Начала алгебры, часть I: алгебраические структуры; комплексные числа; системы линейных уравнений; матрицы; определители

А. А. Михалёв, А. В. Михалёв

Оглавление

Глава 1

Введение: основные алгебраические структуры

В этой главе мы представим вниманию читателя основные алгебраические структуры, с которыми мы встретимся при изложении курса и при решении задач. Детальное знакомство с ними будет происходить по мере нашего продвижения и накопления фактического материала. Преимущество работы с абстрактными математическими понятиями может быть оценено лишь при необходимости рассматривать многочисленные частные примеры.

Предмет алгебры существенно менялся с течением времени: арифметические действия над натуральными и положительными рациональными числами в глубокой древности (3 век н. э.); алгебраические уравнения первой и второй степени (9 век); появление алгебраической символики (15—17 века); к 18-му веку алгебра сложилась в том объёме, который сейчас принято называть «элементарной алгеброй»; в 18—19 веках алгебра — это прежде всего алгебра многочленов; с середины 19-го века центр тяжести алгебраических исследований перемещается на изучение произвольных алгебраических операций. Изучение алгебраических структур (т. е. множеств с определёнными на них операциями) было подготовлено развитием числовых систем (построением комплексных чисел и кватернионов), созданием матричного исчисления, возникновением булевой алгебры, внешней алгебры Грассмана, исследованием групп подстановок. Таким образом, к 20-му веку сформировалась точка зрения

на современную алгебру как на общую теорию алгебраических операций (под влиянием работ Д. Гильберта, Э. Артина, Э. Нётер и с выходом в 1930 г. монографии Б. Л. ван дер Вардена «Современная алгебра»).

1.1. Алгебраические операции

Если M — непустое множество, n — натуральное число, то через M^n обозначим множество упорядоченных последовательностей $(m_1, m_2, \ldots, m_n)$, $m_i \in M$, $1 \leqslant i \leqslant n$. Под *$n$-арной алгебраической операцией* на множестве M понимается отображение

$$\omega\colon M^n \to M,$$

число n называется *арностью алгебраической операции* ω. Исторически сначала возникли бинарные операции ($n = 2$) и унарные операции ($n = 1$). Нульарные операции — это фиксированные элементы множества M, поскольку под $M^{(0)}$ понимается одноэлементное множество.

1.2. Группоиды, полугруппы, моноиды

Непустое множество M с бинарной операцией $\omega\colon M \times M \to M$ называется *группоидом*. Иногда нам удобнее использовать обозначение

$$m_1 \,\omega\, m_2 = \omega((m_1, m_2)), \quad m_1, m_2 \in M.$$

Бинарная операция $\omega\colon M \times M \to M$ называется *ассоциативной*, если

$$(m_1 \,\omega\, m_2) \,\omega\, m_3 = m_1 \,\omega\, (m_2 \,\omega\, m_3) \text{ для всех } m_1, m_2, m_3 \in M,$$

и *коммутативной*, если

$$m_1 \,\omega\, m_2 = m_2 \,\omega\, m_1 \text{ для всех } m_1, m_2 \in M.$$

Упражнение 1.2.1.

1) Бинарная операция разность целых чисел

$$-\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, \quad (m_1, m_2) \mapsto m_1 - m_2,$$

не является ассоциативной и не является коммутативной.

2) Следующие бинарные операции ассоциативны и коммутативны:

2.1)
$$+\colon \mathbb{N}\times\mathbb{N}\to\mathbb{N},\quad (m_1,m_2)\mapsto m_1+m_2$$
(сложение натуральных чисел);
$$\cdot\colon \mathbb{N}\times\mathbb{N}\to\mathbb{N},\quad (m_1,m_2)\mapsto m_1m_2$$
(умножение натуральных чисел);

2.2) пусть $\mathcal{P}(M)$ — множество всех подмножеств (включая пустое) множества M,
$$\cap\colon \mathcal{P}(M)\times\mathcal{P}(M)\to\mathcal{P}(M),\quad (A,B)\mapsto A\cap B$$
(пересечение подмножеств);
$$\cup\colon \mathcal{P}(M)\times\mathcal{P}(M)\to\mathcal{P}(M),\quad (A,B)\mapsto A\cup B$$
(объединение подмножеств);
$$*\colon \mathcal{P}(M)\times\mathcal{P}(M)\to\mathcal{P}(M),$$
$$(A,B)\mapsto A*B=(A\setminus B)\cup(B\setminus A)=(A\cup B)\setminus(A\cap B)$$
(симметрическая разность подмножеств).

3) Пусть $T(M)=M^M=\{f\colon M\to M\}$ — совокупность всех отображений из множества M в множество M,
$$\circ\colon M^M\times M^M\to M^M,\quad (f,g)\mapsto f\circ g,$$
где $(f\circ g)(m)=f(g(m))$ для $m\in M$ (композиция отображений). Тогда $\circ$ — ассоциативная операция (она является коммутативной тогда и только тогда, когда $|M|=1$, т. е. M — одноэлементное множество), подробнее см. задачу 1.7.1.

4) Бинарная операция
$$\mathbb{N}\times\mathbb{N}\to\mathbb{N},\quad (m,n)\mapsto m^n$$
(возведение в степень) неассоциативна и некоммутативна; бинарная операция
$$\omega\colon \mathbb{N}\times\mathbb{N}\to\mathbb{N},\quad (m,n)\mapsto m^n+n^m$$
коммутативна, но не является ассоциативной ($(1\,\omega\,2)\,\omega\,3\neq 1\,\omega\,(2\,\omega\,3)$).

5) Если (M, ω) — группоид с бинарной операцией $\omega\colon M \times M \to M$, то подмножество $L \subseteq M$, для которого

$$l_1 \,\omega\, l_2 \in L \quad \text{для всех} \quad l_1, l_2 \in L$$

(*замкнутое* относительно операции ω), является группоидом,

$$\omega|_{L\times L}\colon L \times L \to L, \quad (l_1, l_2) \mapsto l_1 \,\omega\, l_2,$$

называемым *подгруппоидом*. Например:

а) $(\mathbb{N}, +)$ — подгруппоид в группоиде $(\mathbb{Z}, +)$ (здесь $\mathbb{Z}$ — целые числа);

б) подмножество $\mathbb{Z} \setminus \{0\}$ не является замкнутым в группоиде $(\mathbb{Z}, +)$ относительно операции сложения.

Пусть (M_1, ω_1) и (M_2, ω_2) — группоиды. Отображение

$$f\colon M_1 \to M_2$$

называется *гомоморфизмом группоидов*, если

$$f(x \,\omega_1\, y) = f(x) \,\omega_2\, f(y) \quad \text{для всех} \quad x, y \in M_1.$$

Биективный гомоморфизм группоидов называется *изоморфизмом группоидов* (в случае его наличия группоиды (M_1, ω_1) и (M_2, ω_2) называются *изоморфными*; обозначение $M_1 \cong M_2$).

Лемма 1.2.2.

1) *Пусть* f_1 *и* f_2, *где*

$$(M_1, \omega_1) \xrightarrow{f_1} (M_2, \omega_2) \xrightarrow{f_2} (M_3, \omega_3),$$

являются гомоморфизмами группоидов. Тогда их произведение $f_2 f_1$,

$$f_2 f_1\colon (M_1, \omega_1) \to (M_3, \omega_3), \quad (f_2 f_1)(m_1) = f_2(f_1(m_1)),$$

также является гомоморфизмом группоидов.

2) *Пусть* $f\colon (M_1,\omega_1)\to(M_2,\omega_2)$ *— изоморфизм группоидов, тогда обратное отображение*

$$f^{-1}\colon (M_2,\omega_2)\to(M_1,\omega_1)$$

также является изоморфизмом группоидов.

Доказательство.

1) Для любых $x,y\in M_1$ имеем

$$(f_2f_1)(x\,\omega_1\,y)=f_2(f_1(x\,\omega_1\,y))=f_2(f_1(x)\,\omega_2\,f_1(y))=$$
$$=f_2(f_1(x))\,\omega_3\,f_2(f_1(y))=(f_2f_1)(x)\,\omega_3\,(f_2f_1)(y).$$

2) Пусть $z,w\in M_2$, $z=f(x)$, $w=f(y)$, где $x=f^{-1}(z)$, $y=f^{-1}(w)\in M_1$. Тогда

$$f^{-1}(z\,\omega_2\,w)=f^{-1}(f(x)\,\omega_2\,f(y))=f^{-1}(f(x\,\omega_1\,y))=$$
$$=x\,\omega_1\,y=f^{-1}(z)\,\omega_1\,f^{-1}(w).\quad\square$$

Следствие 1.2.3. *Отношение «быть изоморфными» является отношением эквивалентности на классе группоидов:* $(M,\omega)\cong(M,\omega)$; *если* $(M_1,\omega_1)\cong(M_2,\omega_2)$, *то* $(M_2,\omega_2)\cong(M_1,\omega_1)$; *если* $(M_1,\omega_1)\cong$ $\cong(M_2,\omega_2)$ *и* $(M_2,\omega_2)\cong(M_3,\omega_3)$, *то* $(M_1,\omega_1)\cong(M_3,\omega_3)$.

Упражнение 1.2.4.

1) Тождественное отображение

$$1_M\colon M\to M,\quad 1_M(m)=m,$$

является изоморфизмом

$$1_M\colon (M,\omega)\to(M,\omega)$$

группоидов.

2) Отображения

$$f\colon(\mathbb{N},+)\to(\mathbb{N},\cdot),\qquad f(n)=2^n,$$
$$f_k\colon(\mathbb{N},+)\to(\mathbb{N},+),\quad f_k(n)=kn,\quad k\in\mathbb{N},$$

являются гомоморфизмами группоидов (но отображение

$$f_k\colon(\mathbb{N},\cdot)\to(\mathbb{N},\cdot),\quad f_k(n)=kn,\quad k\neq 1,$$

не является гомоморфизмом группоидов).

Пусть (M, ω) — группоид, элемент $e \in M$ называется (двусторонним) *нейтральным элементом*, если

$$e \,\omega\, m = m = m \,\omega\, e \quad \text{для всех} \quad m \in M.$$

Упражнение 1.2.5. Следующие элементы являются нейтральными:

1) 0 в $(\mathbb{N} \cup \{0\}, +)$, $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$;

2) 1 в $(\mathbb{N}, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$;

3) 1_M в $(M^M, \circ)$;

4) M в $(\mathcal{P}(M), \cap)$;

5) $\varnothing$ в $(\mathcal{P}(M), \cup)$ и в $(\mathcal{P}(M), *)$;

6) в $(\mathbb{N}, +)$ нет нейтральных элементов (в России $\mathbb{N} = \{1, 2, \ldots\}$).

Лемма 1.2.6. *Пусть (M, ω) — группоид, e и e' — нейтральные элементы. Тогда $e = e'$ (другими словами, если в группоиде существует нейтральный элемент, то он единственный).*

Доказательство. $e = e \,\omega\, e' = e'$. □

Замечание 1.2.7. В мультипликативных обозначениях операции ω в моноиде, $m_1 \,\omega\, m_2 = m_1 m_2$, нейтральный элемент часто называют единицей и используют для него обозначение $1 = e_M$; в аддитивных обозначениях, $m_1 \,\omega\, m_2 = m_1 + m_2$, нейтральный элемент обычно называют нулём и используют для него обозначение $0 = 0_M$.

Определение 1.2.8. Группоид (M, ω) с бинарной операцией

$$\omega \colon M \times M \to M$$

называется *полугруппой*, если операция ω ассоциативна; *моноидом*, если операция ассоциативна (т. е. это полугруппа) и в (M, ω) существует нейтральный элемент e.

Замечания 1.2.9.

1) Подгруппоид (L, ω) полугруппы (M, ω) является полугруппой и называется *подполугруппой*.

2) Гомоморфизм (изоморфизм) группоидов, являющихся полугруппами, называется *гомоморфизмом (изоморфизмом) полугрупп*.

3) *Подмоноидом* моноида (M, ω, e_M) называется подполугруппа (L, ω, e_M) (таким образом, подмножество $L \subseteq M$ замкнуто относительно операции ω), содержащая нейтральный элемент e_M.

4) Если (M, ω, e_M) и $(M', \omega', e_{M'})$ — моноиды, то под *гомоморфизмом моноидов* понимается гомоморфизм полугрупп

$$f\colon (M, \omega, e_M) \to (M', \omega', e_{M'})$$

такой, что $f(e_M) = e_{M'}$. Ясно, что произведение гомоморфизмов моноидов — гомоморфизм моноидов, обратное отображение к изоморфизму моноидов — изоморфизм моноидов.

