorname{rk} B. \tag{2.5}$$

Аналогично, для столбцов матрицы C имеем:

$$C^{(j)} = AB^{(j)} = \sum_i A^{(i)}b_{ij},$$

то есть, столбцы матрицы C линейно выражаются чарез столбцы матрицы A, откуда

$$\operatorname{rk} C \leq \operatorname{rk} A. \tag{2.6}$$

Одновременное выполнение неравенств (2.5) и (2.6) дает (2.4). $\lhd$

Следствие 2.12 *Если матрицы B и C обратимы, то* $\operatorname{rk}(BAC) = \operatorname{rk} A$.

Доказательство. Согласно теореме 2.11 имеем

$$\operatorname{rk}(BAC) \leq \operatorname{rk}(BA) \leq \operatorname{rk} A.$$

С другой стороны, $A = B^{-1}(BAC)C^{-1}$, поэтому

$$\operatorname{rk} A = \operatorname{rk}(B^{-1}(BAC)C^{-1}) \leq \operatorname{rk}(BAC). \lhd$$

§3. Перестановки

Теперь перейдем к изучению еще одного важного класса алгебраических объектов — перестановок. Сначала дадим несколько общих определений.

Отображение $\varphi : A \to B$ называется *инъективным* (или *инъекцией*), если $a \neq b$ влечет $\varphi(a) \neq \varphi(b)$ для всех $a, b \in A$. Другими словами, отображение инъективно, если образы различных элементов различны.

Отображение $\varphi : A \to B$ называется *сюръективным* (или *сюръекцией*), если $\varphi(A) = B$, то есть, для каждого элемента $b \in B$ найдется элемент $a \in A$ такой, что $\varphi(a) = b$. Инъективное и сюръективное отображение называют *биективным* (или *биекцией*). *Тождественное*

отображение $id_A : A \to A$, переводящее каждый элемент из A в себя, очевидно является биективным.

Если $\varphi : A \to B$ и $\psi : B \to C$ — отображения, то соответствие $a \mapsto \psi(\varphi(a))$ задает отображение $\psi \circ \varphi : A \to C$, называемое *композицией* отображений φ и ψ.

Лемма 3.1 *Пусть* $\alpha : A \to B$, $\beta : B \to C$ *и* $\gamma : C \to D$ *— отображения. Тогда* $(\gamma \circ \beta) \circ \alpha = \gamma \circ (\beta \circ \alpha)$.

Доказательство. Для каждого $a \in A$ имеем

$$((\gamma \circ \beta) \circ \alpha)(a) = (\gamma \circ \beta)(\alpha(a)) = \gamma(\beta(\alpha(a))) = \gamma((\beta \circ \alpha)(a)) = (\gamma \circ (\beta \circ \alpha))(a). \triangleleft$$

Отображение $\varphi : A \to B$ называется *взаимно однозначным*, если существует *обратное* отображение $\varphi^{-1} : B \to A$, то есть, отображение, удовлетворяющее условиям $\varphi^{-1} \circ \varphi = id_A$ и $\varphi \circ \varphi^{-1} = id_B$.

Упражнение 3.1

1. Приведите примеры отображений, которые инъективны, но не сюръективны; сюръективны, но не инъективны.

2. Докажите, что композиция инъективных (сюръективных) отображений инъективна (сюръективна).

3. Докажите, что если множество A конечно, то инъективность отображения $\alpha : A \to A$ эквивалентна его сюръективности.

3.1 Группа перестановок

Пусть Ω_n — множество, состоящее из n элементов. *Перестановкой* на Ω_n называется произвольная биекция $\alpha : \Omega_n \to \Omega_n$. Обозначим через S_n множество всех перестановок на Ω_n. Композицию $\alpha \circ \beta$ перестановок $\alpha, \beta \in S_n$ будем называть их *произведением* и обозначать через $\alpha\beta$.

Для дальнейших рассуждений природа элементов множества Ω_n не имеет никакого значения, поэтому для определенности будем считать, что Ω_n состоит из чисел 1, 2, ..., n. Тогда каждую перестановку $\alpha \in S_n$ удобно записывать в виде таблицы, содержащей две строки и n столбцов:

$$\alpha = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha(1) & \alpha(2) & \dots & \alpha(n) \end{pmatrix},$$

где в первой строке перечислены элементы множества Ω_n, а во второй — их образы при действии α. Именно с такими таблицами будем в дальнейшем связывать термин "перестановка". Следует подчеркнуть, что порядок расположения столбцов таблицы, задающей перестановку, может быть произвольным.

Теорема 3.2 $(S_n, \cdot)$ — *группа.*

Доказательство. Ассоциативность произведения перестановок вытекает из леммы 3.1. Непосредственно проверяется, что единицей в S_n является тождественная перестановка $e = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$, а обратным к α элементом — перестановка $\alpha^{-1} = \begin{pmatrix} \alpha(1) & \alpha(2) & \dots & \alpha(n) \\ 1 & 2 & \dots & n \end{pmatrix}$.$\triangleleft$

Заметим, что группа S_n при $n \geq 3$ не является коммутативной. В самом деле, например, для $\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ и $\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ имеем $\alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, но $\beta\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$.

Изучим более подробно строение элементов группы S_n. Каждую перестановку $\alpha \in S_n$ можно возвести в степень, причем, полагая $\alpha^0 = e$ и $\alpha^{-k} = (\alpha^{-1})^k$, можно рассматривать степени с любым целым показателем. Вполне очевидны обычные свойства степеней: $\alpha^k\alpha^l = \alpha^{k+l}$, $(\alpha^k)^l = \alpha^{kl}$.

Фиксируем перестановку $\alpha \in S_n$. Будем говорить, что числа $i, j \in \Omega_n$ *α-эквивалентны* (обозначение: $i \sim j$), если $j = \alpha^k(i)$ для некоторого $k \in \mathbb{Z}$. Легко проверяются свойства:

1. $i \sim i$ для любого $i \in \Omega_n$ (рефлексивность).

2. Если $i \sim j$, то $j \sim i$ (симметричность).

3. Если $i \sim j$ и $j \sim k$, то $i \sim k$ (транзитивность).

(Используя терминологию теории бинарных отношений, можно сказать, что $\sim$ есть *отношение эквивалентности* на Ω_n.)

Каждому числу $i \in \Omega_n$ сопоставим множество $[i] = \{j \in \Omega_n : j \sim i\}$, называемое *классом α-эквивалентности* числа i. Заметим, что $[i] \neq \emptyset$, так как $i \in [i]$. Классы α-эквивалентности обладают следующими свойствами:

1°. Если $k \in [i]$, то $[k] = [i]$.

2°. Если $[i] \cap [j] \neq \emptyset$, то $[i] = [j]$.

3°. $\Omega_n = \bigcup\limits_{i \in \Omega_n} [i]$.

Доказательство. 1°. Поскольку $k \in [i]$, то $k \sim i$. Тогда для любого $l \in [k]$ имеем $l \sim k$ и $k \sim i$, откуда $l \sim i$ в силу транзитивности, так что $[k] \subseteq [i]$. Одновременно, с учетом симметричности, для всех $j \in [i]$ имеем $j \sim i$ и $i \sim k$, поэтому $j \sim k$, следовательно, $[i] \subseteq [k]$.

2°. Если $[i] \cap [j] \neq \emptyset$, то найдется $k \in [i] \cap [j]$. Тогда согласно предыдущему свойству $[i] = [k] = [j]$.

3°. Очевидно, $[i] \subseteq \Omega_n$ при любом $i \in \Omega_n$, поэтому $\bigcup\limits_{i \in \Omega_n} [i] \subseteq \Omega_n$. Обратно, так как $i \in [i]$ для всех $i \in \Omega_n$, то $\Omega_n = \bigcup\limits_{i \in \Omega_n} \{i\} \subseteq \bigcup\limits_{i \in \Omega_n} [i]$.$\lhd$

Из свойств 2° и 3° вытекает, что Ω_n разбивается на непересекающиеся друг с другом классы α-эквивалентности $[i_1], \ldots, [i_t]$.

Лемма 3.3 *Если класс* $[i]$ *состоит из* l *элементов, то* $[i] = \{i, \alpha(i), \ldots, \alpha^{l-1}(i)\}$, *причем* $\alpha^l(i) = i$.

Доказательство. Каждое из чисел $i, \alpha(i), \ldots, \alpha^{l-1}(i)$ лежит в $[i]$ согласно определению α-эквивалентности, поэтому для доказательства равенства множеств $[i]$ и $\{i, \alpha(i), \ldots, \alpha^{l-1}(i)\}$ достаточно установить, что все эти числа различны.

Пусть $\alpha^p(i) = \alpha^q(i)$ при $0 \leq p < q \leq l-1$. Подействовав на обе части равенства перестановкой α^{-p}, получаем $i = \alpha^s(i)$, где $0 < s = q - p < l$. Возьмем произвольный элемент $j \in [i]$. По определению для некоторого $k \in \mathbb{Z}$ верно $j = \alpha^k(i)$. Поделив k на s с остатком, получим $k = ms + r$, где $0 \leq r < s < l$. Но тогда

$$j = \alpha^k(i) = \alpha^{ms+r}(i) = \alpha^r((\alpha^s)^m(i)) = \alpha^r(i),$$

следовательно, класс $[i]$ содержит не более s элементов, что ввиду неравенства $s < l$ противоречит условию леммы.

Итак, $[i] = \{i, \alpha(i), \ldots, \alpha^{l-1}(i)\}$. В частности, $\alpha^l(i) \in [i]$, поэтому $\alpha^l(i) = \alpha^k(i)$, где $0 \leq k < l$. Тогда $i = \alpha^{l-k}(i)$, но это равенство, как было показано выше, приводит к противоречию при $k > 0$. Следовательно, $k = 0$ и $\alpha^l(i) = i$.$\lhd$

Для каждого класса $[i_s]$ $(s = 1, \ldots, t)$ рассмотрим перестановку

α_s, действующую по правилу: $\alpha_s(j) = \begin{cases} \alpha(j), & j \in [i_s] \\ j\ , & j \notin [i_s] \end{cases}$, то есть, α_s точно так же, как и α, циклически переставляет элементы из $[i_s]$, но прочие элементы оставляет неизменными. Перестановки такого рода принято называть *циклическими* или *циклами*, а количество переставляемых циклом элементов — его *длиной*. Поскольку каждый цикл α_s переставляет элементы только из своего класса эквивалентности $[i_s]$, циклы $\alpha_1, \ldots, \alpha_t$ называют *независимыми*. Ясно, что $\alpha = \alpha_1 \ldots \alpha_t$. Тем самым, доказана

Теорема 3.4 *Любая перестановка представима в виде произведения независимых циклов.*$\lhd$

Циклическую перестановку длины k, переводящую j_1 в j_2, j_2 в j_3, $\ldots$, j_{k-1} в j_k и j_k снова в j_1, удобно записывать в виде $(j_1, j_2, \ldots, j_k)$. Циклы длины 2 называют *транспозициями*.

Следствие 3.5 *Любая перестановка представима в виде произведения транспозиций.*

Доказательство. С учетом теоремы 3.4 достаточно разложить в произведение транспозиций произвольный цикл. Для упрощения обозначений можно считать, что циклически переставляются числа 1, 2,$\ldots$, k. Непосредственно проверяется, что

$$(1, 2, \ldots, k) = (1, k)(1, k-1) \ldots (1, 3)(1, 2).\lhd$$

3.2 Действие перестановок на функциях n переменных. Четность перестановки

Пусть X — непустое множество. Для каждой функции $f : X^n \to \mathbb{R}$ и каждой перестановки $\alpha \in S_n$ определим новую функцию $\alpha \circ f$ по правилу: $(\alpha \circ f)(x_1, \ldots, x_n) = f(x_{\alpha(1)}, \ldots, x_{\alpha(n)})$. Легко проверяются свойства:

1. $e \circ f = f$, где $e \in S_n$ — тождественная перестановка.
2. $\alpha \circ (\beta \circ f) = (\alpha\beta) \circ f$ для всех $\alpha, \beta \in S_n$.

(Используя терминологию теории групп, говорят, что группа перестановок *действует* на множестве функций.)

Функция f называется *кососимметрической*, если $\alpha \circ f = -f$ для любой транспозиции α вида $(i, i+1)$, иначе говоря, кососимметрическая

функция при перестановке соседних аргументов меняет знак. Простейшим примером кососимметрической функции является функция, тождественно равная нулю. Следующий пример показывает, что существуют и нетривиальные кососимметрические функции.

Пример 3.1 Рассмотрим функцию $f(x_1, \dots, x_n) = \prod\limits_{1 \le i < j \le n} (x_j - x_i)$, где $x_i \in \mathbb{R}$ для всех i. Фиксируем k в интервале от 1 до $n-1$ и разобьем произведение $\frac{n(n-1)}{2}$ входящих в f разностей на блоки:

$$f(x_1, \dots, x_n) = \overbrace{\prod_{1 \le i < j < k} (x_j - x_i)}^{A} \overbrace{\prod_{i=1}^{k-1} (x_k - x_i)}^{B} \overbrace{\prod_{i=1}^{k-1} (x_{k+1} - x_i)}^{C} (x_{k+1} - x_k) \overbrace{\prod_{\substack{1 \le i < j \le n, \\ j > k+1}} (x_j - x_i)}^{D}.$$

Очевидно, что действие транспозиции $(k, k+1)$ не изменит блоки A и D, блок B переведет в C и наоборот, а у разности $x_{k+1} - x_k$ поменяет знак. Следовательно, $(k, k+1) \circ f = -f$, так что функция f — кососимметрическая. Ясно, что $f(x_1, \dots, x_n) \neq 0$, в случае, когда значения всех ее аргументов различны.

Лемма 3.6 *Если функция f — кососимметрическая, то $\alpha \circ f = -f$ для любой транспозиции $\alpha \in S_n$.*

Доказательство. Пусть $\alpha = (i, i+k)$ — произвольная транспозиция. Проведем доказательство индукцией по разности переставляемых чисел.