Определение 1.2.10. Пусть $(M, \cdot, e_M)$ — моноид и $m \in M$.

1) Элемент $m' \in M$, для которого $mm' = e_M$, называется *правым обратным* элемента m.

2) Элемент $m'' \in M$, для которого $m''m = e_M$, называется *левым обратным* элемента m.

3) Элемент $\bar{m} \in M$ называется *двусторонним обратным* элемента m, если $m\bar{m} = e_M = \bar{m}m$ (в этом случае элемент m называется *обратимым*).

Лемма 1.2.11. *Если в моноиде $(M, \cdot, e_M)$ элемент $m \in M$ имеет правый обратный m' и левый обратный m'', то $m' = m''$ и m является обратимым элементом.*

Доказательство.

$$m' = e_M m' = (m''m)m' = m''(mm') = m''e_M = m''. \quad \square$$

Следствие 1.2.12.

1) *Двусторонний обратный элемент $\bar{m}$ элемента m моноида $(M, \cdot, e_M)$ определён (если он существует) однозначно, для него используется мультипликативное обозначение m^{-1}.*

2) *Если для элемента m моноида $(M, \cdot, e_M)$ существует обратный элемент m^{-1}, то $(m^{-1})^{-1} = m$.*

3) *Если элементы x, y моноида $(M, \cdot, e_M)$ обратимы с обратными x^{-1} и y^{-1}, то*

$$(xy)^{-1} = y^{-1}x^{-1}.$$

Действительно (при этом см. теорему 1.3.2),

$$(y^{-1}x^{-1})(xy) = y^{-1}x^{-1}xy = y^{-1}e_M y = y^{-1}y = e_M = \\ = yy^{-1} = ye_M y^{-1} = xyy^{-1}x^{-1} = (xy)(y^{-1}x^{-1}).$$

1.3. Обобщённая ассоциативность (применение ассоциативной операции к n сомножителям при $n \geqslant 3$)

Рассмотрим ассоциативную операцию $*$ на множестве M, $*\colon M \times M \to M$, $(a, b) \mapsto a * b \in M$ для $a, b \in M$, при этом

$$(a * b) * c = a * (b * c) \ \forall a, b, c \in M.$$

Для одного или двух сомножителей нет вопроса о различных расстановках скобок: a, $a * b$. Для трёх сомножителей a, b, c существует всего две расстановки скобок $(a * b) * c$, $a * (b * c)$, обозначающие применение бинарной операции $*$ (каждый раз применяемой к двум элементам). На множестве из четырёх элементов a_1, a_2, a_3, a_4 расстановок скобок уже значительно больше: $((a_1 * a_2) * a_3) * a_4$ (регулярная слева расстановка), $(a_1 * a_2) * (a_3 * a_4)$, $(a_1 * (a_2 * a_3)) * a_4$, $a_1 * ((a_2 * a_3) * a_4)$, $a_1 * (a_2 * (a_3 * a_4))$ (регулярная справа расстановка).

Задача 1.3.1 (трудная). Найти число всех различных расстановок скобок для применения бинарной операции на n сомножителях.

Определение ассоциативной бинарной операции ($(a * b) * c = a * (b * c)$ для всех $a, b, c \in M$) означает, что для трёх сомножителей результат применения операции не зависит от расстановки скобок (т. е. порядка её применения). Наша ближайшая цель — показать, что для ассоциативной бинарной операции это утверждение

верно и для n сомножителей $a_1, a_2, \ldots, a_n$ (при всех расстановках скобок после соответствующего применения операции $*$ мы получаем один и тот же элемент, который можно обозначить $a_1 * a_2 * \ldots * a_n$, без указания расстановки скобок).

Теорема 1.3.2. *Пусть $*$ — бинарная ассоциативная операция на множестве M ($*\colon M \times M \to M$, $(a, b) \mapsto a * b$ для $a, b \in M$, и $(a * b) * c = a * (b * c)$ для всех $a, b, c \in M$), $a_1, a_2, \ldots, a_n \in M$, $n \geqslant 3$. Тогда результат применения операции $*$ к n сомножителям $a_1, a_2, \ldots, a_n$ не зависит от расстановки скобок.*

Доказательство проведём индукцией по n по вполне упорядоченному множеству $\{n \in \mathbb{N} \mid n \geqslant 3\}$, это означает, что любое непустое подмножество этого множества имеет наименьший элемент.

Начало индукции $n = 3$ обеспечено определением ассоциативной операции.

Допустим, что утверждение верно для всех k, $3 \leqslant k < n$. Рассмотрим произвольную расстановку скобок на n сомножителях $a_1, a_2, \ldots, a_n$, соответствующую применениям бинарной операции $*$ (каждый раз к двум элементам). Наша цель — доказать, что результат применения операции $*$ для произвольной расстановки скобок совпадает с результатом применения для *регулярной слева* расстановки скобок $(\ldots((a_1 * a_2) * a_3) * \ldots) * a_n$. При этом для $n = 2$ имеем $a_1 * a_2$, а для $n = 1$ имеем a_1.

В каждой расстановке скобок есть последнее применение операции $*$ (например: $(a*b)\circledast c$; $(a_1*a_2)\circledast(a_3*a_4)$; $((a_1*a_2)*a_3)\circledast(a_4*a_5)$, здесь $\circledast$ обозначает последнее применение операции $*$ в каждой из приведённых расстановок). Таким образом, последнее применение операции $*$ происходит к произведению k сомножителей $a_1, a_2, \ldots, a_k$ с некоторой расстановкой скобок и к произведению $(n - k)$ сомножителей $a_{k+1}, \ldots, a_n$ с некоторой расстановкой скобок, при этом $1 \leqslant k < n$, $1 \leqslant n - k < n$. Результат произведения операции $*$ в левом и правом блоке не зависит от расстановки скобок (возможно, $k = 1$ или $k = 2$; возможно, $n - k = 1$ или $n - k = 2$; если $3 \leqslant k < n$, то в силу индуктивного предположения; если $3 \leqslant n - k < n$, то также в силу индуктивного предположения). Выберем в левом произведении *регулярную слева* расстановку скобок, а в правом произведении — *регулярную справа* расстановку скобок.

Тогда имеем, применяя последовательно ассоциативность для трёх сомножителей,

$$\begin{aligned}
&[(\dots((a_1 * a_2) * a_3)\dots a_{k-1}) * a_k] * \\
&\quad * [a_{k+1} * (a_{k+2} * (\dots(a_{n-1} * a_n)\dots))] = \\
&= [((\dots((a_1 * a_2) * a_3)\dots a_{k-1}) * a_k) * a_{k+1}] * \\
&\quad * [a_{k+2} * (\dots(a_{n-1} * a_n)\dots)] = \\
&\dots \\
&= (\dots((a_1 * a_2) * a_3)\dots a_{n-1}) * a_n
\end{aligned}$$

(в наших обозначениях, если $k = 1$, то $[a_1] = a_1$; аналогично, если $n - k = 1$, то $[a_n] = a_n$).

Итак, результат применения операции $*$ в соответствии с исходной (произвольной) расстановкой скобок совпал с результатом применения при регулярной слева расстановке скобок. Таким образом, результат применения ассоциативной операции не зависит от расстановки скобок. □

1.4. Отображения множеств

Пусть U, V — непустые множества, $f\colon U \to V$ — (однозначное) отображение из множества U в множество V, т. е. каждому элементу $u \in U$ сопоставляется элемент $f(u) \in V$.

Замечание 1.4.1.

1) Сохраняя единообразие с курсом анализа, мы обозначаем применение отображения f к элементу $u \in U$ через $f(u)$, т. е. f пишем слева от u. Возможно (а иногда и удобнее) было бы использовать обозначение uf.

2) Если $f'\colon U \to V$, то $f = f'$, если для любого $u \in U$ имеем $f(u) = f'(u)$.

3) Категория Set, в которой объекты — множества, морфизмы — отображения множеств, является одной из основных категорий в математике.

1.5. Инъективные, сюръективные, биективные отображения

Рассмотрим *образ отображения* $f\colon U \to V$

$$\operatorname{Im} f = \{v \in V \mid v = f(u),\ u \in U\}.$$

Можно рассмотреть также полезное отношение эквивалентности τ_f на множестве U, определяемое отображением $f\colon U \to V$,

$$u_1 \,\tau_f\, u_2 \iff f(u_1) = f(u_2).$$

Определение 1.5.1. Отображение $f\colon U \to V$ называется:

1) *инъективным*, если разные элементы в U при отображении f переходят в разные элементы в V (т. е. $u_1, u_2 \in U,\ u_1 \neq u_2 \implies f(u_1) \neq f(u_2)$),

2) *сюръективным*, если каждый элемент в V является образом некоторого элемента из U (т. е. $\forall v \in V\ \exists u \in U,\ v = f(u)$, другими словами, $\operatorname{Im} f = V$),

3) *биективным*, если отображение f инъективно и сюръективно (т. е. $\forall v \in V\ \exists! u \in U,\ v = f(u)$).

Замечание 1.5.2.

1) В более ранней математической литературе для биективного отображения использовалась более длинная комбинация слов: «взаимно однозначное отображение на»,

2) иногда для сюръективного отображения $f\colon U \to V$ мы будем говорить, что «f отображает множество U *на* множество V».

Задачи 1.5.3.

1) Пусть $|U| = m$, $|V| = n$. Доказать, что $|\{f\colon U \to V\}| = n^m$.

2) Пусть $|U| = m$, $\mathcal{L}(U)$ — совокупность всех подмножеств множества U (включая пустое подмножество). Доказать, что $|\mathcal{L}(U)| = 2^m$.

Указание. Для подмножества $T \subseteq U$ рассмотреть его *характеристическую функцию*

$$C_T\colon U \to \{0,1\}, \quad C_T(u) = \begin{cases} 1, & u \in T, \\ 0, & u \notin T. \end{cases}$$

Следствие. $1 + C_n^1 + \ldots + C_n^n = 2^n$.

3) Найти число инъективных (сюръективных) отображений $f\colon U \to V$, где $|U| = m$, $|V| = n$.

Пример 1.5.4.

1) Отображение $f\colon \mathbb{N} \to \mathbb{N}$, $f(n) = n + 1$, является инъективным, но не является сюръективным.

2) Отображение $f\colon \mathbb{N} \to \mathbb{N}$, $f(1) = 1$ и $f(n) = n - 1$ для $n > 1$, является сюръективным, но не является инъективным.

3) Тождественное отображение $1_U\colon U \to U$, $1_U(u) = u$ для всех $u \in U$, очевидно, является биекцией.

Лемма 1.5.5. *Пусть U — конечное множество, $f\colon U \to U$. Тогда равносильны условия*:

1) *f — инъективное отображение*;

2) *f — сюръективное отображение.*

Доказательство.

1) $\Longrightarrow$ 2) Пусть $|U| = n < \infty$. Так как f — инъективное отображение, то $|\operatorname{Im} f| = n$. Поскольку $\operatorname{Im} f \subseteq U$, $|\operatorname{Im} f| = n = |U|$, то $\operatorname{Im} f = U$, т. е. f — сюръективное отображение.

2) $\Longrightarrow$ 1) Допустим противное, т. е. что f не является инъективным отображением. Тогда $f(u_1) = f(u_2)$ для некоторых $u_1, u_2 \in U$, $u_1 \neq u_2$. Следовательно, $|\operatorname{Im} f| < n = |U|$, поэтому $\operatorname{Im} f < U$, т. е. отображение f не является сюръективным, что приводит к противоречию. □

Замечание 1.5.6. Условие конечности множества U в лемме 1.5.5 существенно, как показывает пример 1.5.4. Более того, это соображение может быть использовано для характеризации конечных множеств в терминах отображений.

1.6. Произведение отображений

Определение 1.6.1. Для диаграммы отображений

$$U \xrightarrow{f} V \xrightarrow{g} W$$

определим *произведение* (иногда называемое *композицией* или *суперпозицией*) $h = gf$ отображений f и g следующим образом:

$$h = gf\colon U \to W, \quad h(u) = g(f(u))$$

для $u \in U$.

Замечание 1.6.2. Не любые два отображения можно перемножить!

Примеры 1.6.3.

1) Если $1_U\colon U \to U$ — тождественное отображение множества U, $1_V\colon V \to V$ — тождественное отображение множества V, $f\colon U \to V$, то
$$f1_U = f = 1_V f.$$

2) Если $f\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, $f(k) = k+1$ для $1 \leqslant k < n$, $f(n) = 1$, то $f^n = 1_U$, где $U = \{1, 2, \ldots, n\}$.

Теорема 1.6.4 (об ассоциативности произведения отображений). *Для диаграммы отображений*

$$U \xrightarrow{f} V \xrightarrow{g} W \xrightarrow{h} Z$$

имеем $h(gf) = (hg)f$.

Доказательство. Ясно, что

$$h(gf)\colon U \to Z, \quad (hg)f\colon U \to Z.$$

Для любого $u \in U$ имеем

$$(h(gf))(u) = h((gf)(u)) = h(g(f(u))),$$
$$((hg)f)(u) = (hg)(f(u)) = h(g(f(u))),$$

таким образом, $(h(gf))(u) = ((hg)f)(u)$ для всех $u \in U$, следовательно, $h(gf) = (hg)f$. □

1.7. Моноид отображений множества

Пусть U — множество, $\mathrm{T}(U) = \{f\colon U \to U\}$ — совокупность отображений с операцией произведения отображений. В силу доказанной теоремы 1.6.4 эта операция ассоциативна. Нейтральным элементом относительно этой операции является тождественное отображение 1_U. Итак, $\mathrm{T}(U)$ — *полугруппа с единицей*, т. е. *моноид*.

Задача 1.7.1. Моноид отображений $\mathrm{T}(U)$ множества U коммутативен тогда и только тогда, когда $|U| = 1$ (т. е. множество U состоит из одного элемента).

Указание. Если $a \in U$, то рассмотрим отображение $f_a\colon U \to U$, $f_a(u) = a$ для всех $u \in U$. Если $a \neq b$, то $f_a f_b = f_a \neq f_b = f_b f_a$.

1.8. Характеризация инъективных, сюръективных и биективных отображений (в терминах произведений отображений)

Теорема 1.8.1. *Пусть $f\colon U \to V$ — отображение непустых множеств. Тогда:*

1) *f — инъективное отображение тогда и только тогда, когда существует отображение $g\colon V \to U$ такое, что $gf = 1_U$ (т. е. существует левый обратный элемент для отображения f),*

2) *f — сюръективное отображение тогда и только тогда, когда существует отображение $g\colon V \to U$ такое, что $fg = 1_V$ (т. е. существует правый обратный элемент для отображения f),*

3) *f — биективное отображение тогда и только тогда, когда существует отображение $g\colon V \to U$ такое, что $gf = 1_U$ и $fg = 1_V$ (т. е. существует левый и правый обратные для отображения f).*

Доказательство.

1а) Пусть $f\colon U \to V$ — инъективное отображение. Построим отображение $g\colon V \to U$ следующим образом. Если $v \in \operatorname{Im} f \subseteq V$ и $v = f(u)$, $u \in U$, то этот элемент u определён единственным образом (в силу инъективности отображения f). В этом случае положим

$g(v) = u$. Для всех элементов $v \in V \setminus \operatorname{Im} f$ положим $g(v) = u_0 \in U$, где u_0 — некоторый фиксированный элемент в U. Тогда для всякого элемента $u \in U$ имеем

$$(gf)(u) = g(f(u)) = u = 1_U(u),$$

т. е. $gf = 1_U$.

1б) Если существует отображение $g\colon V \to U$ такое, что $gf = 1_U$, и $f(u_1) = f(u_2)$ для $u_1, u_2 \in U$, то

$$u_1 = 1_U(u_1) = (gf)(u_1) = g(f(u_1)) = \\ = g(f(u_2)) = (gf)(u_2) = 1_U(u_2) = u_2.$$

Итак, f — инъективное отображение.

2а) Пусть $f\colon U \to V$ — сюръективное отображение. Для каждого элемента $v \in V$ множество $\{u \in U \mid f(u) = v\}$ не является пустым. Выберем в нём один элемент u_v (для интересующихся аксиоматикой теории множеств: это можно сделать в силу аксиомы выбора). Определим отображение $g\colon V \to U$, полагая $g(v) = u_v$. Тогда

$$(fg)(v) = f(g(v)) = f(u_v) = v = 1_V(v).$$

Таким образом, $fg = 1_V$.

2б) Если $fg = 1_V$ для некоторого отображения $g\colon V \to U$, то для всякого $v \in V$ имеем

$$v = 1_V(v) = (fg)(v) = f(g(v)),$$

т. е. $v = f(u)$ для $u = g(v)$, следовательно, $f\colon U \to V$ — сюръективное отображение.

3а) Если $f\colon U \to V$ — биекция, то для всякого элемента $v \in V$ существует, и единственный, элемент $u \in U$ такой, что $v = f(u)$. В этом случае положим $g(v) = u$. Получим отображение $g\colon V \to U$, для которого:

$$(gf)(u) = g(f(u)) = u$$

для всякого $u \in U$, т. е. $gf = 1_U$;

$$(fg)(v) = f(g(v)) = f(g(f(u))) = f(u) = v$$

для всякого $v \in V$, т. е. $fg = 1_V$.

Замечание 1.8.2. Можно было воспользоваться уже доказанными утверждениями 1а), 2а): из инъективности отображения $f: U \to V$ следует существование отображения $g: V \to U$, для которого $gf = 1_U$; из сюръективности отображения $f: U \to V$ следует существование отображения $g': V \to U$, для которого $fg' = 1_V$; но тогда

$$g' = 1_U g' = (gf)g' = g(fg') = g1_V = g;$$

таким образом,

$$gf = 1_U, \quad fg = 1_V.$$

3б) Если существует отображение $g: V \to U$, для которого $gf = 1_U$ и $fg = 1_V$, то в силу 1б), f — инъекция, а в силу 2б), f — сюръекция, т. е. f — биекция. □

Замечание 1.8.3. Отображение g, для которого $gf = 1_U$, $fg = 1_V$, как мы показали, определено однозначно. Оно будет обозначаться $g = f^{-1}$.

Лемма 1.8.4. *Пусть* $U \xrightarrow{f} V \xrightarrow{g} W$.

1) *Если* f, g — *инъекции, то* gf — *инъекция.*

2) *Если* f, g — *сюръекции, то* gf — *сюръекция.*

3) *Если* f, g — *биекции, то* gf — *биекция.*

4) *Если* f — *биекция, то отображение* $g = f^{-1}$ — *биекция.*

Доказательство.

1) Если $u_1, u_2 \in U$, $u_1 \neq u_2$, то $f(u_1) \neq f(u_2)$, и $(gf)(u_1) = g(f(u_1)) \neq g(f(u_2)) = (gf)(u_2)$, т. е. gf — инъекция.

2) Если $w \in W$, то $w = g(v)$ для некоторого $v \in V$; далее, $v = f(u)$ для некоторого $u \in U$; поэтому $w = g(f(u)) = (gf)(u)$, т. е. отображение gf является сюръекцией.

3) следует из 1) и 2).

4) Так как $gf = 1_U$, $fg = 1_V$, то $f = g^{-1}$, и поэтому $g = f^{-1}$ является биекцией. □

1.9. Группы

Одним из основных общематематических понятий является понятие группы.

Определение 1.9.1. Непустое множество G с бинарной операцией $*\colon G \times G \to G$, $(a, b) \to a * b \in G$ для $a, b \in G$, называется *группой*, если:

1) операция ассоциативна (т. е. $(a * b) * c = a * (b * c)$ для всех $a, b, c \in G$);

2) существует нейтральный элемент $e \in G$ (т. е. $g * e = g = e * g$ для всех $g \in G$);

3) для каждого элемента $g \in G$ существует обратный элемент $g^{-1} \in G$ (т. е. $g * g^{-1} = e = g^{-1} * g$).

Замечание 1.9.2. Напомним, что нейтральный элемент (при мультипликативной записи называемый единицей группы) единственный. Обратный элемент g^{-1} для элемента $g \in G$ определён однозначно. Коммутативная группа часто называется *абелевой группой*.

Лемма 1.9.3. *Если G — группа, $a, b \in G$, то*

1) *уравнение $ax = b$ имеет, и только одно, решение $x = a^{-1}b$;*

2) *уравнение $ya = b$ имеет, и только одно, решение $y = ba^{-1}$;*

3) *если $ab = ac$, то $b = c$; если $ba = ca$, то $b = c$;*

4) *если $x^2 = x$, то $x = e$;*

5) $(ab)^{-1} = b^{-1}a^{-1}$; $(a_1 \dots a_n)^{-1} = a_n^{-1} \dots a_1^{-1}$; $(a^{-1})^{-1} = a$.

Доказательство.

1) Ясно, что $a(a^{-1}b) = b$. Если же $ax = b$ для $x \in G$, то $x = a^{-1}ax = a^{-1}b$.

2) Ясно, что $(ba^{-1})a = b$. Если же $ya = b$ для $y \in G$, то $y = (ya)a^{-1} = ba^{-1}$.

3) и 4) следуют из 1) и 2).

5) проверяется непосредственно. □

Примеры 1.9.4 (примеры групп).

1) Целые числа $\mathbb{Z}$, рациональные числа $\mathbb{Q}$, действительные числа $\mathbb{R}$ с операцией сложения. Заметим, что: а) натуральные числа $\mathbb{N}$ с операцией сложения группой не являются (отсутствует нейтральный элемент); б) натуральные числа с нулём $\mathbb{N}_0$ также не являются группой (обратный элемент (в аддитивной записи обычно называемый противоположным элементом) существует только для 0; таким образом, например, 1 уже не имеет обратного элемента).

2) *Группа вычетов* $(\mathbb{Z}_n, +)$ *по модулю* n. Пусть $(\mathbb{Z}, +)$ — группа целых чисел по сложению, $1 < n \in \mathbb{N}$. Для $k \in \mathbb{Z}$ пусть

$$C_k = k + n\mathbb{Z} = \{k + nq \mid q \in \mathbb{Z}\}$$

(сдвиг подгруппы $n\mathbb{Z}$ на элемент k). Ясно, что $C_k = C_l$, $l \in \mathbb{Z}$, тогда и только тогда, когда $k - l = nq$, $q \in \mathbb{Z}$. Так как

$$k = nq + r, \quad \text{где} \quad q \in \mathbb{Z}, \quad 0 \leqslant r < n,$$

то $C_k = C_r$. Таким образом, множество различных сдвигов

$$\mathbb{Z}_n = \{C_0, C_1, \ldots, C_{n-1}\}$$

находится в биективном соответствии с множеством остатков $\{0, 1, 2, \ldots, n-1\}$ при делении на число n.

Определим операцию сложения на множестве $\mathbb{Z}_n$, полагая

$$C_k + C_l = C_{k+l} = C_s, \quad \text{где} \quad k + l = n\bar{q} + s, \quad 0 \leqslant s \leqslant n-1, \quad \bar{q} \in \mathbb{Z}.$$

Проверим корректность этой операции. Если $C_k = C_{k'}$, $C_l = C_{l'}$, то $k' = k + nu$, $l' = l + nv$, $u, v \in \mathbb{Z}$, следовательно,

$$k' + l' = (k + nu) + (l + nv) = (k + l) + n(u + v),$$

и поэтому $C_{k'+l'} = C_{k+l}$.

Так как для $k, l, m \in \mathbb{Z}$

$$(C_k + C_l) + C_m = C_{(k+l)+m} = C_{k+(l+m)} = C_k + (C_l + C_m),$$
$$C_k + C_l = C_{k+l} = C_{l+k} = C_l + C_k,$$

то эта операция ассоциативна и коммутативна. Ясно, что C_0 является нейтральным элементом в $(\mathbb{Z}_n, +)$, а элемент C_{-k} является противоположным элементом для C_k.

Итак, $(\mathbb{Z}_n, +)$ — коммутативная группа, называемая *группой вычетов* по модулю n (операция сложения — это в точности операция сложения остатков при делении на n по модулю числа n: сначала надо сложить остатки как целые числа, а затем взять остаток от деления этой суммы на n). Мы отметили, что $|\mathbb{Z}_n| = n$.

В частности, имеем таблицы сложения для групп $\mathbb{Z}_2$ и $\mathbb{Z}_3$:

+	0	1
0	0	1
1	1	0

,

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

.

3) $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ относительно умножения являются группами (называемыми *мультипликативными группами* соответствующих *полей*).

4) $\mathbb{Q}_+ = \{q \in \mathbb{Q} \mid q > 0\}$, $\mathbb{R}_+ = \{r \in \mathbb{R} \mid r > 0\}$ с операциями умножения являются группами.

5) $G = \{1, -1\}$ с операцией умножения является группой.

Замечание 1.9.5. Множество $T(M)$ всех отображений $f\colon M \to M$ с операцией умножения (композицией) является *полугруппой*, но не является группой при $|M| > 1$ (существуют отображения $f\colon M \to M$, не являющиеся биекциями и, следовательно, не имеющие обратного отображения).

Упражнение 1.9.6.