При $k = 1$ доказываемое утверждение совпадает с определением кососимметрической функции. Предположим, что утверждение уже доказано для всех транспозиций с разностью, строго меньшей k. Тогда

$$(\alpha \circ f)(\dots, x_i, x_{i+1}, \dots, x_{i+k}, \dots) = f(\dots, x_{i+k}, x_{i+1}, \dots, x_i, \dots) =$$
$$-f(\dots, x_{i+1}, x_{i+k}, \dots, x_i, \dots) = f(\dots, x_{i+1}, x_i, \dots, x_{i+k}, \dots) =$$
$$-f(\dots, x_i, x_{i+1}, \dots, x_{i+k}, \dots). \triangleleft$$

Теорема 3.7 *Каждой перестановке $\alpha \in S_n$ соответствует число $\varepsilon_\alpha = \pm 1$, называемое ее четностью, такое, что $\alpha \circ f = \varepsilon_\alpha f$ для любой кососимметрической функции f от n переменных. При этом $\varepsilon_{\alpha\beta} = \varepsilon_\alpha \varepsilon_\beta$.*

Доказательство. Фиксируем неравную тождественно нулю кососимметрическую функцию f от n переменных. (Существование таких

функций установлено в примере 3.1.) Разложим перестановку $\alpha \in S_n$ в произведение транспозиций: $\alpha = \tau_1 \dots \tau_k$. Тогда с учетом леммы 3.6 имеем $\alpha \circ f = \varepsilon_\alpha f$, где $\varepsilon_\alpha = (-1)^k$. Если $\alpha = \sigma_1 \dots \sigma_m$ — еще одно разложение в произведение транспозиций, то $\alpha \circ f = (-1)^m f$, следовательно, $(\alpha \circ f)(x_1, \dots, x_n) = (-1)^m f(x_1, \dots, x_n)$ для всех $x_1, \dots, x_n$. Выберем значения $a_1, \dots, a_n$ переменных так, чтобы $f(a_1, \dots, a_n) \neq 0$. Тогда

$$(-1)^m f(a_1, \dots, a_n) = (\alpha \circ f)(a_1, \dots, a_n) = \varepsilon_\alpha f(a_1, \dots, a_n),$$

откуда, сокращая на $f(a_1, \dots, a_n)$, получаем $(-1)^m = \varepsilon_\alpha$. Тем самым доказано, что число ε_α однозначно определяется самой перестановкой α и не зависит от способа ее разложения. В частности,

$$\varepsilon_{\alpha\beta} f = (\alpha\beta) \circ f = \alpha \circ (\beta \circ f) = \alpha \circ (\varepsilon_\beta f) = \varepsilon_\alpha \varepsilon_\beta f,$$

так что $\varepsilon_{\alpha\beta} = \varepsilon_\alpha \varepsilon_\beta$. $\lhd$

Следствие 3.8 *Если $\alpha = \alpha_1 \dots \alpha_k$ — разложение в произведение независимых циклов, то $\varepsilon_\alpha = (-1)^{\sum_{s=1}^{k} (l_s - 1)}$, где $l_1, \dots, l_k$ — длины циклов.*

Доказательство вытекает из предыдущей теоремы и формулы разложения цикла в произведение транспозиций (см. следствие 3.5). $\lhd$

Приведем еще один способ вычисления четности перестановки α. Говорят, что числа i и j *образуют инверсию* относительно α, если $i < j$, но $\alpha(i) > \alpha(j)$.

Теорема 3.9 *$\varepsilon_\alpha = (-1)^k$, где k — число всех инверсий относительно α.*

Доказательство. Применим метод математической индукции по числу k инверсий.

Если $k = 0$, то α — тождественная перестановка и утверждение тривиально верно.

Пусть $k \geq 1$ и для перестановок с менее, чем k инверсиями, теорема уже доказана. Покажем сначала, что среди чисел, образующих инверсии, есть соседние числа. В самом деле, пусть инверсию образуют числа i и $i + l$. Если $l = 1$, то нужные соседние числа найдены. При $l > 1$ рассмотрим $\alpha(i+1)$. Если $\alpha(i) > \alpha(i+1)$, то инверсию образуют соседние числа i и $i + 1$. Если же $\alpha(i) < \alpha(i + 1)$, то $\alpha(i + 1) > \alpha(i) > \alpha(i + l)$,

следовательно, инверсию образуют числа $i+1$ и $i+l$, находящиеся на одну позицию ближе друг к другу, чем i и $i+l$. Ясно, что через конечное число шагов нужная пара соседних чисел будет найдена.

Итак, пусть i и $i+1$ образуют инверсию. Рассмотрим перестановку $\alpha' = \tau\alpha$, где $\tau = (\alpha(i), \alpha(i+1))$. Нетрудно видеть, что α' содержит ровно на одну инверсию меньше, чем α, значит, по предположению индукции $\varepsilon_{\alpha'} = (-1)^{k-1}$. Но тогда из равенства $\alpha = \tau\alpha'$ выводим $\varepsilon_\alpha = \varepsilon_\tau\varepsilon_{\alpha'} = (-1)(-1)^{k-1} = (-1)^k$.$\triangleleft$

§4. Определители

Пусть $A = (a_{ij})$ — матрица порядка n над полем K. Сопоставим ей элемент поля, вычисляемый по формуле

$$\det A = \sum_{\alpha\in S_n} \varepsilon_\alpha a_{1\alpha(1)}a_{2\alpha(2)}\dots a_{n\alpha(n)} \tag{4.1}$$

и называемый ее *определителем*. Также для обозначения определителя используется запись матрицы в прямых скобках. Выпишем формулу (4.1), называемую *формулой полного развертывания*, в явном виде при малых значениях n:

$n = 1$. $\det A = a_{11}$.

$n = 2$. $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$.

$n = 3$. $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$.

При $n > 3$ количество слагаемых в правой части (4.1) быстро возрастает (при $n = 4$ формула содержит 24 слагаемых, при $n = 5$ — уже 120), поэтому для вычисления определителей больших порядков, как правило, используются специальные приемы, опирающиеся на различные свойства определителей, о которых пойдет речь ниже.

4.1 Определитель транспонированной матрицы

Следующая теорема показывает, что определитель матрицы не меняется при ее транспонировании:

Теорема 4.1 $\det A^t = \det A$.

Доказательство. Пусть $A = (a_{ij})$, $A^t = (a'_{ij})$, где $a'_{ij} = a_{ji}$ при всех i, j. Тогда согласно (4.1) имеем

$$\det A^t = \sum_{\alpha \in S_n} \varepsilon_\alpha a'_{1\alpha(1)} a'_{2\alpha(2)} \dots a'_{n\alpha(n)} = \sum_{\alpha \in S_n} \varepsilon_\alpha a_{\alpha(1)1} a_{\alpha(2)2} \dots a_{\alpha(n)n}.$$

Заметим, что пары индексов элементов, участвующих в произведении $a_{\alpha(1)1} a_{\alpha(2)2} \dots a_{\alpha(n)n}$, отвечают перестановке, обратной к α, поэтому с учетом коммутативности умножения элементов поля получаем $a_{\alpha(1)1} a_{\alpha(2)2} \dots a_{\alpha(n)n} = a_{1\alpha^{-1}(1)} a_{2\alpha^{-1}(2)} \dots a_{n\alpha^{-1}(n)}$, откуда

$$\det A^t = \sum_{\alpha \in S_n} \varepsilon_\alpha a_{1\alpha^{-1}(1)} a_{2\alpha^{-1}(2)} \dots a_{n\alpha^{-1}(n)}. \tag{4.2}$$

Воспользуемся теперь тем, что отображение $\alpha \mapsto \alpha^{-1}$ является биекцией группы перестановок S_n в себя (проверьте это в качестве упражнения(!)), следовательно, если перестановка α пробегает всю группу, то всю группу пробегает и перестановка α^{-1}, поэтому в (4.2) можно заменить индекс суммирования:

$$\sum_{\alpha \in S_n} \varepsilon_\alpha a_{1\alpha^{-1}(1)} a_{2\alpha^{-1}(2)} \dots a_{n\alpha^{-1}(n)} = \sum_{\alpha^{-1} \in S_n} \varepsilon_\alpha a_{1\alpha^{-1}(1)} a_{2\alpha^{-1}(2)} \dots a_{n\alpha^{-1}(n)}.$$

Поскольку $\alpha\alpha^{-1} = e$, то $1 = \varepsilon_e = \varepsilon_{\alpha\alpha^{-1}} = \varepsilon_\alpha \varepsilon_{\alpha^{-1}}$, откуда с учетом $\varepsilon_\alpha = \pm 1$ выводим $\varepsilon_\alpha = \varepsilon_{\alpha^{-1}}$. Обозначив α^{-1} через β, окончательно получаем

$$\det A^t = \sum_{\beta \in S_n} \varepsilon_\beta a_{1\beta(1)} a_{2\beta(2)} \dots a_{n\beta(n)} = \det A. \lhd$$

4.2 Определитель, как полилинейная кососимметрическая функция строк (столбцов) матрицы

Как уже не раз отмечалось, каждую матрицу A порядка n можно рассматривать как набор $(A_{(1)}, \dots, A_{(n)})$ n-мерных векторов — строк матрицы, либо как набор $(A^{(1)}, \dots, A^{(n)})$ векторов — столбцов. Поэтому на определитель матрицы можно смотреть как на функцию от n аргументов. При транспонировании матрицы ее строки становятся столбцами и наоборот, столбцы — строками, определитель же согласно теореме 4.1 не меняется. Следовательно, любое свойство, доказанное для определителя, рассматриваемого как функция строк, автоматически

будет верным и в том случае, когда определитель рассматривается в качестве функции столбцов.

Пусть V — векторное пространство над полем K. (Будем считать, что K — одно из полей $\mathbb{Q}$, $\mathbb{R}$ или $\mathbb{C}$.) Функция $f : V^n \to K$ называется *полилинейной*, если для каждого $i = 1, \dots, n$ справедливо равенство

$$f(a_1, \dots, \lambda a'_i + \mu a''_i, \dots, a_n) = \lambda f(a_1, \dots, a'_i, \dots, a_n) + \mu f(a_1, \dots, a''_i, \dots, a_n),$$

то есть, функция f линейна по каждому аргументу.

D1. $\det(A_{(1)}, \dots, A_{(n)})$ — *полилинейная функция.*

Доказательство вытекает из цепочки равенств

$$\det(A_{(1)}, \dots, \lambda A'_{(i)} + \mu A''_{(i)}, \dots, A_{(n)}) = \sum_{\alpha \in S_n} \varepsilon_\alpha a_{1\alpha(1)} \dots (\lambda a'_{i\alpha(i)} + \mu a''_{i\alpha(i)}) \dots a_{n\alpha(n)} =$$
$$\lambda \sum_{\alpha \in S_n} \varepsilon_\alpha a_{1\alpha(1)} \dots a'_{i\alpha(i)} \dots a_{n\alpha(n)} + \mu \sum_{\alpha \in S_n} \varepsilon_\alpha a_{1\alpha(1)} \dots a''_{i\alpha(i)} \dots a_{n\alpha(n)} =$$
$$\lambda \det(A_{(1)}, \dots, A'_{(i)}, \dots, A_{(n)}) + \mu \det(A_{(1)}, \dots, A''_{(i)}, \dots, A_{(n)}). \triangleleft$$

D2. $\det(A_{(1)}, \dots, A_{(n)})$ — *кососимметрическая функция.*

Доказательство. Обозначим через τ транспозицию $(i, i{+}1)$. Для любой перестановки α верно $\varepsilon_{\alpha\tau} = \varepsilon_\alpha \varepsilon_\tau = -\varepsilon_\alpha$. Следовательно,

$$\det(\dots, A_{(i+1)}, A_{(i)}, \dots) = \sum_{\alpha \in S_n} \varepsilon_\alpha a_{1,\,\alpha(1)} \dots a_{i,\,\alpha(i+1)} a_{i+1,\,\alpha(i)} \dots a_{n,\,\alpha(n)} =$$
$$- \sum_{\alpha \in S_n} \varepsilon_{\alpha\tau} a_{1,(\alpha\tau)(1)} \dots a_{i,(\alpha\tau)(i)} a_{i+1,(\alpha\tau)(i+1)} \dots a_{n,(\alpha\tau)(n)}.$$

Легко проверяется, что отображение $\alpha \mapsto \alpha\tau$ является биекцией группы S_n в себя, поэтому в последнем выражении можно заменить индекс суммирования α на $\beta = \alpha\tau$, так что

$$\det(\dots, A_{(i+1)}, A_{(i)}, \dots) = - \sum_{\beta \in S_n} \varepsilon_\beta a_{1,\,\beta(1)} \dots a_{i,\,\beta(i)} a_{i+1,\,\beta(i+1)} \dots a_{n,\,\beta(n)} =$$
$$- \det(\dots, A_{(i)}, A_{(i+1)}, \dots). \triangleleft$$

Замечание. Ввиду леммы 3.6 свойство D2 означает, что если в матрице поменять местами две строки, то ее определитель поменяет знак.

D3. $\det E = 1$.

Доказательство. Согласно (4.1) имеем

$$\det E = \sum_{\alpha \in S_n} \varepsilon_\alpha \delta_{1\alpha(1)} \dots \delta_{n\alpha(n)}.$$

Если $\alpha(i) \neq i$ для некоторого i, то $\delta_{i\alpha(i)} = 0$, следовательно, равно нулю и все содержащее $\delta_{i\alpha(i)}$ произведение, поэтому под знаком суммы в последнем выражении остается только произведение, отвечающее тождественной перестановке, то есть, $\det E = \varepsilon_e \delta_{11} \dots \delta_{nn}$. Но $\varepsilon_e = 1$ и $\delta_{ii} = 1$ при любом i, значит, $\det E = 1$.$\lhd$

4.3 Свойства полилинейных кососимметрических функций

Пусть $\mathcal{D}$ — произвольная полилинейная кососимметрическая функция строк матрицы A, то есть, функция, обладающая свойствами D1 и D2.

D4. $\mathcal{D}(\lambda A) = \lambda^n \mathcal{D}(A)$.