1) Пусть $c \in \mathbb{R}$, $c > 0$, $G = \{r \in \mathbb{R} \mid -c < r < c\}$ $(= (-c, c))$. Покажите, что $(G, *)$ — группа, где

$$a * b = \frac{a + b}{1 + \frac{ab}{c^2}}$$

(сложение скоростей в специальной теории относительности).

2) Если G — группа, в которой $x^2 = 1$ для всех $x \in G$, то G — абелева группа.

Определение 1.9.7. Пусть G — группа, $a \in G$, $n \in \mathbb{Z}$ — целое число. Положим

$$a^n = \begin{cases} \underbrace{a \cdot a \cdot \ldots \cdot a}_{n}, & \text{если } n > 0, \\ e, & \text{если } n = 0, \\ \underbrace{a^{-1} \cdot a^{-1} \cdot \ldots \cdot a^{-1}}_{m=-n}, & \text{если } n < 0, \text{ где } m = -n > 0. \end{cases}$$

Замечание 1.9.8. Если $m > 0$, то $(a^{-1})^m = (a^m)^{-1}$. Действительно,

$$(\underbrace{a \ldots a}_{m})(\underbrace{a^{-1} \ldots a^{-1}}_{m}) = e = (\underbrace{a^{-1} \ldots a^{-1}}_{m})(\underbrace{a \ldots a}_{m}).$$

Теорема 1.9.9. *Пусть G — группа, $a \in G$, $m, n \in \mathbb{Z}$ — целые числа. Тогда*

$$a^m \cdot a^n = a^{m+n}.$$

Доказательство. Формально, мы должны рассмотреть $3 \times 3 = 9$ случаев.

Случай 1. $m > 0$, $n > 0$ (следовательно, $m + n > 0$). Тогда

$$a^m \cdot a^n = (\underbrace{a \ldots a}_{m}) \cdot (\underbrace{a \ldots a}_{n}) = \underbrace{a \ldots a}_{m+n} = a^{m+n}.$$

Случай 2. $m > 0$, $n < 0$ (поэтому $n' = -n > 0$). Тогда

$$a^m \cdot a^n = (\underbrace{a \ldots a}_{m}) \cdot (\underbrace{a^{-1} \ldots a^{-1}}_{n'=-n}) =$$

$$= \begin{cases} \underbrace{a \cdot a \cdot \ldots \cdot a}_{m-n'=m+n}, & \text{если } m > n' = -n \text{ (т. е. } m + n > 0), \\ e, & \text{если } m = n' = -n \text{ (т. е. } m + n = 0), \\ \underbrace{a^{-1} \ldots a^{-1}}_{n'-m=-n-m}, & \text{если } m < n' = -n \text{ (т. е. } m + n < 0) \end{cases} =$$

$$= a^{m+n}.$$

Аналогично разбираются остальные случаи: 3) $m < 0$, $n > 0$; 4) $m < 0$, $n < 0$; 5) $m = 0$, $n > 0$; 6) $m = 0$, $n = 0$; 7) $m = 0$, $n < 0$; 8) $m > 0$, $n = 0$; 9) $m < 0$, $n = 0$. □

Следствие 1.9.10. $(a^m)^n = a^{mn}$ *для всех* $m, n \in \mathbb{Z}$.

Рассмотрим целые степени элемента a группы G

$$\ldots, a^{-3}, a^{-2}, a^{-1}, a^0 = e, a, a^2, a^3, \ldots$$

Возможны два случая.

Случай 1. Все элементы в этом ряду различны (т. е. $a^k \neq a^l$ для всех целых чисел $k \neq l$). В этом случае будем говорить, что порядок элемента a бесконечный (обозначение: $\mathrm{O}(a) = \infty$).

Случай 2. В этом ряду $a^k = a^l$ для некоторых $k \neq l$. Пусть $k > l$. Тогда $a^{k-l} = e$, где $k - l > 0$, т. е. встретилась и натуральная степень элемента a, равная e. Рассмотрим множество $T = \{t \in \mathbb{Z} \mid t > 0,\ a^t = e\}$. Это непустое подмножество натуральных чисел. Следовательно, в T существует наименьший элемент n, который мы назовём *порядком элемента* a и обозначим через $\mathrm{O}(a)$.

Таким образом:

1) $a^n = e$, $n > 0$;

2) если $a^k = e$, $k > 0$, то $k \geqslant n$.

Пример 1.9.11. $G = \{1, -1\}$, $a = -1$. Тогда $a^1 = -1$, $a^2 = 1$, т. е. $\mathrm{O}(a) = 2$.

Лемма 1.9.12. *Если* $\mathrm{O}(a) = n < \infty$, *то*:

1) *все элементы* $e = a^0, a, a^2, \ldots, a^{n-1}$ *различны*;

2) *для любого* $k \in \mathbb{Z}$ *элемент* a^k *совпадает с одним из* $e, a, a^2, \ldots, a^{n-1}$.

Доказательство.

1) Следует из определения порядка элемента $\mathrm{O}(a)$.

2) Пусть $k \in \mathbb{Z}$. Тогда $k = nq + r$, где $0 \leqslant r < n$. Следовательно, $a^k = (a^n)^q a^r = ea^r = a^r$. □

Лемма 1.9.13. *Пусть* $\mathrm{O}(a) = n < \infty$. *Тогда* $a^k = e$ *в том и только в том случае, когда* $k = nq$.

Доказательство.

1) Если $k = nq$, то $a^k = (a^n)^q = e^q = e$.

2) Допустим противное, т. е. что $k = nq + r$, где $0 < r < n$. Тогда $a^k = (a^n)^q a^r = a^r \neq e$ (по лемме 1.9.12). Получили противоречие. □

Лемма 1.9.14. *Для непустого подмножества H группы G следующие условия эквивалентны:*

1) *H является группой относительно исходной операции в группе G;*

2) *подмножество H удовлетворяет следующим двум условиям:*

   а) *если $h_1, h_2 \in H$, то $h_1 h_2 \in H$;*

   б) *если $h \in H$, то $h^{-1} \in H$.*

Подмножество H группы G, удовлетворяющее эквивалентным условиям 1) и 2), называется *подгруппой* группы G.

Доказательство.

1) $\Longrightarrow$ 2). Если $h_1, h_2 \in H$, то, поскольку операция определена на H (т. е. не выводит из H), имеем $h_1 h_2 \in H$, т. е. 2а).

Если e' — нейтральный элемент группы H, то $e' \cdot e' = e'$. Умножая в группе G обе стороны равенства на $(e')^{-1}$, получаем $e' = e$ (здесь e — нейтральный элемент группы G).

Если $\underset{\sim}{h}^{-1}$ — обратный элемент для элемента $h \in H$, то

$$\underset{\sim}{h}^{-1} \cdot h = e' = e = h \cdot \underset{\sim}{h}^{-1},$$

т. е. $h^{-1} = \underset{\sim}{h}^{-1} \in H$ (условие 2б)).

2) $\Longrightarrow$ 1). Условие 2а) показывает, что операция определена на множестве H. Конечно, она ассоциативна. Далее, для $h \in H$ в силу 2б) $h^{-1} \in H$, и поэтому, в силу 2а), $e = h \cdot h^{-1} \in H$. Ясно, что e — нейтральный элемент в H, а h^{-1} — обратный элемент для h в H. Итак, H — группа относительно операции, индуцированной операцией группы G. □

Следствие 1.9.15. *Если G — группа, $\varnothing \neq F \subset H \subset G$, H — подгруппа группы G, F — подгруппа группы H, то F — подгруппа группы G.* □

Теорема 1.9.16. *Пусть G — группа, $\{H_i \mid i \in I\}$ — любое семейство подгрупп группы G. Тогда их пересечение $H = \bigcap\limits_{i\in I} H_i$ также является подгруппой.*

Доказательство.

1) Если $h_1, h_2 \in H = \bigcap\limits_{i\in I} H_i$, то $h_1, h_2 \in H_i$ для каждого i. Так как H_i — подгруппа, то $h_1h_2 \in H_i$ для каждого i, и поэтому $h_1h_2 \in \bigcap\limits_{i\in I} H_i = H$.

2) Если $h \in H = \bigcap\limits_{i\in I} H_i$, то $h \in H_i$ для каждого i. Так как H_i — подгруппа, то $h^{-1} \in H_i$ для каждого i, и поэтому $h^{-1} \in \bigcap\limits_{i\in I} H_i = H$.

Итак, $H = \bigcap\limits_{i\in I} H_i$ — подгруппа группы G. □

Примеры 1.9.17 (примеры подгрупп).

1) Чётные числа $2\mathbb{Z}$ — подгруппа в группе целых чисел $(\mathbb{Z}, +)$.

2) $\mathbb{Z} \subset (\mathbb{Q}, +)$, $\mathbb{Q} \subset (\mathbb{R}, +)$, $\mathbb{R} \subset (\mathbb{C}, +)$ — подгруппы.

3) В любой группе G имеем наименьшую подгруппу $H = \{e\}$ (и наибольшую подгруппу $H = G$).

Задача 1.9.18. Группа, имеющая лишь конечное число подгрупп, конечна.

Пусть a — элемент группы G. Рассмотрим в G следующее подмножество:

$$(a) = \{a^n \mid n \in \mathbb{Z}\}$$

(т. е. совокупность всех *целых* степеней элемента a).

Лемма 1.9.19.

1) *(a) является коммутативной подгруппой группы G;*

2) *$|(a)| = \mathrm{O}(a)$ (т. е. число элементов в подгруппе (a) равно порядку элемента a).*

Доказательство.

1) Для $m, n \in \mathbb{Z}$

$$a^m a^n = a^{m+n} \in (a): \quad (a^n)^{-1} = a^{-n} \in (a).$$

Таким образом, для (a) выполнены условия предыдущей леммы, т. е. $(a) = \{a^n \mid n \in \mathbb{Z}\}$ — подгруппа группы G. Так как

$$a^m a^n = a^{m+n} = a^n a^m,$$

то (a) — коммутативная группа.

2) Если $\mathrm{O}(a) = \infty$, то

$$(a) = \{\ldots, a^{-1}, e, a, \ldots\},$$

при этом в ряду целых степеней элемента a все элементы различны, т. е. $|(a)| = \infty$. Если же $\mathrm{O}(a) = n < \infty$, то, как мы отметили ранее,

$$(a) = \{e, a, \ldots, a^{n-1}\}$$

и

$$|(a)| = n = \mathrm{O}(a). \quad \square$$

Пример 1.9.20. Если $G = \mathbb{Z}$ и $a = 2$, то

$$(a) = \{na \mid a \in \mathbb{Z}\} = 2\mathbb{Z}$$

(все чётные числа).

Группа G называется *циклической*, если найдётся такой элемент $a \in G$, что $(a) = G$, т. е. все элементы группы G являются (целыми) степенями этого элемента a, называемого в этом случае *циклическим образующим группы* G. Если $\mathrm{O}(a) = n < \infty$, то $G = (a)$ — циклическая группа из n элементов; если же $\mathrm{O}(a) = \infty$, то $G = (a)$ — бесконечная (счётная!) циклическая группа.

Замечание 1.9.21. Любая циклическая группа $G = (a)$ является конечной или счётной коммутативной группой. Поэтому любая некоммутативная группа не является циклической и любая несчётная группа не является циклической группой.

Примеры 1.9.22.

1) $(\mathbb{Z}, +) = (1) = (-1)$ (это показывает, что циклических образующих может быть много!).

2) Группа действительных чисел $(\mathbb{R}, +)$ не является счётной, поэтому она не является циклической.

3) Показать, что счётная группа $(\mathbb{Q}, +)$ рациональных чисел не является циклической.

Пусть G и G' — группы. Напомним, что отображение $f: G \to G'$, для которого $f(ab) = f(a)f(b)$ для всех элементов $a, b \in G$, называется *гомоморфизмом*.

Пример 1.9.23. Пусть $G = \mathbb{R}^+ = \{r \in \mathbb{R} \mid r > 0\}$ с операцией умножения, $G' = (\mathbb{R}, +)$ с операцией сложения. Так как для отображения $\ln: \mathbb{R}^+ \to \mathbb{R}$ имеем $\ln(ab) = \ln(a) + \ln(b)$ для всех $a, b \in \mathbb{R}^+$, то $\ln$ — гомоморфизм групп.

Упражнение 1.9.24. Найти все гомоморфизмы

$$f: G \to G',$$

где $G = (a)$, $\mathrm{O}(a) = m$, $G' = (b)$, $\mathrm{O}(b) = n$ (в частности, для $m = 12$, $n = 15$).

Для гомоморфизмов $f: G \to G'$ определим:

$$\operatorname{Im} f = \{g' \in G' \mid g' = f(g) \text{ для } g \in G\}$$

(*образ* гомоморфизма f);

$$\operatorname{Ker} f = \{g \in G \mid f(g) = e'\},$$

где e' — нейтральный элемент группы G' (*ядро* гомоморфизма f).

Упражнение 1.9.25. В рассмотренных выше примерах найти образ и ядро гомоморфизма.

Задача 1.9.26. Доказать, что не существует сюръективного гомоморфизма $(\mathbb{Q}, +) \to (\mathbb{Z}, +)$.

Указание. В $(\mathbb{Q}, +)$ уравнение $nx = a$ имеет (и единственное) решение для любых $n \in \mathbb{N}$, $a \in \mathbb{Q}$.

Теорема 1.9.27 (свойства гомоморфизма групп). *Пусть G и G' — группы, e и e' соответственно — их нейтральные элементы, $f: G \to G'$ — гомоморфизм групп. Тогда:*

1) $f(e) = e'$;

2) $f(x^{-1}) = (f(x))^{-1}$ *для всех* $x \in G$;

3) $H' = \operatorname{Im} f$ *— подгруппа группы* G';

4) *если* $G = (a)$ *— циклическая группа, то* $\operatorname{Im} f = (f(a))$ *также циклическая группа*;

5) $f(g^{-1}hg) = (f(g))^{-1} f(h) f(g)$;

6) $\operatorname{Ker} f$ *— подгруппа группы* G, *при этом* $g^{-1}(\operatorname{Ker} f)g \subseteq \operatorname{Ker} f$ *для всех элементов* $g \in G$.

Доказательство.

1) Так как $u = f(e) = f(e^2) = f(e)f(e) = u^2$, то $u = e'$, т. е. $f(e) = e'$.

2) Так как $f(x^{-1})f(x) = f(x^{-1}x) = f(e) = e'$ и $f(x)f(x^{-1}) = f(xx^{-1}) = f(e) = e'$, то $f(x^{-1}) = (f(x))^{-1}$.

3) Если $h_1' = f(g_1)$ и $h_2' = f(g_2)$ — элементы из $\operatorname{Im} f$, где $g_1, g_2 \in G$, то

$$h_1' h_2' = f(g_1) f(g_2) = f(g_1 g_2) \in \operatorname{Im} f.$$

Если $h' = f(g) \in \operatorname{Im} f$, $g \in G$, то

$$(h')^{-1} = (f(g))^{-1} = f(g^{-1}) \in \operatorname{Im} f.$$

Итак, $\operatorname{Im} f$ — подгруппа группы G'.

4) Если $G = (a)$ и $h' \in \operatorname{Im} f$, $h' = f(g)$, $g \in G$, то $g = a^n$, $n \in \mathbb{Z}$, и поэтому

$$h' = f(g) = f(a^n) = (f(a))^n.$$

Итак, $\operatorname{Im} f = (f(a))$ — циклическая группа с образующим $f(a)$.

5) следует из 2).

6) Если $h_1, h_2 \in H = \operatorname{Ker} f$, то $f(h_1) = e'$, $f(h_2) = e'$. Поэтому $f(h_1 h_2) = f(h_1) f(h_2) = e' \cdot e' = e'$, т. е. $h_1 h_2 \in \operatorname{Ker} f$.

Если $h \in \operatorname{Ker} f$, то $f(h) = e'$, и поэтому $f(h^{-1}) = (f(h))^{-1} = (e')^{-1} = e'$, т. е. $h^{-1} \in \operatorname{Ker} f$. Таким образом, $\operatorname{Ker} f$ — подгруппа группы G.

Если $h \in H = \operatorname{Ker} f$, то $f(h) = e'$. Для любого элемента $g \in G$ имеем

$$f(g^{-1}hg) = f(g^{-1})f(h)f(g) = f(g)^{-1}e'f(g) = e'.$$

Таким образом, $g^{-1} \operatorname{Ker} f g \subseteq \operatorname{Ker} f$ для всех элементов $g \in G$. □

Лемма 1.9.28. *Если G, G', G'' — группы, $f\colon G \to G'$, $g\colon G' \to G''$ — гомоморфизмы, то $gf\colon G \to G''$ — гомоморфизм.*

Доказательство. Пусть $a, b \in G$. Тогда

$$(gf)(ab) = g[f(ab)] = g[f(a)f(b)] = \\ = [g(f(a))]\,[g(f(b))] = [(gf)(a)]\,[(gf)(b)]. \quad \square$$

Лемма 1.9.29. *Пусть G, G' — группы, $f\colon G \to G'$ — гомоморфизм групп. Тогда:*

1) *f — инъекция в том и только в том случае, когда $\operatorname{Ker} f = \{e\}$;*

2) *f — биекция в том и только в том случае, когда $\operatorname{Ker} f = \{e\}$, $\operatorname{Im} f = G'$.*

Доказательство. Достаточно доказать 1). Если f — инъекция, то, учитывая равенство $f(e) = e'$, видим, что $\operatorname{Ker} f = \{e\}$. Пусть теперь $\operatorname{Ker} f = \{e\}$. Если $f(a) = f(b)$ для $a, b \in G$, то $f(a^{-1}b) = f(a^{-1})f(b) = [f(a)]^{-1}f(b) = e'$, т. е. $a^{-1}b \in \operatorname{Ker} f = \{e\}$. Поэтому $a^{-1}b = e$, т. е. $a = b$. Итак, f — инъекция. □

Определение 1.9.30. Пусть G, G' — группы. Отображение $f\colon G \to G'$ назовём *изоморфизмом групп*, если:

1) f — гомоморфизм;

2) f — биекция.

Группы G и G' называются *изоморфными*, если существует какой-либо изоморфизм $f\colon G \to G'$ (обозначение $G \cong G'$).

Примеры 1.9.31. Следующие отображения — изоморфизмы групп:

1) $(\mathbb{R}^+, \cdot) = (\{r \in \mathbb{R} \mid r > 0\}, \cdot) \xrightarrow{\ln} (\mathbb{R}, +)$;

2) $\mathbb{Z} \to 2\mathbb{Z}$, $n \mapsto 2n$.

Лемма 1.9.32. *Если G, G', G'' — группы, $f\colon G \to G'$, $g\colon G' \to G''$ — изоморфизмы, то gf и f^{-1} — изоморфизмы* (см. лемму 1.2.2).

Доказательство.

а) По лемме 1.9.29, gf — гомоморфизм. Так как gf и биекция, то gf — изоморфизм.

б) Мы знаем, что f^{-1} — биекция. Пусть $w, z \in G'$. Тогда $w = f(x)$, $z = f(y)$, где $x, y \in G$. Следовательно, $wz = f(x)f(y) = f(xy)$. Поэтому $f^{-1}(wz) = f^{-1}(f(xy)) = xy = f^{-1}(w)f^{-1}(z)$, т. е. f^{-1} — гомоморфизм. Итак, f^{-1} — изоморфизм. □

Следствие 1.9.33. *Отношение $G \cong G'$ является отношением эквивалентности на классе групп.*

Замечание 1.9.34. Изоморфные группы обладают одинаковыми «алгебраическими» свойствами.

Пример 1.9.35. Если группы G и G' изоморфны и G — коммутативная группа, то G' — также коммутативная группа. Действительно, пусть $f\colon G \to G'$ — некоторый изоморфизм. Если $z, w \in G'$, то $z = f(a)$, $w = f(b)$ для некоторых $a, b \in G$. Тогда

$$zw = f(a)f(b) = f(ab) = f(ba) = f(b)f(a) = wz. \quad \square$$

1.10. Кольца

Множество R с двумя бинарными операциями (сложением $+$ и умножением $\cdot$) называется *ассоциативным кольцом с единицей*, если:

1) относительно сложения $(R, +)$ — абелева (т. е. коммутативная) группа;

2) умножение — ассоциативная операция, и существует нейтральный элемент 1 (т. е. $1 \cdot r = r = r \cdot 1$ для всех $r \in R$), называемый единицей;

3) сложение и умножение связаны *законами дистрибутивности*

$$(a + b)c = ac + bc, \quad c(a + b) = ca + cb$$

для всех $a, b, c \in R$.

Если операция умножения коммутативна, то кольцо $(R, +, \cdot)$ называется *коммутативным* кольцом. Коммутативные кольца являются одним из главных объектов изучения в коммутативной алгебре и алгебраической геометрии.

Замечания 1.10.1.

1) Исследуются и неассоциативные кольца. Например, если вместо ассоциативности 2) умножение удовлетворяет *тождеству Якоби*

$$a(bc) + b(ca) + c(ab) = 0$$

для всех $a, b, c \in R$ и

$$ab = -ba$$

для всех $a, b \in R$, то такое кольцо называется *кольцом Ли*.

2) Рассматриваются также и ассоциативные кольца без единицы. Например, чётные числа $R = 2\mathbb{Z}$ являются ассоциативным коммутативным кольцом без единицы.

Примеры 1.10.2 (примеры ассоциативных колец).

1) *Кольцо* $(\mathbb{Z}, +, \cdot)$ *целых чисел*; поля $\mathbb{Q}$, $\mathbb{R}$.

2) *Кольцо непрерывных вещественных функций* $C[0, 1]$ *на отрезке* $[0, 1]$ (для $f, g \in C[0, 1]$, $x \in [0, 1]$: $(f+g)(x) = f(x)+g(x)$, $(fg)(x) = f(x)g(x)$).

3) Кольцо многочленов $\mathbb{R}[x]$ с действительными коэффициентами.

4) *Кольцо вычетов* $(\mathbb{Z}_n, +, \cdot)$ по модулю n.

Мы уже убедились, что *группа вычетов*

$$(\mathbb{Z}_n, +) = \{C_0, C_1, \ldots, C_{n-1}\}, \quad C_k = k + n\mathbb{Z},$$

по модулю n с операцией сложения

$$C_k + C_l = C_{k+l} = C_r, \quad \text{где} \quad k + l = nq + r, \ 0 \leqslant r \leqslant n - 1,$$

является коммутативной группой (см. пример 1.9.4, 2)).

Определим операцию *умножения*, полагая

$$C_k \cdot C_l = C_{kl} = C_s, \quad \text{где} \quad kl = n\tilde{q} + s, \ 0 \leqslant s \leqslant n-1.$$

Проверим корректность этой операции. Если $C_k = C_{k'}$, $C_l = C_{l'}$, то $k' = k + nu$, $l' = l + nv$, $k' \cdot l' = kl + n(kv + ul + nuv)$, и поэтому $C_{k'l'} = C_{kl}$.

Так как

$$(C_k C_l)C_m = C_{(kl)m} = C_{k(lm)} = C_k(C_l C_m),$$
$$C_k C_l = C_{kl} = C_{lk} = C_l C_k,$$
$$C_1 C_k = C_k = C_k C_1,$$
$$(C_k + C_l)C_m = C_{(k+l)m} = C_{km+lm} = C_k C_m = C_l C_m,$$

то $(\mathbb{Z}_n, +, \cdot)$ является ассоциативным коммутативным кольцом с единицей C_1 (называемым *кольцом вычетов по модулю* n).

Свойства колец $(R, +, \cdot)$

1. Так как $(R, +)$ — абелева группа, то: существует, и единственный, нейтральный элемент относительно сложения 0; для любого $a \in R$ существует, и единственный, противоположный элемент $-a$ (т. е. $a + (-a) = 0$); уравнение $x + b = a$ имеет, и единственное, решение $x = a - b = a + (-b)$.

2. Справедлив обобщённый закон ассоциативности для умножения, т. е. результат произведения для n сомножителей не зависит от расстановки скобок; единичный элемент 1 — единственный нейтральный элемент (см. теорему 1.3.2).

3. Проводя индукцию по n, убеждаемся в том, что

$$(a_1 + \ldots + a_n)b = a_1 b + \ldots + a_n b;$$
$$b(a_1 + \ldots + a_n) = ba_1 + \ldots + ba_n.$$

4. Так как $a0 = a(0+0) = a0 + a0$, то $a0 = 0$. Аналогично, $0a = 0$.

5. Так как $ab + (-a)b = (a + (-a))b = 0b = 0$, то $(-a)b = -ab$. Аналогично, $a(-b) = -ab$. Поэтому $(-a)(-b) = -(a(-b)) = -(-ab) = ab$.

6. $(a - b)c = (a + (-b))c = ac + (-b)c = ac - bc$, $c(a - b) = c(a + (-b)) = ca + c(-b) = ca - cb$, т. е. дистрибутивность для разности.