Доказательство. Воспользовавшись n раз свойством D1, получаем

$$\mathcal{D}(\lambda A) = \mathcal{D}(\lambda A_{(1)}, \dots, \lambda A_{(n)}) = \lambda^n \mathcal{D}(A_{(1)}, \dots, A_{(n)}) = \lambda^n \mathcal{D}(A).\lhd$$

D5. *Если A содержит нулевую строку, то $\mathcal{D}(A) = 0$.*

Доказательство. Пусть $A_{(i)} = 0$. Тогда $A_{(i)} = 0 \cdot A_{(i)}$, следовательно, в силу D1 имеем

$$\mathcal{D}(A) = \mathcal{D}(A_{(1)}, \dots, A_{(i)}, \dots, A_{(n)}) = \\ \mathcal{D}(A_{(1)}, \dots, 0 \cdot A_{(i)}, \dots, A_{(n)}) = 0 \cdot \mathcal{D}(A) = 0.\lhd$$

D6. *Если матрица A содержит две одинаковых строки, то $\mathcal{D}(A) = 0$.*

Доказательство. Пусть $i \neq j$ и $A_{(i)} = A_{(j)}$. Согласно D2, если в матрице A поменять местами i-ю и j-ю строки, то $\mathcal{D}(A)$ поменяет знак. С другой стороны, от перестановки одинаковых строк матрица A не изменится. Следовательно, $\mathcal{D}(A) = -\mathcal{D}(A)$, откуда, $\mathcal{D}(A) = 0$.

D7. *При элементарных преобразованиях строк II-го рода матрицы A значение $\mathcal{D}(A)$ не меняется.*

Доказательство. Воспользуемся свойствами D1 и D6:

$$\mathcal{D}(\dots, A_{(i)}, \dots, \lambda A_{(i)} + A_{(j)}, \dots) = \lambda \mathcal{D}(\dots, A_{(i)}, \dots, A_{(i)}, \dots) +$$

$$\mathcal{D}(\dots, A_{(i)}, \dots, A_{(j)}, \dots) = \lambda \cdot 0 + \mathcal{D}(A) = \mathcal{D}(A). \lhd$$

Подчеркнем, что свойства D4–D7 справедливы для всех полилинейных кососимметрических функций строк матриц, а значит, справедливы для определителя как частного случая таких функций. На самом деле, между определителями и полилинейными кососимметрическими функциями существует и обратная связь: каждую полилинейную кососимметрическую функцию строк матрицы можно выразить через ее определитель. Предварительно докажем вспомогательную лемму.

Лемма 4.2 *Если $A = (a_{ij})$ — верхнетреугольная матрица порядка n, то $\mathcal{D}(A) = a_{11}a_{22}\dots a_{nn}\mathcal{D}(E)$.*

Доказательство. Последняя строка матрицы A имеет вид $A_{(n)} = (0, 0, \dots, a_{nn})$. Поскольку $A_{(n)} = a_{nn}(0, 0, \dots, 1)$, то согласно D1

$$\mathcal{D}(A) = a_{nn}\mathcal{D}\begin{pmatrix} a_{11} & \dots & a_{1,n-1} & a_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \dots & 0 & 1 \end{pmatrix}. \tag{4.3}$$

Легко видеть, что элементарными преобразованиями II-го рода строк полученной в (4.3) матрицы ее последний столбец можно привести к виду $(0, \dots, 0, 1)^t$, поэтому в силу D7

$$\mathcal{D}(A) = a_{nn}\mathcal{D}\begin{pmatrix} a_{11} & \dots & a_{1,n-1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & a_{n-1,n-1} & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Повторив аналогичные преобразования матрицы для строк с номерами $n-1, \dots, 1$, получим требуемое равенство. $\lhd$

Теорема 4.3 *Если $\mathcal{D}$ — полилинейная кососимметрическая функция строк матрицы A, то $\mathcal{D}(A) = \mathcal{D}(E)\det A$.*

Доказательство. С помощью элементарных преобразований I-го и II-го рода приведем матрицу A к ступенчатому виду $\tilde{A}$. Пусть при этом было выполнено k преобразований I-го рода. В силу D7 преобразования II-го рода не меняют значения $\mathcal{D}(A)$, а каждое преобразование I-го рода

согласно D2 меняет знак, так что

$$\mathcal{D}(A) = (-1)^k \mathcal{D}(\tilde{A}).$$

Матрица $\tilde{A} = (\tilde{a}_{ij})$ имеет ступенчатый вид (см. (2.1)), следовательно, она является верхнетреугольной матрицей. Тогда по лемме 4.2 имеем $\mathcal{D}(\tilde{A}) = \tilde{a}_{11} \dots \tilde{a}_{nn} \mathcal{D}(E)$, откуда

$$\mathcal{D}(A) = \mathcal{D}(E)(-1)^k \tilde{a}_{11} \dots \tilde{a}_{nn}. \tag{4.4}$$

Заметим, что приведенные выше рассуждения остаются верными, если функцию $\mathcal{D}$ заменить определителем, при этом (4.4) с учетом D3 примет вид

$$\det A = (-1)^k \tilde{a}_{11} \dots \tilde{a}_{nn},$$

следовательно, $\mathcal{D}(A) = \mathcal{D}(E)(-1)^k \tilde{a}_{11} \dots \tilde{a}_{nn} = \mathcal{D}(E) \det A$. $\lhd$

В качестве применения теоремы 4.3 докажем две следующих теоремы об определителях.

Теорема 4.4 *Пусть* $A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$ *— блочно-верхнетреугольная матрица, причем блоки* B *и* C *— квадратные. Тогда*

$$\det A = \det B \, \det C.$$

Доказательство. Фиксируем блоки B, D и рассмотрим функцию $\mathcal{D}(C) = \begin{vmatrix} B & D \\ 0 & C \end{vmatrix}$. Обозначим через k и l порядки блоков B и C соответственно. Очевидно, если $C_{(i)} = \lambda C'_{(i)} + \mu C''_{(i)}$, то $A_{(k+i)} = \lambda A'_{(k+i)} + \mu A''_{(k+i)}$, поэтому

$$\begin{gathered} \mathcal{D}(C_{(1)}, \dots, \lambda C'_{(i)} + \mu C''_{(i)}, \dots, C_{(l)}) = \\ \det(A_{(1)}, \dots, \lambda A'_{(k+i)} + \mu A''_{(k+i)}, \dots, A_{(k+l)}) = \\ \lambda \det(A_{(1)}, \dots, A'_{(k+i)}, \dots, A_{(k+l)}) + \mu \det(A_{(1)}, \dots, A''_{(k+i)}, \dots, A_{(k+l)}) = \\ \lambda \mathcal{D}(C_{(1)}, \dots, C'_{(i)}, \dots, C_{(l)}) + \mu \mathcal{D}(C_{(1)}, \dots, C''_{(i)}, \dots, C_{(l)}), \end{gathered}$$

следовательно, функция $\mathcal{D}$ — полилинейная. Аналогично,

$$\mathcal{D}(\dots, C_{(i+1)}, C_{(i)}, \dots) = \det(\dots, A_{(k+i+1)}, A_{(k+i)}, \dots) =$$

$$- \det(\dots, A_{(k+i)}, A_{(k+i+1)}, \dots) = -\mathcal{D}(\dots, C_{(i)}, C_{(i+1)}, \dots),$$

так что $\mathcal{D}$ — кососимметрическая. Тогда по теореме 4.3

$$\mathcal{D}(C) = \mathcal{D}(E) \det C.$$

Осталось показать, что $\mathcal{D}(E) = \det B$. Положим $\tilde{A} = \begin{pmatrix} B & D \\ 0 & E \end{pmatrix}$. Имеем

$$\mathcal{D}(E) = \det \tilde{A} = \sum_{\alpha \in S_{k+l}} \varepsilon_\alpha \tilde{a}_{1,\alpha(1)} \dots \tilde{a}_{k+l,\alpha(k+l)}. \tag{4.5}$$

Заметим, что $\tilde{a}_{ij} = \delta_{ij}$ при $i > k$, поэтому можно считать, что суммирование в (4.5) ведется только по тем перестановкам α, для которых $\alpha(k+1) = k+1, \dots, \alpha(k+l) = k+l$, то есть, фактически, по перестановкам β чисел 1, 2,..., k. Полагая $\beta(i) = \alpha(i)$ при $i \le k$, с учетом $\tilde{a}_{ij} = b_{ij}$ $(i, j \le k)$ получаем

$$\sum_{\alpha \in S_{k+l}} \varepsilon_\alpha \tilde{a}_{1,\alpha(1)} \dots \tilde{a}_{k+l,\alpha(k+l)} = \sum_{\beta \in S_k} \varepsilon_\beta \tilde{a}_{1,\beta(1)} \dots \tilde{a}_{k,\beta(k)} \delta_{k+1,k+1} \dots \delta_{k+l,k+l} =$$

$$\sum_{\beta \in S_k} \varepsilon_\beta b_{1,\beta(1)} \dots b_{k,\beta(k)} = \det B.$$

Подставив полученное выражение в (4.5), выводим требуемое равенство $\mathcal{D}(E) = \det B$.$\lhd$

Теорема 4.5 *Пусть A, B — квадратные матрицы одного порядка. Тогда*

$$\det(AB) = \det A \det B.$$

Доказательство. Фиксируем матрицу B и рассмотрим функцию $\mathcal{D}(A) = \det(AB)$. Легко видеть, что $(AB)_{(i)} = A_{(i)}B$ для всех i. Тогда

$$\begin{gathered} \mathcal{D}(A_{(1)}, \dots, \lambda A'_{(i)} + \mu A''_{(i)}, \dots, A_{(n)}) = \\ \det(A_{(1)}B, \dots, \lambda A'_{(i)}B + \mu A''_{(i)}B, \dots, A_{(n)}B) = \\ \lambda \det(A_{(1)}B, \dots, A'_{(i)}B, \dots, A_{(n)}B) + \mu \det(A_{(1)}B, \dots, A''_{(i)}B, \dots, A_{(n)}B) = \\ \lambda \mathcal{D}(A_{(1)}, \dots, A'_{(i)}, \dots, A_{(n)}) + \mu \mathcal{D}(A_{(1)}, \dots, A''_{(i)}, \dots, A_{(n)}), \end{gathered}$$

так что функция $\mathcal{D}$ — полилинейная. Аналогично,

$$\mathcal{D}(\dots, A_{(i+1)}, A_{(i)}, \dots) = \det(\dots, A_{(i+1)}B, A_{(i)}B, \dots) =$$

$$-\det(\dots, A_{(i)}B, A_{(i+1)}B, \dots) = -\mathcal{D}(\dots, A_{(i)}, A_{(i+1)}, \dots),$$

значит, $\mathcal{D}$ — кососимметрическая. Ввиду теоремы 4.3

$$\det(AB) = \mathcal{D}(A) = \mathcal{D}(E) \det A = \det(EB) \det A = \det A \det B.\lhd$$

4.4 Разложение определителя по строке (столбцу)

Пусть $A = (a_{ij})$ матрица порядка n. *Дополняющим минором* M_{ij} элемента a_{ij} называется определитель матрицы, полученной вычеркиванием из матрицы A i-й строки и j-го столбца. Число $A_{ij} = (-1)^{i+j} M_{ij}$ называется *алгебраическим дополнением* элемента a_{ij}.

Теорема 4.6 *Пусть $A = (a_{ij})$ — матрица порядка n. Для всех i справедливы равенства*

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik} = \sum_{k=1}^{n} a_{ik} A_{ik}, \tag{4.6}$$

$$\det A = \sum_{k=1}^{n} (-1)^{k+i} a_{ki} M_{ki} = \sum_{k=1}^{n} a_{ki} A_{ki}, \tag{4.7}$$

называемые формулами разложения определителя по строке и, соответственно, столбцу.

Доказательство. Сначала докажем (4.7). Представим i-й столбец матрицы A в виде суммы n столбцов:

$$A^{(i)} = (a_{1i}\, 0 \ldots 0)^t + \cdots + (0\, 0 \ldots a_{ni})^t.$$

Тогда

$$\det A = \sum_{k=1}^{n} \begin{vmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ & \ldots & & \ldots & \\ a_{k1} & \ldots & a_{ki} & \ldots & a_{kn} \\ & \ldots & & \ldots & \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{vmatrix}.$$

Вычислим отдельно k-е слагаемое получившейся суммы. Последовательно меняя местами i-й столбец с каждым предыдущим столбцом, переставим его на место первого столбца. При этом мы $i-1$ раз применили элементарное преобразование I-го рода, поэтому

$$\begin{vmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ & \ldots & & \ldots & \\ a_{k1} & \ldots & a_{ki} & \ldots & a_{kn} \\ & \ldots & & \ldots & \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{vmatrix} = (-1)^{i-1} \begin{vmatrix} 0 & a_{11} & \ldots & a_{1n} \\ & \ldots & \ldots & \\ a_{ki} & a_{k1} & \ldots & a_{kn} \\ & \ldots & \ldots & \\ 0 & a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Теперь аналогичным образом переместим k-ю строку на место первой строки:

$$(-1)^{i-1}\begin{vmatrix} 0 & a_{11} & \dots & a_{1n} \\ & \dots & \dots & \\ a_{ki} & a_{k1} & \dots & a_{kn} \\ & \dots & \dots & \\ 0 & a_{n1} & \dots & a_{nn} \end{vmatrix} = (-1)^{i-1+k-1}\begin{vmatrix} a_{ki} & a_{k1} & \dots & a_{kn} \\ 0 & a_{11} & \dots & a_{1n} \\ & \dots & \dots & \\ 0 & a_{n1} & \dots & a_{nn} \end{vmatrix} =$$

$$(-1)^{k+i}\begin{vmatrix} a_{ki} & a_{k1} & \dots & a_{kn} \\ 0 & a_{11} & \dots & a_{1n} \\ & \dots & \dots & \\ 0 & a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Заметим, что матрица, стоящая под знаком определителя в правой части последней цепочки равенств, имеет блочно-верхнетреугольный вид $\begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$, где $B = (a_{ki})$ — матрица порядка 1, а блок C получается вычеркиванием из матрицы A k-й строки и i-го столбца. Следовательно, согласно теореме 4.4

$$\begin{vmatrix} a_{ki} & a_{k1} & \dots & a_{kn} \\ 0 & a_{11} & \dots & a_{1n} \\ & \dots & \dots & \\ 0 & a_{n1} & \dots & a_{nn} \end{vmatrix} = \det B \, \det C = a_{ki}M_{ki},$$

так что окончательно получаем

$$\det A = \sum_{k=1}^{n}\begin{vmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ & \dots & & \dots & \\ a_{k1} & \dots & a_{ki} & \dots & a_{kn} \\ & \dots & & \dots & \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{vmatrix} = \sum_{k=1}^{n}(-1)^{k+i}a_{ki}M_{ki} = \sum_{k=1}^{n} a_{ki}A_{ki},$$

то есть, верно (4.7).

Легко видеть, что если в матрице A вычеркнуть i-ю строку и j-й столбец, а в матрице A^t — j-ю строку и i-й столбец, то получившиеся матрицы порядка $n-1$ являются транспонированными друг к другу, следовательно, их определители равны. Таким образом, минор M_{ij}

матрицы A равен минору M'_{ji} матрицы $A^t = (a'_{ij})$. Применив (4.7) к определителю матрицы A^t, получаем

$$\det A = \det A^t = \sum_{k=1}^{n} (-1)^{k+i} a'_{ki} M'_{ki} = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} M_{ik} = \sum_{k=1}^{n} a_{ik} A_{ik},$$

то есть, справедливо (4.6). $\lhd$

4.5 Применение определителей

Воспользуемся полученными выше результатами об определителях для решения следующих задач.

1. *Критерий обратимости матрицы в терминах определителя.*

Пусть $A = (a_{ij})$ — матрица порядка n. Для каждого элемента матрицы A найдем его алгебраическое дополнение и составим из них *присоединенную* матрицу

$$A^{\vee} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}.$$

Лемма 4.7 $AA^{\vee} = (\det A)E$.

Доказательство. С учетом формулы (4.6) имеем

$$(AA^{\vee})_{ii} = \sum_{k=1}^{n} a_{ik} A_{ik} = \det A.$$

Пусть теперь $i \neq j$. Обозначим через $\bar{A}$ матрицу, которая получается, если в A j-ю строку заменить на i-ю. Так как $\bar{A}$ содержит две одинаковых строки, ее определитель равен 0. С другой стороны, разложив $\det \bar{A}$ по j-й строке, получаем

$$\det \bar{A} = \sum_{k=1}^{n} \bar{a}_{jk} A_{jk} = \sum_{k=1}^{n} a_{ik} A_{jk} = (AA^{\vee})_{ij},$$

так что $(AA^{\vee})_{ij} = 0$ при $i \neq j$. Следовательно, $AA^{\vee} = (\det A)E$. $\lhd$

Теорема 4.8 (критерий обратимости) *Матрица A обратима тогда и только тогда, когда* $\det A \neq 0$. *Если* $\det A \neq 0$, *то*

$$A^{-1} = \frac{1}{\det A} A^{\vee}. \tag{4.8}$$

Доказательство. Если матрица A обратима, то $AB = E$ для некоторой матрицы B. Тогда в силу теоремы 4.5

$$1 = \det E = \det(AB) = \det A \det B,$$

следовательно, $\det A \neq 0$.

Теперь докажем обратное утверждение. Пусть $\det A \neq 0$. Положим $B = \dfrac{1}{\det A} A^{\vee}$. Тогда с учетом леммы 4.7

$$AB = A\Big(\frac{1}{\det A} A^{\vee}\Big) = \frac{1}{\det A}(AA^{\vee}) = E,$$

то есть, A обратима справа. Поскольку для квадратной матрицы одно- и двусторонняя обратимость эквивалентны (см. теорему 2.6), получаем, что A обратима и $A^{-1} = B = \dfrac{1}{\det A} A^{\vee}$.$\lhd$

Из теорем 4.8 и 2.10 вытекает

Следствие 4.9 *Пусть A — матрица порядка n. Тогда* $\det A \neq 0 \Leftrightarrow \operatorname{rk} A = n$.$\lhd$

2. *Формулы Крамера.*

Рассмотрим систему линейных уравнений, записанную в матричном виде $Ax = b$. Как известно (см. п. 2.3 "Обратимые матрицы"), единственное решение данной системы в случае обратимости матрицы A находится по формуле $x = A^{-1}b$. Воспользовавшись формулой (4.8) для обратной матрицы, выводим

$$x = \frac{1}{\det A} A^{\vee} b. \tag{4.9}$$

Обозначим через $\triangle$ определитель матрицы A, а через $\triangle_i$ $(i = 1, \ldots, n)$ — определители матриц, которые получаются, если в матрице A заменить i-й столбец столбцом b. В частности, производя разложение определителя $\triangle_i$ по i-му столбцу, имеем

$$\triangle_i = \sum_k b_k A_{ki}.$$

Переходя в (4.9) к поэлементной записи, получаем

$$x_i = \frac{1}{\triangle} \sum_k b_k A_{ki} = \frac{\triangle_i}{\triangle}, \quad i = 1, \ldots, n. \tag{4.10}$$

Равенства (4.10) называются **формулами Крамера**.

3. *Метод окаймляющих миноров.*

Пусть $A = (a_{ij})$ — $m \times n$-матрица. Фиксируем индексы $1 \le i_1 \le i_2 \le \cdots \le i_k \le m$ и $1 \le j_1 \le j_2 \le \cdots \le j_k \le n$. Определитель

$$\begin{vmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_k} \\ \vdots & \ddots & \vdots \\ a_{i_k j_1} & \dots & a_{i_k j_k} \end{vmatrix}$$

называется *минором* порядка k матрицы A. Минор $\tilde{M}$ порядка $k+1$ называется *окаймляющим* для минора M, если M получается из $\tilde{M}$ вычеркиванием одной крайней (первой или последней) строки и одного крайнего столбца.

Алгоритм нахождения ранга матрицы.

Шаг 1. Ищем в матрице A ненулевой элемент a_{ij}. Если такого нет, то $\operatorname{rk} A = 0$, в противном случае найден отличный от нуля минор $M = a_{ij}$ порядка 1.

Шаг 2. Пусть уже найден минор $M \neq 0$ порядка k. Ищем отличный от нуля минор $\tilde{M}$ порядка $k+1$ среди миноров, окаймляющих M. Если такой минор есть, то повторяем шаг 2 для минора $\tilde{M}$, а если все окаймляющие миноры равны нулю, то $\operatorname{rk} A = k$.

Обоснование алгоритма. Ясно, что алгоритм может остановиться на шаге 1 только в том случае, когда матрица A — нулевая и, следовательно, ее ранг равен 0.

Пусть теперь $A \neq 0$. Покажем сначала, что алгоритм не может работать бесконечно долго. В самом деле, в таком случае из строк и столбцов матрицы A можно было бы составлять ненулевые миноры сколь угодно большого порядка. Но матрица A имеет m строк, поэтому любой ее минор порядка $m+1$ содержит по крайней мере две одинаковых строки и, следовательно, равен нулю. (Аналогично показывается, что порядок отличного от нуля минора матрицы не может быть больше количества ее столбцов, поэтому, если k — порядок ненулевого минора матрицы A, то $k \le \min(m, n)$.)

Итак, через конечное число шагов алгоритм завершит работу и будет найден минор $M \neq 0$ порядка r, все окаймляющие миноры которого равны 0. Докажем, что $\operatorname{rk} A = r$. Для простоты обозначений будем считать, что минор M образован элементами первых r строк

и первых r столбцов матрицы A. (Такого расположения минора M всегда можно добиться подходящими перестановками строк и столбцов матрицы, при этом ее ранг не меняется.) Разобьем матрицу A на блоки: $A = \begin{pmatrix} B & C \\ D & F \end{pmatrix}$, где блок B — матрица порядка r (размеры остальных блоков легко вычисляются, исходя из размеров B и A, в частности, если $r = m$, то $A = (B|C)$). Ясно, что $\det B = M \neq 0$, откуда ввиду следствия 4.9 выводим $\operatorname{rk} B = r$. Тогда и $\operatorname{rk}(B|C) = r$, так как $r = \operatorname{rk} B \leq \operatorname{rk}(B|C) = \operatorname{rk}\{A_{(1)}, \ldots, A_{(r)}\} \leq r$. Если $r = m$, то требуемое равенство $\operatorname{rk} A = r$ получено.

Пусть теперь $r < m$. Фиксируем номера i, j строки и столбца матрицы A и рассмотрим окаймляющий M минор

$$\tilde{M} = \begin{vmatrix} a_{11} & \ldots & a_{1r} & a_{1j} \\ \vdots & \ddots & \vdots & \vdots \\ a_{r1} & \ldots & a_{rr} & a_{rj} \\ a_{i1} & \ldots & a_{ir} & a_{ij} \end{vmatrix}.$$

Согласно предположению $\tilde{M} = 0$, значит, его столбцы $\tilde{M}_1, \ldots, \tilde{M}_r, \tilde{M}_{r+1}$ линейно зависимы (предположение о том, что $\operatorname{rk}\{\tilde{M}_1, \ldots, \tilde{M}_r, \tilde{M}_{r+1}\} = r + 1$ ввиду следствия 4.9 означало бы, что $\tilde{M} \neq 0$). В то же время первые r столбцов минора $\tilde{M}$ линейно независимы, так как $r \geq \operatorname{rk}\{\tilde{M}_1, \ldots, \tilde{M}_r\} \geq \operatorname{rk} B = r$. Из курса линейной алгебры известно, что если добавление вектора к линейно независимой системе превращает ее в линейно зависимую, то добавленный вектор есть линейная комбинация исходных, так что столбец $\tilde{M}_{r+1} = (a_{1j}, \ldots, a_{rj}, a_{ij})^t$ линейно выражается через первые r столбцов минора $\tilde{M}$. В силу произвольности номера j заключаем, что через линейно независимые столбцы $\tilde{M}_1 = (a_{11}, \ldots, a_{r1}, a_{i1})^t, \ldots, \tilde{M}_r = (a_{1r}, \ldots, a_{rr}, a_{ir})^t$ линейно выражаются все столбцы матрицы $\tilde{A} = \begin{pmatrix} a_{11} & \ldots & a_{1r} & \ldots & a_{1n} \\ & \ldots & & \ldots & \\ a_{r1} & \ldots & a_{rr} & \ldots & a_{rn} \\ a_{i1} & \ldots & a_{ir} & \ldots & a_{in} \end{pmatrix}$, следовательно, ее ранг равен r. Но $\tilde{A}$ составлена из строк $A_{(1)}, \ldots, A_{(r)}, A_{(i)}$ матрицы A, поэтому $\operatorname{rk}\{A_{(1)}, \ldots, A_{(r)}, A_{(i)}\} = \operatorname{rk} \tilde{A} = r$, откуда, в частности, вытекает линейная зависимость последней системы строк. Заметим, что полученное ранее равенство $\operatorname{rk}(B|C) = r$ означает, что строки

$A_{(1)}, \ldots, A_{(r)}$ линейно независимы, следовательно, строка $A_{(i)}$ является их линейной комбинацией. Воспользовавшись произвольностью номера i, приходим к выводу, что все строки матрицы A линейно выражаются через первые ее r линейно независимых строк, откуда $\operatorname{rk} A = r$.$\lhd$

Следствие 4.10 *Ранг матрицы равен наибольшему из порядков ее ненулевых миноров.*

Доказательство. Пусть $\operatorname{rk} A = r$ и k — наибольший из порядков ненулевых миноров матрицы A. Если минор $M \neq 0$ порядка k составлен из элементов строк $A_{(i_1)}, \ldots, A_{(i_k)}$, то эти строки линейно независимы, откуда $r \geq k$. С другой стороны, применение метода окаймляющих миноров дает ненулевой минор порядка r, следовательно, $k \geq r$.$\lhd$

§5. Системы линейных уравнений

Рассмотрим систему линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \quad \cdots \quad \cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases} . \tag{5.1}$$

Легко видеть, что систему (5.1) можно записать в векторном виде

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \tag{5.2}$$

а также в матричном

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{5.3}$$

или сокращенно $Ax = b$, где $A = (a_{ij})$ — матрица системы, x — столбец неизвестных, b — столбец свободных членов. Приписав справа к матрице A столбец b, получаем *расширенную* матрицу $\bar{A} = (A|b)$ системы.

5.1 Классификация систем. Критерий совместности

Системы линейных уравнений классифицируются по нескольким признакам. Система, имеющая хотя бы одно решение, называется *совместной*, если же решений нет — *несовместной*. Если совместная система имеет ровно одно решение, то это *определенная* система, в противном случае — *неопределенная*. Критерий совместности системы дает

Теорема Кронекера–Капелли. *Система $Ax = b$ совместна тогда и только тогда, когда ранги основной и расширенной матриц системы совпадают, то есть,* $\operatorname{rk} A = \operatorname{rk} \bar{A}$.

Доказательство. ($\Rightarrow$) : Пусть система $Ax = b$ совместна, то есть, существует решение x_0. Записав равенство $Ax_0 = b$ в векторном виде (5.2), получаем, что столбец b является линейной комбинацией столбцов $A^{(1)}, \ldots, A^{(n)}$ матрицы A, следовательно, $\operatorname{rk} A = \operatorname{rk}\{A^{(1)}, \ldots, A^{(n)}\} = \operatorname{rk}\{A^{(1)}, \ldots, A^{(n)}, b\} = \operatorname{rk} \bar{A}$.

($\Leftarrow$): Фиксируем базис $\{A^{(i_1)}, \ldots, A^{(i_r)}\}$ системы столбцов матрицы A. Он одновременно является и базисом системы столбцов расширенной матрицы, поскольку $\operatorname{rk} A = \operatorname{rk} \bar{A}$. Следовательно, столбец b линейно выражается через столбцы $A^{(i_1)}, \ldots, A^{(i_r)}$ и тем более выражается через все столбцы матрицы A. С учетом (5.2) это означает совместность системы. $\lhd$

Докажем полезную лемму, которая не раз понадобится в дальнейшем.