Лемма 1.10.3 (*бином Ньютона*). *Пусть R — кольцо с 1, $n \in \mathbb{N}$, $a, b, a_1, a_2, \ldots, a_s \in R$. Тогда:*

1) *если $ab = ba$, то*

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}, \quad \text{где} \quad C_n^k = \frac{n!}{k!(n-k)!}, \quad C_n^0 = 1;$$

2) *если $a_i a_j = a_j a_i$ для всех i, j, то*

$$(a_1 + a_2 + \ldots + a_s)^n = \sum \frac{n!}{(i_1!)\ldots(i_s!)} a_1^{i_1} \ldots a_s^{i_s},$$

где суммирование происходит по всем s-строчкам $(i_1, i_2, \ldots, i_s)$ таким, что $i_1 + i_2 + \ldots + i_s = n$.

Доказательство.

1) Индукция по n с учётом равенства $C_n^k + C_n^{k-1} = C_{n+1}^{k+1}$ для $k < n$ и применением перестановочности элементов a и b и закона дистрибутивности.

2) Индукция по s; $s = 2$ — пункт 1); если утверждение верно для s, то по 1):

$$(a_1 + \ldots + a_s + a_{s+1})^n = ((a_1 + \ldots + a_s) + a_{s+1})^n =$$

$$= \sum_{k=0}^{n} C_n^k (a_1 + \ldots + a_s)^k a_{s+1}^{n-k} = \sum_{k+j=n} \frac{n!}{k!j!} (a_1 + \ldots + a_s)^k a_{s+1}^j =$$

$$= \sum_{k+j=n} \frac{n!}{k!j!} \sum_{\substack{(i_1,\ldots,i_s) \\ i_1+\ldots+i_s=k}} \frac{k!}{(i_1!)\ldots(i_s!)} a_1^{i_1} a_2^{i_2} \ldots a_s^{i_s} a_{s+1}^j =$$

$$= \sum_{\substack{(i_1,\ldots,i_{s+1}) \\ i_1+\ldots+i_{s+1}=n}} \frac{n!}{(i_1!)\ldots(i_{s+1}!)} a_1^{i_1} a_2^{i_2} \ldots a_{s+1}^{i_{s+1}} \quad (j = i_{s+1}). \quad \square$$

Определение 1.10.4. Подмножество S кольца R называется *подкольцом*, если:

а) S — подгруппа относительно сложения в группе $(R, +)$;

б) для $a, b \in S$ имеем $ab \in S$;

в) для кольца R с 1 предполагается, что $1 \in S$.

Примеры 1.10.5 (примеры подколец). $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Задача 1.10.6. Описать все подкольца в кольце вычетов $\mathbb{Z}_n$ по модулю n.

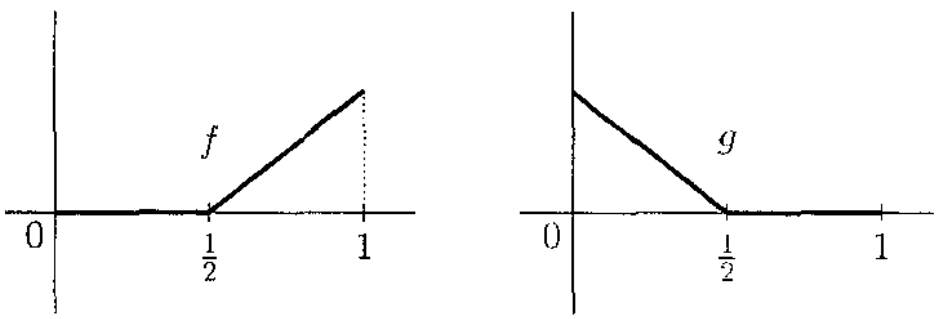
Замечание 1.10.7. В кольце $\mathbb{Z}_{10}$ элементы, кратные 5, образуют кольцо с 1, не являющееся подкольцом в $\mathbb{Z}_{10}$ (у этих колец различные единичные элементы).

Определение 1.10.8. Если R — кольцо, $a, b \in R$ и $a \neq 0$, $b \neq 0$, $ab = 0$, то элемент a называется *левым делителем нуля* в R, элемент b называется *правым делителем нуля* в R.

Замечание 1.10.9. В коммутативных кольцах, естественно, нет различий между левыми и правыми делителями нуля.

Пример 1.10.10. В $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ нет делителей нуля.

Пример 1.10.11. Кольцо непрерывных функций $C[0,1]$ имеет делители нуля. Действительно, если



то $f \neq 0$, $g \neq 0$, $fg = 0$.

Пример 1.10.12. Если $n = kl$, $1 < k, l < n$, то $C_k \neq C_0$, $C_l \neq C_0$, но $C_k C_l = C_0$, т. е. кольцо вычетов $\mathbb{Z}_n$ по составному числу n имеет делители нуля.

Лемма 1.10.13. *Если в кольце R нет (левых) делителей нуля, то из $ab = ac$, где $0 \neq a \in R$, $b, c \in R$, следует, что $b = c$ (т. е. возможность сокращать на ненулевой элемент слева, если нет левых делителей нуля; и справа, если нет правых делителей нуля).*

Доказательство. Если $ab = ac$, то $a(b - c) = 0$. Так как a не является левым делителем нуля, то $b - c = 0$, т. е. $b = c$. □

Определение 1.10.14. Элемент $x \in R$ называется *нильпотентным*, если $x^n = 0$ для некоторого $0 < n \in \mathbb{N}$. Наименьшее такое натуральное число n называется *степенью нильпотентности элемента*.

Ясно, что нильпотентный элемент является делителем нуля (если $n > 1$, то $x \cdot x^{n-1} = 0$, $x^{n-1} \neq 0$). Обратное утверждение неверно (в $\mathbb{Z}_6$ нет нильпотентных элементов, однако 2, 3, 4 — ненулевые делители нуля).

Упражнение 1.10.15. Кольцо $\mathbb{Z}_n$ содержит нильпотентные элементы тогда и только тогда, когда n делится на m^2, где $m \in \mathbb{N}$, $m \neq 1$.

Определение 1.10.16. Элемент x кольца R называется *идемпотентом*, если $x^2 = x$. Ясно, что $0^2 = 0$, $1^2 = 1$. Если $x^2 = x$ и $x \neq 0$, $x \neq 1$, то $x(x-1) = x^2 - x = 0$, и поэтому нетривиальные идемпотенты являются делителями нуля.

Через $U(R)$ обозначим множество обратимых элементов ассоциативного кольца R, т. е. тех $r \in R$, для которых существует обратный элемент $s = r^{-1}$ (т. е. $rr^{-1} = 1 = r^{-1}r$).

Лемма 1.10.17. *$U(R)$ является группой относительно операции умножения.*

Доказательство.
1) Если $r, s \in U(R)$, то $rs \in U(R)$, поскольку $(rs)^{-1} = s^{-1}r^{-1}$.
2) $1 \in U(R)$.
3) Если $r \in U(R)$, то $(r^{-1})^{-1} = r$, т. е. $r^{-1} \in U(R)$. □

Пример 1.10.18. $U(\mathbb{Z}) = \{1, -1\}$, $U(\mathbb{Q}) = \mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, $U(\mathbb{R}) = \mathbb{R}^*$.

Пример 1.10.19. $U(C[0,1]) = \{f \in C[0,1] \mid f(x) \neq 0 \ \forall x \in [0,1]\}$.

Пример 1.10.20. Пусть $\mathbb{Z}_m = \{C_0, C_1, \ldots, C_{m-1}\}$, $C_k = k + m\mathbb{Z}$, — кольцо вычетов по модулю m. Отметим, что $k + m\mathbb{Z} \in U(\mathbb{Z}_m)$, $k \in \mathbb{Z}$, тогда и только тогда, когда $(k+m\mathbb{Z})(l+m\mathbb{Z}) = 1+m\mathbb{Z}$ для некоторого $l \in \mathbb{Z}$, т. е. $kl + m\mathbb{Z} = 1 + m\mathbb{Z}$, что означает $kl = 1 + mq$, $q \in \mathbb{Z}$, т. е. $(k, m) = 1$.

Итак, $|U(\mathbb{Z}_m)| = \varphi(m)$, где $\varphi(m)$ — число натуральных чисел $1 \leqslant k < m$, не имеющих нетривиальных общих делителей с числом m (*функция Эйлера*). В частности, $\varphi(1) = 1$, $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$, $\varphi(p) = p - 1$ для простого числа p. Более того, если $p \in \mathbb{N}$, то $\varphi(p) = p - 1$ тогда и только тогда, когда p — простое число.

Задача 1.10.21. Докажите, что группа $U(\mathbb{Z}_n)$ циклическая тогда и только тогда, когда $n \in \{2, 4, p^\alpha, 2p^\alpha\}$, где p — нечётное простое число.

1.11. Поля

Определение 1.11.1. Ассоциативное коммутативное кольцо K с 1, в котором для любого ненулевого элемента $a \in K$ существует обратный элемент a^{-1}, называется *полем*.

Лемма 1.11.2. *Если K — поле, то уравнение $ax = b$, где $a \neq 0$, имеет одно и только одно решение (именно, $a^{-1}b$).*

Доказательство. Если $ax = b$, то $x = a^{-1}ax = a^{-1}b$. Если $x = a^{-1}b$, то $ax = a(a^{-1}b) = b$. □

Теорема 1.11.3. *В поле K нет делителей нуля.*

Доказательство. Допустим, что $a, b \in K$, $a \neq 0$, $b \neq 0$ и $ab = 0$. Тогда $b = a^{-1}(ab) = a^{-1}0 = 0$, противоречие. □

Замечание 1.11.4. Обратное утверждение неверно. В кольце $\mathbb{Z}$ целых чисел нет делителей нуля, но оно не является полем.

Пример 1.11.5. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Z}_2$ — поля.

Теорема 1.11.6. *Кольцо вычетов $\mathbb{Z}_n$ является полем (полем вычетов) тогда и только тогда, когда $n = p$ является простым числом.*

Доказательство. Если $n = p$ — простое число, то $\mathbb{Z}_p$ — кольцо без делителей нуля (действительно, если $C_k C_l = C_0$, $C_k \neq C_0$, $C_l \neq C_0$, то $kl = pq$, но k и l не делятся на p, что приводит к противоречию). Доказательство завершает следующая лемма.

Лемма 1.11.7. Конечное *коммутативное кольцо без делителей нуля является полем.*

Доказательство. Пусть $R = \{r_0 = 0, r_1 = 1, \ldots, r_{n-1}\}$ — кольцо из n элементов без делителей нуля. Для $r_k \neq 0$, $1 \leqslant k \leqslant n-1$, все произведения $r_k r_1, \ldots, r_k r_{n-1}$ различны, поскольку r_k не является делителем нуля. Следовательно, найдётся i, для которого $r_k r_i = 1$, т. е. $r_i = r_k^{-1}$. □

Лемма 1.11.8. *Пересечение* $\bigcap\limits_{i \in I} K_i$ *любого семейства подполей* K_i, $i \in I$, *поля* K *является подполем.* □

Упражнение 1.11.9. Через $\mathbb{Q}[\sqrt{2}]$ обозначим наименьшее подполе в $\mathbb{R}$, содержащее поле $\mathbb{Q}$ и элемент $\sqrt{2}$ (существующее по лемме 1.11.8). Покажите, что поля $\mathbb{Q}[\sqrt{2}]$ и $\mathbb{Q}[\sqrt{3}]$ не являются изоморфными.

Определение 1.11.10. Рассмотрим поле K как абелеву группу $(K, +)$ относительно сложения, пусть $\mathrm{O}(1)$ — порядок элемента 1 в этой группе. Если $\mathrm{O}(1) = \infty$, то говорят, что характеристика $\operatorname{char} K$ поля K равна 0 (т. е. для любых целых чисел $k, l \in \mathbb{Z}$ из $k \neq l$ следует, что $k \cdot 1 \neq l \cdot 1$ в K). Если $\mathrm{O}(1) = p < \infty$, то полагают $\operatorname{char} K = p$ и говорят, что K — поле конечной характеристики p (т. е. p — наименьшее натуральное число, для которого $p \cdot 1 = \underbrace{1 + \ldots + 1}_{p} = 0$).

Примеры 1.11.11.

1) $\operatorname{char} \mathbb{Q} = 0$, $\operatorname{char} \mathbb{R} = 0$.

2) $\operatorname{char} \mathbb{Z}_p = p$ (для простого числа p).