Лемма 5.1 *Элементарные преобразования строк расширенной матрицы $\bar{A}$ не меняют множества решений системы $Ax = b$.*

Доказательство. Как известно (см. леммы 2.2 и 2.3), элементарные преобразования строк матрицы $\bar{A}$ равносильны ее домножению слева на некоторую обратимую матрицу F, поэтому после элементарных преобразований система $Ax = b$ примет вид $FAx = Fb$. Ясно, что если x_0 — некоторое решение системы $Ax = b$, то $Ax_0 = b$, откуда $FAx_0 = Fb$, то есть, x_0 — решение преобразованной системы. Обратно, если $FAx_0 = Fb$, то ввиду обратимости матрицы F получаем $Ax_0 = F^{-1}FAx_0 = F^{-1}Fb = b$, то есть, x_0 — решение исходной системы

$Ax = b$. Следовательно, множества решений исходной и преобразованной систем совпадают.$\lhd$

5.2 Однородные системы

Система $Ax = b$ называется *однородной*, если ее столбец свободных членов — нулевой, то есть, $b = 0$. В противном случае система называется *неоднородной*. Однородная система $Ax = 0$ имеет тривиальное решение $x = 0$, поэтому она всегда совместна.

Пусть A — $m \times n$-матрица. Легко проверяется

Лемма 5.2 *Множество V_A решений системы $Ax = 0$ является подпространством пространства V n-мерных столбцов.*$\lhd$

Теорема 5.3 Пусть $s = \dim V_A$, $r = \operatorname{rk} A$. Тогда $r + s = n$.

Доказательство. Выберем в V_A базис $\{X_1, \ldots, X_s\}$ и дополним его до базиса $\{X_1, \ldots, X_s, \ldots, X_n\}$ пространства V. Любой столбец $X \in V$ линейно выражается через векторы базиса:

$$X = \lambda_1 X_1 + \cdots + \lambda_s X_s + \cdots + \lambda_n X_n,$$

откуда

$$AX = \lambda_1 AX_1 + \cdots + \lambda_s AX_s + \cdots + \lambda_n AX_n = \lambda_{s+1} AX_{s+1} + \cdots + \lambda_n AX_n,$$

поскольку $AX_1 = \cdots = AX_s = 0$. Таким образом, любой столбец вида AX есть линейная комбинация столбцов $\tilde{X}_1 = AX_{s+1}, \ldots, \tilde{X}_{n-s} = AX_n$.

Покажем, что система $\{\tilde{X}_1, \ldots, \tilde{X}_{n-s}\}$ линейно независима. В самом деле,

$$0 = \lambda_1 \tilde{X}_1 + \cdots + \lambda_{n-s} \tilde{X}_{n-s} = \lambda_1 AX_{s+1} + \cdots + \lambda_{n-s} AX_n = A(\lambda_1 X_{s+1} + \cdots + \lambda_{n-s} X_n)$$

влечет $\lambda_1 X_{s+1} + \cdots + \lambda_{n-s} X_n \in V_A$, следовательно,

$$\lambda_1 X_{s+1} + \cdots + \lambda_{n-s} X_n = \mu_1 X_1 + \cdots + \mu_s X_s$$

для некоторых $\mu_1, \ldots, \mu_s$, откуда

$$\mu_1 X_1 + \cdots + \mu_s X_s - \lambda_1 X_{s+1} - \cdots - \lambda_{n-s} X_n = 0.$$

Но $\{X_1, \ldots, X_s, \ldots, X_n\}$ — базис V, поэтому $\mu_1 = \cdots = \mu_s = \lambda_1 = \cdots = \lambda_{n-s} = 0$.

Заметим, что если $X = (0 \dots \overset{\overset{i}{\downarrow}}{1} \dots 0)^t$, то $AX = A^{(i)}$, поэтому любой столбец матрицы A (как столбец вида AX) линейно выражается через столбцы $\tilde{X}_1, \dots, \tilde{X}_{n-s}$, следовательно, с учетом леммы 2.7 имеем $\operatorname{rk} A \le \operatorname{rk}\{\tilde{X}_1, \dots, \tilde{X}_{n-s}\} = n - s$. С другой стороны, непосредственно из определения операции умножения матриц следует, что любой столбец вида AX является линейной комбинацией столбцов матрицы A, так что в силу той же леммы 2.7 получаем $\operatorname{rk}\{\tilde{X}_1, \dots, \tilde{X}_{n-s}\} \le \operatorname{rk}\{A^{(1)}, \dots, A^{(n)}\} = \operatorname{rk} A$. В итоге, $r = \operatorname{rk} A = n - s$, откуда $r + s = n$.$\lhd$

Следствие 5.4 *Однородная система имеет только нулевое решение тогда и только тогда, когда ранг ее матрицы совпадает с количеством неизвестных.*

Доказательство. Ввиду теоремы 5.3 имеем $V_A = \{0\} \Leftrightarrow n - r = \dim V_A = 0 \Leftrightarrow n = r$.$\lhd$

Итак, задача о нахождении общего решения однородной системы сводится к поиску базиса подпространства V_A (он называется *фундаментальным набором решений* (ФНР) системы): если $\{X_1, \dots, X_{n-r}\}$ — ФНР, то

$$X^{\text{одн}}_{\text{общ}} = C_1 X_1 + \dots + C_{n-r} X_{n-r}, \tag{5.4}$$

где $C_1, \dots, C_{n-r}$ — произвольные числовые коэффициенты.

Существуют различные способы нахождения ФНР. Опишем один из них.

Рассмотрим однородную систему $Ax = 0$. Ввиду лемм 2.1 и 5.1 можно считать, что матрица A имеет ступенчатый вид. Более того, перенумеровав при необходимости неизвестные (с учетом (5.2) это равносильно перестановке столбцов матрицы A), матрицу можно привести к виду

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2r} & \dots & a_{2n} \\ & \dots & & \dots & & \dots \\ 0 & 0 & \dots & a_{rr} & \dots & a_{rn} \\ 0 & 0 & \dots & 0 & \dots & 0 \\ & \dots & & \dots & & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{pmatrix},$$

где $a_{11}\dots a_{rr} \neq 0$. Объявим неизвестные $x_1,\dots,\ x_r$ главными, а прочие — свободными. Перенесем свободные неизвестные в правые части уравнений. Получим

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r &= -a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n \\ a_{22}x_2 + \dots + a_{2r}x_r &= -a_{2,r+1}x_{r+1} - \dots - a_{2n}x_n \\ \dots \qquad\qquad \dots &\qquad \dots \\ a_{rr}x_r &= -a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n \end{cases} \tag{5.5}$$

Каждому набору $(c_1,\dots,c_{n-r})$ значений свободных неизвестных отвечает частное решение $X_0 = (x_1,\dots,x_r,c_1,\dots,c_{n-r})$ исходной системы (его удобно искать, двигаясь по системе (5.5) снизу вверх — из последнего уравнения находим значение неизвестной x_r, подставляем его в предыдущее уравнение, находим значение x_{r-1} и так далее). Тогда решения $X_1,\ X_2,\dots,\ X_{n-r}$, отвечающие наборам $(1,0,\dots,0)$, $(0,1,\dots,0),\dots,\ (0,0,\dots,1)$ значений свободных неизвестных, образуют ФНР. В самом деле, векторы $X_1,\dots,X_{n-r}$ линейно независимы, так как в составленной из их координат матрице последние $n-r$ столбцов образуют единичную матрицу, определитель которой, разумеется, отличен от нуля, следовательно, с учетом следствия 4.10 имеем $n-r \geq \mathrm{rk}\{X_1,\dots,X_{n-r}\} \geq n-r$. Осталось заметить, что количество векторов $X_1,\dots,\ X_{n-r}$ совпадает с размерностью пространства решений V_A.

5.3 Неоднородные системы

Вернемся к неоднородным системам общего вида. Мы уже знаем способ проверки совместности систем с помощью теоремы Кронекера–Капелли. Получим теперь формулу общего решения неоднородной системы в случае ее совместности.

Итак, пусть $Ax = b$ — совместная система линейных уравнений. Следовательно, существует удовлетворяющий ей вектор X'. Тогда для любого частного решения X_0 однородной системы $Ax = 0$ имеем

$$A(X' + X_0) = AX' + AX_0 = b + 0 = b,$$

то есть, $X' + X_0$ — снова решение системы $Ax = b$. Обратно, пусть X'' — некоторое решение системы $Ax = b$. Тогда

$$A(X'' - X') = AX'' - AX' = b - b = 0,$$

значит, вектор $X_0 = X'' - X'$ есть решение однородной системы, а решение X'' исходной системы можно представить в виде $X'' = X' + X_0$. Тем самым получена формула общего решения для неоднородных систем

$$X^{\text{неодн}}_{\text{общ}} = X^{\text{неодн}}_{\text{част}} + X^{\text{одн}}_{\text{общ}}. \tag{5.6}$$

Следствие 5.5 *Совместная система $Ax = b$ является определенной тогда и только тогда, когда система $Ax = 0$ имеет только нулевое решение.* $\lhd$

Итак, общее решение неднородной системы есть сумма ее частного решения и общего решения соответствующей однородной системы. О решении однородных систем было рассказано выше, следовательно, осталось указать способ нахождения частного решения совместной неоднородной системы $Ax = b$.

Как и в случае однородных систем, воспользовавшись элементарными преобразованиями строк и перенумеровав при необходимости неизвестные, можно считать, что расширенная матрица системы имеет вид

$$\bar{A} = \left(\begin{array}{ccccccc|c} a_{11} & a_{12} & \dots & a_{1r} & \dots & a_{1n} & b_1 \\ 0 & a_{22} & \dots & a_{2r} & \dots & a_{2n} & b_2 \\ & \dots & & \dots & & \dots & \\ 0 & 0 & \dots & a_{rr} & \dots & a_{rn} & b_r \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \\ & \dots & & \dots & & \dots & \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 \end{array}\right), \tag{5.7}$$

где $a_{11} \dots a_{rr} \neq 0$, неизвестные $x_1, \dots, x_r$ — главные, $x_{r+1}, \dots, x_n$ — свободные. Перенеся в правые части уравнений свободные неизвестные, придав им произвольные значения (как правило, для простоты вычислений значения свободных неизвестных полагают равными 0) и решив получившуюся определенную систему из r уравнений, получим требуемое частное решение исходной системы.

Суммируя приведенные выше результаты, получаем следующий АЛГОРИТМ РЕШЕНИЯ НЕОДНОРОДНЫХ СИСТЕМ:

Шаг 1. С помощью теоремы Кронекера–Капелли проверяем совместность системы. Если система несовместна, то решений нет, в случае совместности переходим к шагу 2.

Шаг 2. Находим ФНР соответствующей однородной системы и по формуле (5.4) получаем ее общее решение $X^{\text{одн}}_{\text{общ}}$.

Шаг 3. Находим частное решение $X^{\text{неодн}}_{\text{част}}$ неоднородной системы и согласно (5.6) получаем ее общее решение $X^{\text{неодн}}_{\text{общ}}$.

Следующая таблица отражает зависимость числа решений системы m уравнений с n неизвестными от ее типа.

Система	Неоднородная		Однородная	
	Общая	$m < n$	Общая	$m < n$
Число решений	0,1,∞	0,∞	1,∞	∞

Таб. 2

В заключение опишем еще один способ решения неоднородных систем, известный как метод Гаусса.

Приведем расширенную матрицу $\bar{A}$ системы $Ax = b$ к ступенчатому виду. Если полученная матрица содержит хотя бы одну строку вида $(0 \dots 0 \,|\, b_i)$, где $b_i \neq 0$, то система несовместна, в противном случае, перенумеровав при необходимости неизвестные, можно считать, что расширенная матрица имеет вид (5.7). Объявим первые r неизвестных главными, а прочие — свободными, и перенесем свободные неизвестные в правые части уравнений. Получим систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r &= b_1 - a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n \\ a_{22}x_2 + \dots + a_{2r}x_r &= b_2 - a_{2,r+1}x_{r+1} - \dots - a_{2n}x_n \\ \dots \quad\quad \dots &\quad\quad \dots \\ a_{rr}x_r &= b_r - a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n \end{cases} \tag{5.8}$$

Двигаясь снизу вверх по системе (5.8), решаем ее, обращаясь с правыми частями уравнений как с буквенными выражениями. В результате главные неизвестные будут выражены через свободные: $x_1 = f_1(x_{r+1}, \dots, x_n), \dots, x_r = f_r(x_{r+1}, \dots, x_n)$. Тогда общее решение исходной системы $Ax = b$ имеет вид

$$X^{\text{неодн}}_{\text{общ}} = (f_1(c_1, \dots, c_{n-r}), \dots, f_r(c_1, \dots, c_{n-r}), c_1, \dots, c_{n-r}),$$

где $c_1, \dots, c_{n-r}$ — произвольные числа.

§6. Многочлены

В предыдущих параграфах были рассмотрены примеры различных алгебраических систем, такие как поле комплексных чисел, кольцо матриц, группа перестановок. Данный параграф посвящен еще одному важному классу алгебраических объектов — колец многочленов.

6.1 Построение кольца многочленов. Степень многочлена

Пусть K — поле. Рассмотрим множество P, состоящее их всех последовательностей элементов поля K, в которых все члены, начиная с некоторого номера, равны 0, то есть, $P = \{(f_0, f_1, f_2, \dots) : \exists n \, \forall k > n \quad f_k = 0\}$. Зададим на P операции сложения и умножения:

$$(f_0, f_1, f_2, \dots) + (g_0, g_1, g_2, \dots) = (f_0 + g_0, f_1 + g_1, f_2 + g_2, \dots),$$

$$(f_0, f_1, f_2, \dots)(g_0, g_1, g_2, \dots) = (h_0, h_1, h_2, \dots),$$

где $h_k = \sum\limits_{i+j=k} f_i g_j$ при всех k. Например, $h_0 = f_0 g_0$, $h_1 = f_0 g_1 + f_1 g_0$, $h_2 = f_0 g_2 + f_1 g_1 + f_2 g_0$ и так далее.

Теорема 6.1 $(P, +, \cdot)$ — *коммутативное ассоциативное кольцо с единицей.*

Доказательство. Пусть $f, g \in P$. Тогда $f_k = 0$ при $k > n$ для некоторого n и $g_k = 0$ при $k > m$ для некоторого m. Ясно, что $(f + g)_k = 0$ при $k > \max(n, m)$, значит, $f + g \in P$. Если $i + j > n + m$, то $i > n$ или $j > m$, следовательно, $f_i g_j = 0$, поскольку по крайней мере один из элементов f_i и g_j равен нулю. Поэтому $(fg)_k = \sum\limits_{i+j=k} f_i g_j = 0$ при $k > n + m$, значит, $fg \in P$. Таким образом, P замкнуто относительно введенных на нем операций сложения и умножения.

Перейдем к проверке аксиом кольца. Поскольку сложение последовательностей из P производится поэлементно, оно наследует коммутативность и ассоциативность сложения в K, нулем в P будет последовательность $(0, 0, 0, \dots)$, противоположным к $(f_0, f_1, f_2, \dots)$ элементом — последовательность $(-f_0, -f_1, -f_2, \dots)$. Таким образом, $(P, +)$ — абелева группа. Коммутативность умножения в P вытекает из коммутативности умножения в K:

$$(fg)_k = \sum_{i+j=k} f_i g_j = \sum_{j+i=k} g_j f_i = (gf)_k, \quad k = 0, 1, 2, \dots$$

Проверим дистрибутивность:

$$[(f+g)h]_k = \sum_{i+j=k} (f+g)_i h_j = \sum_{i+j=k} (f_i+g_i)h_j = \sum_{i+j=k} f_i h_j + \sum_{i+j=k} g_i h_j =$$

$$(fh)_k + (gh)_k = (fh+gh)_k, \quad k = 0, 1, 2, \dots$$

следовательно, $(f+g)h = fh+gh$. Воспользовавшись коммутативностью умножения, получаем и второй закон дистрибутивности: $f(g+h) = (g+h)f = gf + hf = fg + fh$. Теперь убедимся в ассоциативности уможения:

$$[(fg)h]_k = \sum_{i+j=k} (fg)_i h_j = \sum_{i+j=k} \Big(\sum_{s+t=i} f_s g_t \Big) h_j = \sum_{s+t+j=k} f_s g_t h_j.$$

Аналогичные выкладки дают то же выражение и для $[f(gh)]_k$. Наконец, непосредственная проверка показывает, что единицей в P является последовательность $(1, 0, 0, \dots)$.$\lhd$

Заметим, что с последовательностями вида $(a, 0, 0, \dots)$ операции сложения и умножения выполняются так же, как и с элементами поля K: $(a, 0, 0, \dots) + (b, 0, 0, \dots) = (a+b, 0, 0, \dots)$, $(a, 0, 0, \dots)(b, 0, 0, \dots) = (ab, 0, 0, \dots)$, кроме того, $(a, 0, 0, \dots)(f_0, f_1, f_2, \dots) = (af_0, af_1, af_2, \dots)$ и $(f_0, f_1, f_2, \dots)(a, 0, 0, \dots) = (af_0, af_1, af_2, \dots)$, поэтому такие элементы кольца P удобно отождествлять с элементами из K и вместо $(a, 0, 0, \dots)$ писать просто a. В частности, единица $(1, 0, 0, \dots)$ кольца P при этом приобретает привычную форму записи — 1.

Обозначим элемент $(0, 1, 0, 0, 0, \dots) \in P$ через X. Имеем:

$$\begin{aligned} X &= (0, 1, 0, 0, 0, \dots) \\ X^2 &= (0, 0, 1, 0, 0, \dots) \\ X^3 &= (0, 0, 0, 1, 0, \dots) \end{aligned}$$

и так далее. Удобно также считать, что $X^0 = 1$. Тогда

$$(f_0, f_1, \dots, f_n, 0, 0, \dots) =$$

$$(f_0, 0, \dots, 0, 0, 0, \dots) + (0, f_1, \dots, 0, 0, 0, \dots) + \dots + (0, 0, \dots, f_n, 0, 0, \dots) =$$

$$f_0 + f_1 X + \dots + f_n X^n = \sum_k f_k X^k.$$

Элемент X называется *переменной* или *неизвестной*, выражение $f = f_0 + f_1 X + \dots + f_n X^n$ — *многочленом* (*полиномом*) от X,

элементы $f_0, f_1, \ldots, f_n \in K$ — *коэффициентами* многочлена f, а кольцо P обозначается через $K[X]$ и называется *кольцом многочленов* от переменной X над полем K.

Легко видеть, что введенные выше операции сложения и умножения многочленов при переходе к новой форме записи соответствуют обычным правилам сложения и умножения выражений, содержащих переменную X. В последних порядок нумерации коэффициентов не имеет существенного значения. Договоримся в дальнейшем нумеровать коэффициенты многочленов в порядке убывания степеней переменной X: $f = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n$ $(a_0 \neq 0)$. Наибольшая из степеней переменной X, коэффициент при которой отличен от 0, называется *степенью* многочлена f и обозначается $\deg f$, коэффициент a_0 при X^n, где $n = \deg f$, называется *старшим* коэффициентом, а коэффициент a_n — *свободным членом*. Степень нулевого многочлена обозначается через $-\infty$. Множество степеней многочленов естественным образом упорядочено: $-\infty < 0 < 1 < 2 < \ldots$, при этом считается, что $-\infty + \deg f = -\infty$ для любого $f \in K[X]$. Вполне очевидны следующие

Свойства степеней:

1°. $\deg(f \pm g) \leq \max(\deg f, \deg g)$.

2°. $\deg(fg) = \deg f + \deg g$.

Говорят, что кольцо не содержит *делителей нуля*, если для любых его элементов a, b из $ab = 0$ следует $a = 0$ или $b = 0$. Коммутативное кольцо без делителей нуля называется *целостным*. Легко видеть, что любое поле целостно.

Предложение 6.2 *Кольцо $K[X]$ — целостно.*

Доказательство. Пусть $f, g \in K[X]$ и $fg = 0$. Если $f \neq 0$, то $\deg f \geq 0$, поэтому с учетом 2° имеем $-\infty = \deg 0 = \deg(fg) = \deg f + \deg g \geq \deg g \geq -\infty$, откуда $\deg g = -\infty$ и, следовательно, $g = 0$.$\lhd$

6.2 Деление многочленов с остатком

Пусть K — поле, $f, g \in K[X]$. Если $f = qg + r$, где $q, r \in K[X]$, $\deg r < \deg g$, то говорят, что f *делится на g с остатком r*.

Теорема (о делении с остатком). *Пусть $f, g \in K[X]$, $g \neq 0$. Тогда*

существует единственная пара многочленов $q, r \in K[X]$ *такая, что*

$$f = qg + r, \quad \deg r < \deg g.$$

Доказательство. Сначала докажем существование пары (q, r), воспользовавшись методом математической индукции по степени n многочлена f.

Основание индукции. Согласно условию теоремы $g \neq 0$, так что $\deg g > -\infty$, поэтому $f = 0 \cdot g + f$ — требуемое представление f при любом $n < \deg g$.

Пусть для всех многочленов степени, строго меньшей n, утверждение уже доказано. Обозначим степень g через m. С учетом сказанного выше, можно считать, что $m \leq n$. Запишем многочлены f и g в явном виде:

$$f = a_0 X^n + a_1 X^{n-1} + \cdots + a_{n-1} X + a_n,$$
$$g = b_0 X^m + b_1 X^{m-1} + \cdots + b_{m-1} X + b_m.$$

Рассмотрим многочлен $\tilde{g} = a_0 b_0^{-1} X^{n-m} g$. Очевидно, что $\deg \tilde{g} = n$ и что старшие члены многочленов f и $\tilde{g}$ совпадают, поэтому степень многочлена $\tilde{f} = f - \tilde{g}$ строго меньше n. По предположению индукции найдутся многочлены $\tilde{q}, \tilde{r}$ такие, что $\tilde{f} = \tilde{q} g + \tilde{r}$, $\deg \tilde{r} < \deg g$. Тогда

$$f = \tilde{g} + \tilde{f} = a_0 b_0^{-1} X^{n-m} g + \tilde{q} g + \tilde{r} = (a_0 b_0^{-1} X^{n-m} + \tilde{q}) g + \tilde{r}$$

— требуемое представление f.

Докажем единственность. Пусть $f = q_i g + r_i$, $\deg r_i < \deg g$ $(i = 1, 2)$ — два представления многочлена f. Тогда $(q_1 - q_2) g = r_2 - r_1$. Если $q_1 - q_2 \neq 0$, то $\deg(q_1 - q_2) \geq 0$, следовательно,

$$\deg g \leq \deg(q_1 - q_2) + \deg g = \deg((q_1 - q_2) g) = \deg(r_2 - r_1) \leq$$
$$\max(\deg r_2, \deg r_1) < \deg g,$$

откуда $\deg g < \deg g$ — противоречие. Значит, $q_1 - q_2 = 0$, то есть, $q_1 = q_2$. Тогда $q_1 g = q_2 g$, поэтому из равенств $q_1 g + r_1 = f = q_2 g + r_2$ выводим $r_1 = r_2$. $\lhd$

6.3 Делимость в кольце многочленов

Пусть R — целостное кольцо с единицей, $a, b \in R$. Говорят, что a *делит* b (обозначение: $a \mid b$), если $b = ac$ для некоторого $c \in R$.

Выражение $a \nmid b$ означает, что a не делит b. Делители единицы называются *обратимыми* элементами.

Если многочлен $f \in K[X]$ обратим, то $fg = 1$ для некоторого $g \in K[X]$. Тогда $0 = \deg(fg) = \deg f + \deg g$, откуда $\deg f = \deg g = 0$, то есть, $f, g \in K$. Следовательно, обратимыми элементами кольца $K[X]$ являются только ненулевые элементы поля K.

Если $a \mid b$ и $b \mid a$, то элементы $a, b \in R$ называются *ассоциированными*. Покажем, что в этом случае $b = ua$, где u — обратимый элемент. В самом деле, если $a = 0$, то $b = ac = 0c = 0$, так что $a = 1 \cdot b$, а если $a \neq 0$, то $a = bd = acd$, откуда $a(1 - cd) = 0$. Тогда в силу целостности R получаем $1 - cd = 0$, так что $cd = 1$ и, следовательно, c, d — обратимые элементы.

Легко проверяются следующие

Свойства делимости:

1. $a \mid b,\ b \mid c \Rightarrow a \mid c$.
2. $c \mid a,\ c \mid b \Rightarrow c \mid (a \pm b)$.
3. $a \mid b \Rightarrow a \mid bc$.
4. $a \mid b_1, \ldots, a \mid b_k \Rightarrow a \mid (b_1c_1 + \cdots + b_kc_k)$ при любых $c_1, \ldots, c_k \in R$.

Элемент $p \in R$, $p \nmid 1$, называется *простым*, если его нельзя представить в виде $p = ab$, где $a \nmid 1$, $b \nmid 1$. Простые элементы кольца многочленов обычно называют *неприводимыми* многочленами. Многочлен степени 1 всегда неприводим. В самом деле, если $p = ab$, то $1 = \deg p = \deg a + \deg b$, откуда либо $\deg a = 0$ и, следовательно, $a \mid 1$, либо $\deg b = 0$ и $b \mid 1$.

6.4 НОД и НОК в кольце многочленов

Пусть R — целостное кольцо. Говорят, что $d \in R$ является *наибольшим общим делителем* элементов $a, b \in R$ (обозначение: $d = \text{НОД}(a, b)$), если

1) $d \mid a,\ d \mid b$;

2) $c \mid a,\ c \mid b \Rightarrow c \mid d$.

Ясно, что НОД определяется с точностью до обратимого множителя. Справедливы следующие свойства:

1. $\text{НОД}(a, b) = a \Leftrightarrow a \mid b$.
2. $\text{НОД}(a, 0) = a$.

3. $\text{НОД}\,(ta, tb) = t\,\text{НОД}\,(a, b)$.

4. $\text{НОД}\,(\text{НОД}\,(a, b), c) = \text{НОД}\,(a, \text{НОД}\,(b, c))$.

Докажем свойство 3 (проверка остальных свойств предлагается в качестве упражнения). Введем обозначения: $\text{НОД}\,(a, b) = d$, $\text{НОД}\,(ta, tb) = f$. Так как $d \mid a$, $d \mid b$, то $td \mid ta$, $td \mid tb$. Следовательно, $td \mid f$, тем самым, $f = tdg$ для некоторого g. Тогда $tdg = f \mid ta$, что с учетом целостности влечет $dg \mid a$, аналогично, $tdg = f \mid tb$ влечет $dg \mid b$. Значит, $dg \mid \text{НОД}\,(a, b) = d$, откуда $g \mid 1$. Получаем, что f и td — ассоциированные элементы. $\lhd$

Элементы $a, b \in R$, для которых $\text{НОД}\,(a, b) = 1$, называются *взаимно простыми*.

Двойственным образом к НОД определяется понятие *наименьшего общего кратного* (НОК): $m = \text{НОК}\,(a, b)$, если

1') $a \mid m$, $b \mid m$;

2') $a \mid c$, $b \mid c \Rightarrow m \mid c$.

Предложение 6.3 *Пусть R — целостное кольцо, $a, b \in R$. Если $d = \text{НОД}\,(a, b)$, $m = \text{НОК}\,(a, b)$, то $ab = md$.*

Доказательство. Если $a = 0$ или $b = 0$, то $m = 0$ и равенство $ab = md$ тривиально верно, поэтому можно считать, что $a \neq 0$, $b \neq 0$ и, следовательно, $d \neq 0$.

Согласно условию верны равенства $a = a'd$, $b = b'd$, откуда $ab = (a'b'd)d$. Положим $m = a'b'd$ и докажем, что $m = \text{НОК}\,(a, b)$. Имеем: $a \mid ab' = a'b'd = m$ и, аналогично, $b \mid a'b = a'b'd = m$, тем самым выполнено условие 1') определения НОК.