Теорема 1.11.12. *Если* K *— поле и* $\operatorname{char} K = p > 0$, *то* p *— простое число.*

Доказательство. Допустим противное, т. е. что $p = st$, где $1 < s, t < p$. Тогда

$$(s \cdot 1)(t \cdot 1) = (\underbrace{1 + \ldots + 1}_{s})(\underbrace{1 + \ldots + 1}_{t}) = st \cdot 1 = p \cdot 1 = 0,$$

но $s \cdot 1 \neq 0$, $t \cdot 1 \neq 0$ в поле K, что противоречит отсутствию делителей нуля в поле. □

1.12. Идеалы и гомоморфизмы колец

Определение 1.12.1. Пусть R — кольцо. Подмножество $\varnothing \neq I \subset R$ называется *левым идеалом* кольца R, если:

1) I — подгруппа аддитивной группы $(R, +)$ кольца R;

2) $rI \subseteq I$ для любого элемента $r \in R$ (т. е. $ri \in I$ для всех $i \in I$).

Аналогично определяется *правый идеал*: вместо 2) условие

2') $Ir \subseteq I$ для любого элемента $r \in R$ (т. е. $ir \in I$ для всех $i \in I$).

Если подмножество I в кольце R является и левым и правым идеалом, то I называется *двусторонним идеалом* кольца R (т. е. I — подгруппа в $(R, +)$, $rI \subseteq I$, $Ir \subseteq I$ для всех $r \in R$). Для двустороннего идеала I кольца R будем использовать обозначение $I \lhd R$.

Примеры 1.12.2.

1) $\{0\}$ и R — идеалы кольца R.

2) $\mathbb{Z}n \lhd \mathbb{Z}$ для любого $n \in \mathbb{Z}$.

3) $I_a = \{f \in C[0,1] \mid f(a) = 0\} \lhd C[0,1]$ для любого $a \in [0,1]$.

4) Если R — коммутативное кольцо, $a \in R$, то подмножество

$$Ra = \{ra \mid r \in R\}$$

является идеалом кольца R, называемым *главным идеалом*, порождённым элементом $a \in R$.

Упражнение 1.12.3. Покажите, что в кольце целых чисел $(\mathbb{Z}, +, \cdot)$ каждый идеал имеет вид $\mathbb{Z}n$, $n \in \mathbb{Z}$, т. е. каждый идеал является главным (такие коммутативные кольца называются *кольцами главных идеалов*).

Пусть R и R' — кольца. Отображение $f\colon R \to R'$ называется *гомоморфизмом колец*, если $f(a+b) = f(a) + f(b)$ и $f(ab) = f(a)f(b)$ для всех $a, b \in R$.

Через $\operatorname{Im} f$ обозначим *образ гомоморфизма* f, т. е.

$$\operatorname{Im} f = \{f(r) \in R' \mid r \in R\};$$

через $\operatorname{Ker} f$ — *ядро гомоморфизма* f, т. е.

$$\operatorname{Ker} f = \{a \in R \mid f(a) = 0\}.$$

Если гомоморфизм f является биекцией, то f называется *изоморфизмом колец*.

Отметим ряд свойств гомоморфизмов колец $f\colon R \to R'$.

1. Так как f — гомоморфизм абелевых групп $(R, +)$, $(R', +)$, то $f(0) = 0'$, $f(-a) = -f(a)$.

2. Если $R \ni 1$, $R' \ni 1'$ и $\operatorname{Im} f = R'$, то $f(1) = 1'$, $f(a^{-1}) = f(a)^{-1}$ для обратимого элемента a. Действительно, если $a' \in R'$, то $a' = f(a)$, $a \in R$. Тогда

$$f(1)a' = f(1)f(a) = f(1 \cdot a) = f(a) = a',$$
$$a'f(1) = f(a)f(1) = f(a \cdot 1) = f(a) = a',$$

т. е. $f(1) = 1'$;

$$f(a^{-1})f(a) = f(a^{-1}a) = f(1) = 1',$$
$$f(a)f(a^{-1}) = f(aa^{-1}) = f(1) = 1',$$

т. е. $f(a^{-1}) = f(a)^{-1}$.

Это утверждение может не быть верным, если $\operatorname{Im} f \neq R'$.

3. Если $f\colon R \to R'$ — гомоморфизм колец, то $\operatorname{Ker} f$ — двусторонний идеал кольца R.

Доказательство. Так как $f\colon (R, +) \to (R', +)$ — гомоморфизм групп, то $\operatorname{Ker} f$ — подгруппа в $(R, +)$.

Если $a \in \operatorname{Ker} f$, т. е. $f(a) = 0$, $r, s \in R$, то

$$f(ra) = f(r)f(a) = f(r) \cdot 0 = 0,$$
$$f(as) = f(a)f(s) = 0 \cdot f(s) = 0,$$

итак, $ra \in \operatorname{Ker} f$, $as \in \operatorname{Ker} f$, т. е. $\operatorname{Ker} f \lhd R$. □

4. Гомоморфизм колец $f\colon R \to R'$ является изоморфизмом тогда и только тогда, когда $\operatorname{Ker} f = \{0\}$ и $\operatorname{Im} f = R'$ (следует вспомнить критерий изоморфизма для гомоморфизмов групп, см. лемму 1.9.29).

Ясно, что изоморфные кольца обладают одинаковыми кольцевыми свойствами. Например, если $f\colon R \to R'$ — изоморфизм колец, R — поле, то R' также поле.

Упражнение 1.12.4. Если R — коммутативное кольцо, то R — поле тогда и только тогда, когда в R нет идеалов, отличных от $\{0\}$ и R.

Упражнение 1.12.5. Отображение $\mathbb{Z}_3 \to \mathbb{Z}_6$, при котором $k + 3\mathbb{Z} \mapsto 4k + 6\mathbb{Z}$, является инъективным гомоморфизмом колец.

1.13. Кольцо многочленов от одной переменной

Пусть K — произвольное поле.

Под *многочленом* (ненулевым) от одной переменной x с коэффициентами из поля K будем понимать формальное выражение вида

$$f(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + a_n x^n$$

(иногда удобнее записывать эту сумму *одночленов* $a_i x^i$ в другом порядке: $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$), $a_i \in K$, $a_n \neq 0$ — *старший коэффициент* ($a_n x^n$ — *старший член* многочлена $f(x)$), a_0 — *свободный член*, $n = \deg f(x)$ — *степень* ненулевого многочлена $f(x)$ (*нулевой многочлен* — это $f(x) = a_0 = 0$).

Можно было вместо формальных выражений рассматривать счётные последовательности

$$(a_0, a_1, \ldots, a_n, 0, 0, \ldots), \quad a_i \in K,$$

в которых *почти все* a_i (т. е. все, кроме конечного числа) равны нулю (нулевой многочлен — это последовательность, в которой все компоненты равны нулю).

Два многочлена $f(x)$ и $g(x)$ называются *равными*, если равны соответствующие коэффициенты при каждой степени x^k переменной x.

Через $K[x]$ обозначим множество всех многочленов $f(x)$ с коэффициентами из поля K.

На множестве $K[x]$ введём операции сложения и умножения, для

$$f(x)=\sum_{i=0}^{n}a_ix^i,\quad g(x)=\sum_{i=0}^{s}b_ix^i$$

полагая

$$f(x)+g(x)=\sum_{i\geqslant 0}d_ix^i,\quad f(x)g(x)=\sum_{i\geqslant 0}t_ix^i,$$

где

$$d_i=a_i+b_i,\quad t_i=\sum_{\substack{k+l=i\\0\leqslant k,l\leqslant i}}a_kb_l.$$

Теорема 1.13.1. *Множество $K[x]$ с операциями сложения и умножения — коммутативное ассоциативное кольцо с единицей.*

Доказательство.

1) Так как при сложении складываются коэффициенты при одной степени x^i, т. е. $d_i=a_i+b_i$, то ясно, что $K[x]$ с операцией сложения — коммутативная группа.

2) Учитывая определение коэффициента

$$t_i=\sum_{\substack{k+l=i\\0\leqslant k,l\leqslant i}}a_kb_l,$$

заключаем, что операция умножения коммутативна.

Пусть теперь

$$h(x)=\sum_{i\geqslant 0}c_ix^i.$$

Тогда, подсчитывая коэффициенты при степени x^i в $(f(x)g(x))h(x)$ и в $f(x)(g(x)h(x))$, видим, что

$$\sum_{u+m=i}\Bigg(\sum_{k+l=u}a_kb_l\Bigg)c_m=\sum_{k+l+m=i}a_kb_lc_m=\sum_{k+v=i}a_k\Bigg(\sum_{l+m=v}b_lc_m\Bigg).$$

Итак, мы проверили ассоциативность умножения многочленов.

Ясно, что $f(x)=1$ (т. е. $a_0=1$) является нейтральным элементом для операции умножения.

3) Подсчитывая коэффициенты при степени x^i в $(f(x)+g(x))h(x)$ и $f(x)h(x)+g(x)h(x)$, видим, что

$$\sum_{k+l=i}(a_k+b_k)c_l=\sum_{k+l=i}a_kc_l+\sum_{k+l=i}b_kc_l,$$

т. е. установлен закон дистрибутивности в $K[x]$. □

Замечание 1.13.2. Отображение $K\to K[x]$, для которого

$$a\mapsto f(x)=a_0=a,$$

является инъективным гомоморфизмом колец (т. е. получили вложение поля K в кольцо многочленов $K[x]$).

Лемма 1.13.3. *Пусть K — поле, $f(x),g(x)\in K[x]$, $0\neq f(x)$, $0\neq g(x)$. Тогда*

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

б) $\deg(f(x)g(x))=\deg f(x)+\deg g(x)$.

Доказательство.

а) Если $i>\max(\deg f(x),\deg g(x))$, то $c_i=a_i+b_i=0$.

б) Если $\deg f(x)=n$, $\deg g(x)=s$ и $i>n+s$, то

$$d_i=\sum_{\substack{k+l=i\\0\leqslant k,l\leqslant i}}a_kb_l=0.$$

При этом $d_{n+s}=a_nb_s\neq 0$ (поскольку $a_n\neq 0$, $b_s\neq 0$ и в поле K нет делителей нуля). Итак, $d_{n+s}=a_nb_s\neq 0$ — старший коэффициент многочлена $f(x)g(x)$ — является произведением старших коэффициентов многочленов $f(x)$ и $g(x)$. Таким образом, $\deg(f(x)g(x))=n+s=\deg f(x)+\deg g(x)$. □

Следствие 1.13.4. *Пусть K — поле. В кольце многочленов $K[x]$ нет делителей нуля.*

Доказательство. Как мы видели, если $f(x)\neq 0$, $\deg f(x)=n$, $a_n\neq 0$ — старший коэффициент многочлена $f(x)$, $g(x)\neq 0$, $\deg f(x)=s$, $b_s\neq 0$ — старший коэффициент многочлена $g(x)$, то $a_nb_s\neq 0$ — старший коэффициент многочлена $f(x)g(x)$, т. е. $f(x)g(x)\neq 0$. □

Следствие 1.13.5. *Пусть K — поле. В кольце $K[x]$ (как в любом кольце без делителей нуля) можно сокращать на ненулевой многочлен, т. е. из $f(x)g(x) = f(x)h(x)$, $f(x) \neq 0$, следует, что $g(x) = h(x)$.*

Следствие 1.13.6. *Пусть K — поле. $U(K[x]) = K \setminus \{0\}$ (здесь $U(R)$ — группа обратимых элементов кольца R).*

Доказательство. Если $0 \neq a \in K$, то $a^{-1} \in K \subseteq K[x]$, т. е. $a \in U(K[x])$.

Если $f(x)g(x) = 1$, то $f(x) \neq 0$, $g(x) \neq 0$, $\deg f(x) + \deg g(x) = 0$, и поэтому $\deg f(x) = 0 = \deg g(x)$, т. е. $f(x) = a_0 \neq 0$, $a_0 \in K$. □

Упражнение 1.13.7. Произведение двух линейных многочленов

$$(ax + b)(cx + d) = acx^2 + (ad + bc)x + bd,$$

требующее четырёх умножений (ac, ad, bc, bd) и одного сложения ($ad + bc$), может быть вычислено с помощью трёх умножений и четырёх сложений и вычитаний:

$$ac, \quad bd, \quad u = (a + b)(c + d), \quad ad + bc = u - ac - bd.$$

А. А. Карацуба использовал это соображение для построения быстрых алгоритмов умножения чисел и многочленов.

Теорема 1.13.8 (алгоритм деления с остатком в кольце многочленов). *Для любых многочленов $f(x), g(x) \in K[x]$, $g(x) \neq 0$, существуют (и притом единственные) многочлены $q(x), r(x) \in K[x]$ такие, что:*

1) $f(x) = g(x)q(x) + r(x)$;

2) *либо* $r(x) = 0$, *либо* $r(x) \neq 0$, $\deg r(x) < \deg g(x)$.