Пусть теперь $a \mid c$ и $b \mid c$. Тогда с учетом свойства 3 НОД из $ab \mid ca$, $ab \mid cb$ выводим $md = ab \mid \text{НОД}\,(ca, cb) = c\,\text{НОД}\,(a, b) = cd$. Таким образом, $md \mid cd$, следовательно, $m \mid c$, то есть, для m выполнено и условие 2'). $\lhd$

Согласно доказанному предложению для вычисления НОК элементов достаточно найти их НОД и поделить на него произведение исходных элементов. В кольце $K[X]$ для вычисления $\text{НОД}\,(a, b)$ обычно используют метод, называемый АЛГОРИТМОМ ЕВКЛИДА. В силу свойства 2 НОД можно ограничиться случаем, когда a и b отличны от нуля.

Введем обозначения $a = r_{-1}$, $b = r_0$ и положим $k = 0$.

Шаг k: Поделим r_{k-1} с остатком на r_k. Получим

$$r_{k-1} = q_k r_k + r_{k+1}, \quad \deg r_{k+1} < \deg r_k.$$

Если $r_{k+1} = 0$, то алгоритм завершен и НОД $(a, b) = r_k$, в противном случае переходим к шагу $k + 1$.

Обоснование алгоритма. Поскольку последовательность степеней остатков r_0, r_1,... является строго убывающей, через некоторое число m шагов ($m \leq \deg r_0 + 1$) получим $\deg r_{m+1} = -\infty$, то есть, $r_{m+1} = 0$ и алгоритм закончит свою работу. Покажем, что $r_m = \text{НОД}\,(a, b)$.

Имеем систему равенств:

$$\begin{aligned} a &= q_0 b + r_1, \\ b &= q_1 r_1 + r_2, \\ \dots &\quad \dots \\ r_{m-2} &= q_{m-1} r_{m-1} + r_m, \\ r_{m-1} &= q_m r_m. \end{aligned} \tag{6.1}$$

Очевидно, $r_m \mid r_m$. Согласно последнему равенству системы (6.1) $r_m \mid r_{m-1}$. Тогда из предпоследнего равенства вытекает $r_m \mid r_{m-2}$. Воспользовавшись третьим с конца равенством, выводим $r_m \mid r_{m-3}$ и так далее. В итоге, получаем $r_m \mid b$ и $r_m \mid a$, тем самым, для r_m выполено условие 1) определения НОД.

Пусть теперь $c \mid a$ и $c \mid b$. Двигаясь по системе (6.1) сверху вниз, последовательно получаем $c \mid r_1$, $c \mid r_2$,..., $c \mid r_m$, то есть, верно и условие 2). Следовательно, $r_m = \text{НОД}\,(a, b)$. $\lhd$

Перепишем равенства системы (6.1) (кроме последнего) в виде

$$\begin{aligned} a - q_0 b &= r_1, \\ b - q_1 r_1 &= r_2, \\ \dots &\quad \dots \\ r_{m-2} - q_{m-1} r_{m-1} &= r_m. \end{aligned} \tag{6.2}$$

Двигаясь по системе (6.2) сверху вниз, замечаем, что r_1 выражается через многочлены a и b, многочлен r_2 выражается через b, r_1 и поэтому опять выражается через a и b, и так далее. В итоге, получим выражение многочлена r_m через a и b. Таким образом, опираясь дополнительно на предложение 6.3 и алгоритм Евклида, получаем доказательство следующей теоремы.

Теорема 6.4 *В кольце* $K[X]$ *любые многочлены* a *и* b *имеют* НОД *и* НОК, *причем существуют многочлены* $u, v \in K[X]$ *такие, что*

$$\text{НОД}\,(a, b) = ua + vb.$$

В частности, a *и* b *взаимно просты* $\Leftrightarrow$ $ua + vb = 1$ *для некоторых* $u, v \in K[X]\,.\lhd$

Следствие 6.5 *Для всех* $a, b, c \in K[X]$ *справедливы импликации:*

1. $\text{НОД}\,(a, b) = 1,\ \text{НОД}\,(a, c) = 1 \ \Rightarrow\ \text{НОД}\,(a, bc) = 1\,.$
2. $a \mid bc,\ \text{НОД}\,(a, b) = 1 \ \Rightarrow\ a \mid c\,.$
3. $b \mid a,\ c \mid a,\ \text{НОД}\,(b, c) = 1 \ \Rightarrow\ bc \mid a\,.$

Доказательство. 1. Согласно теореме 6.4 имеем: $u_1a + v_1b = 1$, $u_2a + v_2c = 1$. Тогда $1 = (u_1a+v_1b)(u_2a+v_2c) = (u_1u_2a+u_1v_2c+v_1u_2b)a+(v_1v_2)bc$, следовательно, $\text{НОД}\,(a, bc) = 1\,.$

2. Если $a \mid bc$, $\text{НОД}\,(a, b) = 1$, то $bc = aw$, $ua + vb = 1$ для некоторых $u, v, w \in K[X]$. Тогда $c = c(ua+vb) = acu+bcv = acu+awv = a(cu + wv)$, откуда $a \mid c$.

3. Так как $b \mid a$ и $c \mid a$, то $\text{НОК}\,(b, c) \mid a$. Поэтому с учетом равенства $\text{НОД}\,(b, c) = 1$ и предложения 6.3 получаем $\text{НОК}\,(b, c) = \text{НОК}\,(b, c)\text{НОД}\,(b, c) = bc$, следовательно, $bc \mid a\,.\lhd$

6.5 Факториальность кольца $K[X]$

Пусть R — целостное кольцо. Говорят, что R — *факториально* или является кольцом *с однозначным разложением на простые множители*, если для любого $a \in R$ $(a \neq 0)$, существуют обратимый элемент $u \in R$ и (не обязательно различные) простые элементы $p_1, \ldots, p_r \in R$ такие, что

$$a = up_1 \ldots p_r, \qquad (*)$$

причем если $a = wq_1 \ldots q_s$ — еще одно разложение вида $(*)$, то $s = r$ и при подходящей нумерации q_i ассоциирован с p_i для всех i.

Замечание. Допуская случай $r = 0$, можно считать, что обратимые элементы тоже обладают представлением вида $(*)$.

Теорема 6.6 *Кольцо* R *с разложением на простые множители факториально тогда и только тогда, когда для всех* $a, b \in R$ *и для любого простого* $p \in R$ *из того, что* $p \mid ab$ *вытекает* $p \mid a$ *или* $p \mid b$.

Замечание. Из присутствующего в формулировке теоремы 6.6 условия о делимости простым элементом одного из двух сомножителей вытекает справедливость аналогичного условия для любого конечного числа сомножителей: *если* $p \in R$ *— прост и* $p \mid \prod\limits_{i=1}^{k} a_k$*, то* $p \mid a_i$ *для некоторого* i. Доказательство является несложным упражнением на применение метода математической индукции.

Доказательство теоремы. ($\Rightarrow$) : Пусть R факториально и простой элемент $p \in R$ делит ab, то есть, $ab = pc$ для некоторого $c \in R$. Разложим элементы a, b, c на простые множители: $a = u\prod\limits_{i} a_i$, $b = v\prod\limits_{j} b_i$, $c = w\prod\limits_{k} c_k$. Тогда $(uv)\prod\limits_{i} a_i \prod\limits_{j} b_j = ab = wp\prod\limits_{k} c_k$, откуда ввиду факториальности R вытекает ассоциированность p с одним из простых множителей a_i или b_j. Следовательно, $p \mid a$ или $p \mid b$.

($\Leftarrow$) : Докажем единственность разложения вида $(*)$. Применим метод математической индукции по количеству n участвующих в разложении простых сомножителей.

При $n = 0$ утверждение очевидно.

Предположим, что для всех элементов кольца, допускающих разложение с менее, чем n простыми сомножителями, единственность разложения уже доказана. Пусть $a \neq 0$ и

$$a = u\prod_{i=1}^{n} p_i = v\prod_{j=1}^{m} q_j,$$

где все множители p_i, q_j — простые и $m \geq n$. В частности, p_n прост и $p_n \mid v\prod\limits_{j=1}^{m} q_j$. Следовательно, элемент p_n делит некоторый множитель q_j и поэтому, в силу простоты q_j, ассоциирован с ним, то есть, $q_j = wp_n$ для некоторого обратимого w. Перенумеровав при необходимости сомножители второго разложения, можно считать, что $j = m$. Сократив на p_n, получаем

$$u\prod_{i=1}^{n-1} p_i = (vw)\prod_{j=1}^{m-1} q_j.$$

Количество простых сомножителей в левой части последнего равенства меньше n, следовательно, по предположению индукции $n - 1 = m - 1$, то есть, $n = m$ и, при подходящей нумерации, p_i ассоциирован с q_i для

всех i.$\lhd$

Лемма 6.7 *Каждый ненулевой многочлен кольца $K[X]$ обладает разложением на простые множители.*

Доказательство. Установим существование разложения вида $(*)$ для ненулевого многочлена $a \in K[X]$, воспользовавшись методом математической индукции по степени n многочлена.

Для многочленов нулевой степени, то есть, обратимых элементов кольца $K[X]$, утверждение тривиально верно.

Пусть утверждение уже доказано для всех многочленов степени, меньшей n. Если a прост, то доказывать нечего, в противном случае $a = bc$ для некоторых необратимых $b, c \in K[X]$. Тогда $\deg a = \deg b + \deg c$, откуда $1 \leq \deg b, \deg c < \deg a = n$, следовательно, в силу предположения индукции b и c обладают разложением вида $(*)$, значит, таким разложением обладает и $bc = a$.$\lhd$

Теорема 6.8 *Кольцо $K[X]$ факториально.*

Доказательство. В силу леммы 6.7 и теоремы 6.6 достаточно показать, что для любого неприводимого многочлена $p \in K[X]$ верна импликация $p \mid ab \Rightarrow p \mid a$ или $p \mid b$. Справедливость данной импликации при $a = 0$ или $b = 0$ очевидна, поэтому можно считать, что $ab \neq 0$.

Обозначим НОД многочленов a и p через d. Так как p неприводим, то либо $d = p$, либо $d = 1$. В первом случае $p = d = \text{НОД}\,(p, a) \mid a$. Во втором случае согласно п.2 следствия 6.5 из $p \mid ab$ и $\text{НОД}\,(p, a) = d = 1$ вытекает $p \mid b$.$\lhd$

Из факториальности кольца $K[X]$ следует, что любые два ненулевых многочлена $a, b \in K[X]$ обладают "общим" разложением

$$a = up_1^{k_1} \dots p_r^{k_r}, \quad b = vp_1^{l_1} \dots p_r^{l_r},$$

где некоторые неприводимые сомножители могут присутствовать с нулевыми степенями. Нетрудно убедиться в справедливости следующих утверждений (докажите их в качестве упражнения):

1. $a \mid b \Leftrightarrow k_i \leq l_i$ для всех i.
2. $\text{НОД}\,(a, b) = p_1^{s_1} \dots p_r^{s_r}$, где $s_i = \min(k_i, l_i)$ для всех i.
3. $\text{НОК}\,(a, b) = p_1^{t_1} \dots p_r^{t_r}$, где $t_i = \max(k_i, l_i)$ для всех i.

6.6 Корни многочленов

Пусть $f \in K[X]$, $f = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n$. Многочлену f отвечает функция, действующая из K в K по правилу: $c \mapsto a_0c^n + a_1c^{n-1} + \cdots + a_{n-1}c + a_n$. Ее значение на элементе $c \in K$ обозначается через $f(c)$ и называется *значением многочлена f в точке c*. Следует подчеркнуть принципиальное различие алгебраической и функциональной точек зрения на многочлены, поскольку для некоторых полей разным многочленам может отвечать одна и та же функция. Однако, для многих полей, в том числе, для полей $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, соответствие между многочленами и отвечающими им функциями является биективным. (Подробнее см. [1], гл. 5, 6.)

Говорят, что элемент $c \in K$ является *корнем* многочлена $f \in K[X]$, если $f(c) = 0$. Элемент $c \in K$ называют также *корнем уравнения* $f(x) = 0$.

Теорема (Безу) *$c \in K$ — корень многочлена $f \in K[X]$* $\Leftrightarrow$ $(X - c) \mid f$.

Доказательство. Поделим f с остатком на $X - c$:

$$f = (X - c)q + r,$$

где $\deg r < \deg(X - c) = 1$, следовательно, $r \in K$.

($\Rightarrow$) : Пусть c — корень. Тогда $0 = f(c) = (c - c)q(c) + r = 0{\cdot}q(c) + r = r$, так что $r = 0$ и, следовательно, $(X - c) \mid f$.

($\Leftarrow$) : Если $(X - c) \mid f$, то $f = (X - c)q$ для некоторого $q \in K[X]$. Тогда $f(c) = 0{\cdot}q(c) = 0$, то есть, c — корень. $\lhd$

Замечание. $f(c) = r$, то есть, значение многочлена f на элементе c равно остатку от деления f на $X - c$.

Существует алгоритм "быстрого" деления f на $X-c$, известный как схема Горнера. Он состоит в следующем: пусть $f = a_0X^n+a_{n-1}X^{n-1}+\cdots+a_{n-1}X+a_0$, $q = b_0X^{n-1}+b_1X^{n-2}+\cdots+b_{n-2}X+b_{n-1}$ и $f = (X-c)q+r$, тогда $b_0 = a_0$, $b_k = a_k + b_{k-1}c$ при $k = 1, 2, \ldots, n-1$, $r = a_n + b_{n-1}c$. Эти вычисления удобно записывать в виде таблицы:

	a_0	a_1	$\ldots$	a_{n-1}	a_n
c	$b_0 = a_0$	$b_1 = a_1 + b_0c$	$\ldots$	$b_{n-1} = a_{n-1} + b_{n-2}c$	$r = a_n + b_{n-1}c$

Заполняется таблица так: в верхней строке (кроме первой ячейки) записываются коэффициенты многочлена f, в первой ячейке нижней

строки пишется элемент c, в следующей ячейке — a_0, а все последующие ячейки заполняются слева направо по правилу: к содержимому верхней ячейки прибавляется содержимое предыдущей (левой) ячейки, умноженное предварительно на c. Корректность данного алгоритма легко установить, подставив выражения для коэффициентов многочлена q и остатка r в равенство $f = (X - c)q + r$ (проверьте это в качестве упражнения(!)).

Элемент $c \in K$ называется *k-кратным корнем* многочлена $f \in K[X]$, если $(X - c)^k \mid f$ и $(X - c)^{k+1} \nmid f$. В частности, при $k = 1$ корень называется *простым*, при $k = 2$ — *двойным*, при $k = 3$ — *тройным* и так далее. Таким образом, $c \in K$ есть k-кратный корень многочлена f тогда и только тогда, когда $f = (X - c)^k g$, где $\text{НОД}\,(X - c, g) = 1$, то есть, $g(c) \neq 0$. Отметим, что при этом $\deg f = k + \deg g$.

Теорема 6.9 *Пусть $f \in K[X]$, $\deg f > 0$ и $c_1, \ldots, c_r$ — корни f кратностей $k_1, \ldots, k_r$. Тогда найдется $g \in K[X]$ такой, что*

$$f = (X - c_1)^{k_1} \ldots (X - c_r)^{k_r} g, \text{ где } g(c_i) \neq 0 \quad (i = 1, \ldots, r).$$

В частности, $k_1 + \cdots + k_r \leq \deg f$, то есть, сумма кратностей корней ненулевого многочлена не превосходит его степени.

Доказательство проведем индукцией по r. При $r = 1$ доказывать нечего, так как утверждение совпадает с определением k_1-кратного корня.

Предположим, что утверждение уже доказано для $r - 1$ корней, то есть, существует $h \in K[X]$ такой, что $f = (X - c_1)^{k_1} \ldots (X - c_{r-1})^{k_{r-1}} h$ и $h(c_i) \neq 0$ при $i = 1, \ldots, r - 1$. Так как $c_r - c_i \neq 0$ при $i < r$, то $(c_r - c_1)^{k_1} \ldots (c_r - c_{r-1})^{k_{r-1}} \neq 0$, но c_r — корень f, поэтому $0 = f(c_r) = (c_r - c_1)^{k_1} \ldots (c_r - c_{r-1})^{k_{r-1}} h(c_r)$ влечет $h(c_r) = 0$, то есть, c_r — корень h некоторой кратности s. Следовательно, $h = (X - c_r)^s v$, $v(c_r) \neq 0$, для подходящего $v \in K[X]$. Ясно, что $1 \leq s \leq k_r$.

Поскольку c_r является k_r-кратным корнем f, для некоторого $u \in K[X]$ верно $f = (X - c_r)^{k_r} u$. Тогда

$$(X - c_r)^{k_r} u = f = (X - c_1)^{k_1} \ldots (X - c_{r-1})^{k_{r-1}} (X - c_r)^s v.$$

Сокращая обе части последнего равенства на $(X - c_r)^s$, с учетом условия $(c_r - c_1)^{k_1} \ldots (c_r - c_{r-1})^{k_{r-1}} v(c_r) \neq 0$ приходим к $s = k_r$. Итак,

$$f = (X - c_1)^{k_1} \ldots (X - c_r)^{k_r} v,$$

где $v(c_r) \neq 0$ по выбору v и $v(c_i) \neq 0$ при $i = 1, \ldots, r-1$, поскольку $v \mid h$, а $h(c_i) \neq 0$ для всех $i = 1, \ldots, r-1$.

Наконец, $\deg f = k_1 + \cdots + k_r + \deg v \geq k_1 + \cdots + k_r$.$\lhd$

Следствие 6.10 *Пусть* $f, g \in K[X]$, $\deg f$, $\deg g \leq n$. *Если существуют различные элементы* $c_0, c_1, \ldots, c_n \in K$ *такие, что* $f(c_i) = g(c_i)$ *для всех* $i = 0, 1, \ldots, n$, *то* $f = g$.

Доказательство. Обозначим $f - g$ через h. Тогда $\deg h \leq \max(\deg f, \deg g) \leq n$ и $h(c_i) = f(c_i) - g(c_i) = 0$ $(i = 0, 1, \ldots, n)$, то есть, многочлен h степени не выше n имеет $n+1$ корней. В силу второй части утверждения теоремы 6.9 это возможно только при $h = 0$. Следовательно, $f = g$.$\lhd$

6.7 Производная многочлена

Пусть $f \in K[X]$, $f = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n$. *Производной* многочлена f называется многочлен

$$f' = na_0X^{n-1} + (n-1)a_1X^{n-2} + \cdots + a_{n-1}.$$

Замечание. В произвольном поле K выражение na для $n \in \mathbb{N}$ и $a \in K$ понимается как сумма n одинаковых слагаемых: $a + \cdots + a$.

Свойства:

1. $(\alpha f + \beta g)' = \alpha f' + \beta g'$ для всех $\alpha, \beta \in K$ и $f, g \in K[X]$.
2. $(fg)' = f'g + fg'$ для всех $f, g \in K[X]$.
3. $(f^k)' = kf^{k-1}f'$.

Доказательство. Свойство 1 очевидно. Справедливость свойства 2 ввиду свойства 1 достаточно установить для многочленов вида $f = X^n$, $g = X^m$ $(n, m \geq 1)$. Имеем:

$$(fg)' = (X^{n+m})' = (n+m)X^{n+m-1} = nX^{n-1}X^m + X^n(mX^{m-1}) = f'g + fg'.$$

Свойство 3 с учетом свойства 2 легко доказывается индукцией по k.$\lhd$

В дальнейшем будем считать, что K — одно из полей $\mathbb{Q}$, $\mathbb{R}$ или $\mathbb{C}$. Заметим, что в этом случае $\deg f' = \deg f - 1$.

Ранее было доказано (см. теорему 6.8), что кольцо $K[X]$ факториально, следовательно, для ненулевого многочлена $f \in K[X]$ существует разложение на неприводимые множители: $f = p_1^{k_1} \ldots p_r^{k_r}$. Неприводимые многочлены p_i, $(i = 1, \ldots, r)$, будем называть k_i-*кратными*

множителями многочлена f. Другими словами, p есть k-кратный множитель для f, если $p^k \mid f$, но $p^{k+1} \nmid f$.

Теорема 6.11 *Пусть неприводимый многочлен $p \in K[X]$ является k-кратным множителем многочлена f $(k \geq 1,\ \deg f \geq 1)$. Тогда p есть $(k-1)$-кратный множитель многочлена f'. В частности, $p \nmid f'$ при $k = 1$.*

Доказательство. Согласно условию теоремы имеем $f = p^k g$, где $p \nmid g$. Тогда

$$f' = kp^{k-1}p'g + p^k g' = p^{k-1}(kp'g + pg'),$$

откуда $p^{k-1} \mid f'$. Покажем, что $p^k \nmid f'$. Действительно, если $p^k \mid f'$, то $p \mid (kp'g + pg')$, значит, $p \mid kp'g$. Тогда в силу неприводимости p и факториальности $K[X]$ имеем $p \mid g$ или $p \mid (kp')$, но ни то, ни другое невозможно, так как $p \nmid g$ в силу выбора g и $p \nmid (kp')$, поскольку $\deg p > \deg(kp') > -\infty$. $\lhd$

Следствие 6.12 *Корень $c \in K$ многочлена $f \in K[X]$ не является простым $\Leftrightarrow$ $f(c) = f'(c) = 0$.*

Доказательство. $(\Rightarrow)$: $f(c) = 0$, так как c — корень f. Поскольку c является k-кратным корнем и $k \geq 2$, то $f = (X-c)^k g$, так что $X-c$ есть k-кратный неприводимый множитель f. Согласно теореме 6.11 получаем $(X-c)^{k-1} \mid f'$, следовательно, $f'(c) = 0$.

$(\Leftarrow)$: Согласно теореме Безу из $f(c) = 0$ выводим $(X-c) \mid f$. Пусть k — кратность множителя $X-c$. Тогда $X-c$ является $(k-1)$-кратным множителем для f'. Но $f'(c) = 0$ влечет $(X-c) \mid f'$, откуда $k-1 \geq 1$, значит, $k \geq 2$ и, следовательно, корень c не является простым. $\lhd$

Следствие 6.13 *Если $f = p_1^{k_1} \dots p_r^{k_r}$ — разложение f на неприводимые множители, то* $\text{НОД}\,(f, f') = p_1^{k_1-1} \dots p_r^{k_r-1}$.

Доказательство. По теореме 6.11 $p_i^{k_i} \mid f$ влечет $p_i^{k_i-1} \mid f'$. Следовательно, с учетом неприводимости множителей p_i имеем $f' = p_1^{k_1-1} \dots p_r^{k_r-1} g$ для некоторого $g \in K[X]$, где $\text{НОД}\,(p_i, g) = 1$ $(i = 1, \dots, r)$, откуда и вытекает требуемое утверждение. $\lhd$

Отметим, что с помощью следствия 6.13 можно найти многочлен h, являющийся произведением всех неприводимых множителей многочлена f, а именно,

$$h = \frac{f}{\text{НОД}\,(f, f')} = p_1 \dots p_r,$$

причем для нахождения h не обязательно знать исходные разложения f и f', достаточно лишь использовать алгоритм Евклида.

6.8 Неприводимые многочлены над полями $\mathbb{R}$ и $\mathbb{C}$

Вопрос о строении неприводимых многочленов с комплексными коэффициентами легко решается с помощью следующей теоремы, долгое время носившей название ОСНОВНОЙ ТЕОРЕМЫ АЛГЕБРЫ. Мы не будем приводить ее доказательство, поскольку оно опирается на некоторые факты теории функций комплексного переменного. (Желающие могут ознакомиться с ним, например, по книге [1], гл. 6.)

Теорема 6.14 *Любой многочлен* $f \in \mathbb{C}[X]$ $(\deg f \geq 1)$ *имеет корень в* $\mathbb{C}$.$\lhd$

Следствие 6.15 *Многочлен* $f \in \mathbb{C}[X]$ *неприводим* $\Leftrightarrow \deg f = 1$.

Доказательство. $(\Leftarrow)$: Неприводимость многочленов степени 1 над произвольным полем была доказана ранее (см. п. 6.3 “Делимость в кольце многочленов”).

$(\Rightarrow)$: Пусть f — комплексный многочлен и $\deg f \geq 1$. Согласно предыдущей теореме у f есть корень c, следовательно, по теореме Безу $f = (X - c)g$ для некоторого $g \in \mathbb{C}[X]$. С учетом неприводимости f последнее возможно только при $\deg g = 0$, откуда $\deg f = 1$.$\lhd$

Следствие 6.16 *Количество корней комплексного многочлена* f *(с учетом их кратностей) равно степени многочлена* f.

Доказательство. Разложим f на неприводимые множители: $f = p_1^{k_1} \dots p_r^{k_r}$. В силу следствия 6.15 имеем $\deg p_i = 1$ для всех i, так что каждый многочлен p_i отвечает некоторому корню c_i кратности k_i $(i = 1, \dots, r)$ и $\deg f = k_1 + \dots + k_r$.$\lhd$

Перейдем теперь к изучению неприводимых многочленов с вещественными коэффициентами. Разумеется, таковыми будут все многочлены 1-й степени. Кроме того, неприводимым будет любой многочлен 2-й степени вида $f = aX^2 + bX + c$, где $b^2 - 4ac < 0$, поскольку в этом случае уравнение $f(x) = 0$ не имеет вещественных корней и, следовательно, f невозможно разложить в произведение многочленов

степени 1. Покажем, что других неприводимых многочленов над полем $\mathbb{R}$ не существует.

Лемма 6.17 *Пусть $f \in \mathbb{R}[X]$ и $c = \alpha + i\beta$ — его комплексный корень. Тогда $\bar{c} = \alpha - i\beta$ — тоже корень f.*

Доказательство. Пусть $f = a_0X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n$. Согласно свойствам операции комплексного сопряжения с учетом вещественности коэффициентов многочлена f имеем:

$$0 = \bar{0} = \overline{f(c)} = a_0\bar{c}^n + a_1\bar{c}^{n-1} + \cdots + a_{n-1}\bar{c} + a_n = f(\bar{c}).\triangleleft$$

Предложение 6.18 *Если $f \in \mathbb{R}[X]$ неприводим, то либо $\deg f = 1$, либо f — многочлен степени 2, не имеющий вещественных корней.*

Доказательство. Ясно, что если f имеет вещественный корень c, то неприводимым он может быть только в случае ассоциированности с многочленом $X - c$, следовательно, $\deg f = 1$.

Таким образом, остается рассмотреть случай, когда вещественных корней у f нет. Очевидно, тогда $\deg f \geq 2$. Поскольку $\mathbb{R} \subset \mathbb{C}$, на f можно смотреть, как на многочлен над полем $\mathbb{C}$, следовательно, у f есть комплексный корень $c = \alpha + i\beta$ ($\beta \neq 0$). Применяя лемму 6.17, получаем, что $\bar{c} = \alpha - i\beta$ также является корнем f. В силу теоремы Безу f делится на $X - c$ и на $X - \bar{c}$, а ввиду взаимной простоты этих линейных многочленов, f делится на их произведение. Но $g = (X - c)(X - \bar{c}) = (X - \alpha - i\beta)(X - \alpha + i\beta) = X^2 - 2\alpha X + (\alpha^2 + \beta^2)$ — многочлен с вещественными коэффициентами, поэтому из неприводимости f и условия $g \mid f$ вытекает ассоциированность f и g.$\triangleleft$

Литература

1. Кострикин А.И. Введение в алгебру. Ч.1. Основы алгебры. М.: Наука, 2001.

2. Кострикин А.И. Введение в алгебру. Ч.2. Линейная алгебра. М.: Наука, 2001.

3. Винберг Э.Б. Курс алгебры. М.: Факториал Пресс, 2001.

4. Сборник задач по алгебре (под редакцией Кострикина А.И.). М.: Физматлит, 2001.

5. Курош А.Г. Курс высшей алгебры. М.: Наука, 1975.

6. Проскуряков И.В. Сборник задач по линейной алгебре. М.: Лаб. баз. зн. 2001.

7. Корешков Н.А. Линейные операторы. Учебное пособие. Изд-во КГУ. Казань, 2005.

Содержание