Доказательство-алгоритм (деление многочленов столбиком). Пусть

$$f(x) = a_n x^n + \ldots + a_1 x + a_0,$$
$$g(x) = b_s x^s + \ldots + b_1 x + b_0, \quad b_s \neq 0.$$

Если $n < s$, то утверждение 1) очевидно:

$$f(x) = g(x) \cdot 0 + f(x).$$

Пусть $n \geqslant s$. Тогда:

$$f(x) - \frac{a_n}{b_s}x^{n-s}g(x) = f_1(x) = a_{1,n_1}x^{n_1} + \ldots, \quad s \leqslant n_1 < n,$$

$$f_1(x) - \frac{a_{1,n_1}}{b_s}x^{n_1-s}g(x) = f_2(x) = a_{2,n_1}x^{n_2} + \ldots, \quad s \leqslant n_2 < n_1,$$

$$\ldots$$

$$f_{k-2}(x) - \frac{a_{k-2,n_{k-2}}}{b_s}x^{n_{k-2}-s}g(x) = f_{k-1}(x) = a_{k-1,n_{k-1}}x^{n_{k-1}} + \ldots, \quad s \leqslant n_{k-1} < n_{k-2},$$

$$f_{k-1}(x) - \frac{a_{k-1,n_{k-1}}}{b_s}x^{n_{k-1}-s}g(x) = f_k(x) = a_{k,n_k}x^{n_k} + \ldots,$$

$$\begin{cases} f_k(x) = 0 \text{ или} \\ n_k < s, \ n_k < n_{k-1}. \end{cases}$$

Складывая все эти равенства и сокращая, получаем

$$f(x) - \left(\frac{a_n}{b_s}x^{n-s} + \ldots + \frac{a_{k-1,n_{k-1}}}{b_s}x^{n_{k-1}-s}\right)g(x) = f_k(x),$$

т. е.

$$f(x) = q(x)g(x) + r(x),$$

где

$$q(x) = \frac{a_n}{b_s}x^{n-s} + \ldots + \frac{a_{k-1,n_{k-1}}}{b_s}x^{n_{k-1}-s},$$

$$r(x) = f_k(x), \quad r(x) = 0 \text{ или } \deg(r(x)) < s = \deg g(x).$$

Если $f(x) = g(x)q(x)+r(x) = g(x)q'(x)+r'(x)$, при этом $r(x), r'(x)$ или равны нулю, или имеют степень, меньшую чем $\deg g(x) = s$, то

$$g(x)(q(x) - q'(x)) = r'(x) - r(x).$$

Если $q(x) - q'(x) \neq 0$, то получаем противоречие, поскольку степень левой части $\geqslant \deg g(x)$, а многочлен в правой части или нулевой, или его степень $< \deg g(x)$. Итак, $q(x) = q'(x)$, и поэтому $r'(x) = r(x)$. □

Замечание 1.13.9. Если K — подполе поля K' (например, $K = \mathbb{Q} \subset \mathbb{R} = K'$), $f(x), g(x) \in K[x] \subseteq K'[x]$, $f(x) = g(x)q(x) + r(x)$ — деление с остатком в кольце многочленов $K'[x]$, то $q(x), r(x) \in K[x]$.

Определение 1.13.10. Пусть $f(x), \varphi(x) \in K[x]$, $\varphi(x) \neq 0$. Будем говорить, что многочлен $f(x)$ *делится* на $\varphi(x)$, если $f(x) = \varphi(x)q(x)$ (т. е. остаток $r(x)$ при делении на $\varphi(x)$ равен нулю).

Замечание 1.13.11. Совокупность $\varphi(x)K[x] = \{\varphi(x)f(x) \mid f(x) \in K[x]\}$ всех многочленов, делящихся на $\varphi(x)$, является идеалом в кольце $K[x]$ (называемым *главным идеалом*, порождённым $\varphi(x)$).

Упражнение 1.13.12. Пусть K — поле. Покажите, что кольцо многочленов $K[x]$ является коммутативным кольцом главных идеалов.

Отметим ряд свойств делимости многочленов.

Лемма 1.13.13. *Если $f(x)$ делится на $g(x)$, $g(x)$ делится на $h(x)$, то $f(x)$ делится на $h(x)$.*

Доказательство. Действительно, если $f(x) = g(x)q(x)$, $g(x) = h(x)\tilde{q}(x)$, то $f(x) = h(x)\tilde{q}(x)q(x)$. □

Лемма 1.13.14. *Если $f(x)$ и $g(x)$ делятся на $h(x)$, то $f(x) + g(x)$, $f(x) - g(x)$ делятся на $h(x)$.*

Доказательство. Действительно, если $f(x) = h(x)q(x)$, $g(x) = h(x)\tilde{q}(x)$, то $f(x) \pm g(x) = h(x)(q(x) \pm \tilde{q}(x))$. □

Лемма 1.13.15. *Если многочлен $f(x)$ делится на $h(x)$, $g(x) \in K[x]$, то $f(x)g(x)$ делится на $h(x)$.*

Доказательство. Действительно, если $f(x) = h(x)q(x)$, то $f(x)g(x) = h(x)(q(x)g(x))$. □

Лемма 1.13.16. *Если $f_1(x), \ldots, f_k(x)$ делятся на $h(x)$, $g_1(x), \ldots, g_k(x) \in K[x]$, то $f_1(x)g_1(x) + \ldots + f_k(x)g_k(x)$ делится на $h(x)$.*

Доказательство. Действительно, это вытекает из лемм 1.13.15 и 1.13.14. □

Лемма 1.13.17. *Если $0 \neq c \in K$, то любой многочлен $f(x) \in K[x]$ делится на c.*

Доказательство. Действительно, $f(x) = c(c^{-1}f(x))$. □

Лемма 1.13.18. *Если $f(x)$ делится на $\varphi(x)$ и $0 \neq c \in K$, то $f(x)$ делится на $c\varphi(x)$.*

Доказательство. Действительно, если $f(x) = \varphi(x)q(x)$, то $f(x) = (c\varphi(x))(c^{-1}q(x))$. □

Лемма 1.13.19. *Многочлены вида $cf(x)$, $0 \neq c \in K$, и только они являются делителями многочлена $f(x)$, имеющими степень $\deg f(x)$.*

Лемма 1.13.20. *Многочлен $f(x)$ делится на $g(x)$ и $g(x)$ делится на $f(x)$ тогда и только тогда, когда $g(x) = cf(x)$, $0 \neq c \in K$.*

Лемма 1.13.21. *Многочлены $f(x)$ и $cf(x)$, $0 \neq c \in K$, обладают одинаковым запасом делителей в кольце $K[x]$.*

Определение 1.13.22. Пусть $f(x), g(x) \in K[x]$. Многочлен $d(x) \in K[x]$ называется *наибольшим общим делителем* (НОД) *многочленов* $f(x)$ и $g(x)$, если:

1) $d(x)$ — общий делитель многочленов $f(x)$ и $g(x)$ (т. е. $f(x) = d(x)q(x)$, $g(x) = d(x)\tilde{q}(x)$);

2) для любого общего делителя $d'(x)$ многочленов $f(x)$ и $g(x)$ многочлен $d(x)$ делится на $d'(x)$.

Обозначение: $d(x) = \text{НОД}(f(x), g(x)) = (f(x), g(x))$.

Замечание 1.13.23. Из 2) следует, что $\deg d(x) \geqslant \deg d'(x)$, т. е. что $d(x)$ — общий делитель наибольшей степени. Правда, нам ещё надо установить существование НОД в нашем смысле.

Теорема 1.13.24 (алгоритм Евклида). *Для любых $f(x), g(x) \in K[x]$:*

1) *существует наибольший общий делитель $d(x)$ многочленов $f(x)$ и $g(x)$;*

2) $d(x) = \text{НОД}(f(x), g(x))$ *находится по процедуре последовательного деления, восходящей к Евклиду;*

3) *наибольший делитель* $d(x)$ *определён однозначно с точностью до ненулевой константы* $0 \neq c \in K$.

Доказательство. 1), 2) Рассмотрим процедуру Евклида:

$$
\begin{aligned}
&f(x) = g(x)q_1(x) + r_1(x), && \deg r_1(x) < \deg g(x);\\
&g(x) = r_1(x)q_2(x) + r_2(x), && \deg r_2(x) < \deg r_1(x);\\
&r_1(x) = r_2(x)q_3(x) + r_3(x), && \deg r_3(x) < \deg r_2(x);\\
&\dots\\
&r_{k-3}(x) = r_{k-2}(x)q_{k-1}(x) + r_{k-1}(x), && \deg r_{k-1}(x) < \deg r_{k-2}(x);\\
&r_{k-2}(x) = r_{k-1}(x)q_k(x) + r_k(x), && \deg r_k(x) < \deg r_{k-1}(x);\\
&r_{k-1}(x) = r_k(x)q_{k+1}(x).
\end{aligned}
$$

а) Поднимаясь последовательно вверх, мы видим, что $r_k(x)$ — общий делитель многочленов $g(x)$ и $f(x)$.

б) Если $d'(x)$ — общий делитель многочленов $f(x)$ и $g(x)$, то, опускаясь последовательно вниз, мы видим, что $d'(x)$ — делитель многочлена $d(x)$.

3) Если $d(x)$ и $d'(x)$ — два наибольших общих делителя, то они делятся друг на друга, и поэтому $d'(x) = cd(x)$, $0 \neq c \in P$. Ясно, что если $d(x)$ — наибольший общий делитель и $0 \neq c \in P$, то $cd(x)$ — также наибольший общий делитель. □

Теорема 1.13.25 (о выражении наибольшего общего делителя через исходные многочлены). *Если* $f(x), g(x) \in K[x]$ *и* $d(x) = \text{НОД}(f(x), g(x))$, *то существуют многочлены* $u(x), v(x) \in K[x]$ *такие, что*

$$d(x) = f(x)u(x) + g(x)v(x)$$

(*если при этом* $\deg f(x) > 0$, $\deg g(x) > 0$, *то можно считать, что*

$$\deg u(x) < \deg g(x),$$
$$\deg v(x) < \deg f(x);$$

это позволяет искать многочлены $u(x)$, $v(x)$ *с неопределёнными коэффициентами*).

Доказательство. Существование таких многочленов $u(x)$, $v(x)$ следует из алгоритма Евклида нахождения $d(x) = r_k(x)$. Мы выражаем последовательно $r_k(x)$ сначала через $r_{k-2}(x)$ и $r_{k-1}(x)$, потом, подставляя выражение $r_{k-1}(x)$ через $r_{k-3}(x)$ и $r_{k-2}(x)$, через $r_{k-3}(x)$ и $r_{k-2}(x)$ и, завершая подъём, через $g(x)$ и $f(x)$.

Если найдены «плохие» $u(x)$ и $v(x)$, пусть, например, $\deg u(x) \geqslant \deg g(x)$, то $u(x) = g(x)q(x) + r(x)$, и поэтому $d(x) = f(x)r(x) + g(x)[v(x) + f(x)q(x)]$. Из сравнения степеней следует, что $\deg(v(x) + f(x)q(x)) < \deg f(x)$, поскольку $\deg(f(x)r(x)) < \deg f(x) + \deg g(x)$, $\deg d(x) \leqslant \deg f(x)$, $\deg d(x) \leqslant \deg g(x)$. □

Определение 1.13.26. Многочлены $f(x), g(x) \in K[x]$ из кольца многочленов $K[x]$ над полем K называются *взаимно простыми*, если их наибольший делитель $d(x)$ равен 1 (т. е. их общие делители — это лишь ненулевые многочлены нулевой степени $0 \neq c \in K$).

Теорема 1.13.27. *Многочлены $f(x), g(x) \in K[x]$ взаимно просты тогда и только тогда, когда существуют такие многочлены $u(x), v(x) \in K[x]$, что*

$$f(x)u(x) + g(x)v(x) = 1.$$

Доказательство.

1) Если многочлены $f(x)$ и $g(x)$ взаимно просты, то для их наибольшего делителя $d(x)$ имеем равенство

$$d(x) = 1.$$

Принимая во внимание выражение многочлена $d(x)$ через $f(x)$ и $g(x)$, получаем, что для некоторых $u(x), v(x) \in K[x]$

$$f(x)u(x) + g(x)v(x) = 1.$$

2) Если для $u(x), v(x) \in K[x]$ имеем

$$f(x)u(x) + g(x)v(x) = 1,$$

то любой общий делитель многочленов $f(x)$ и $g(x)$ является делителем многочлена 1. Таким образом,

$$\text{НОД}(f(x), g(x)) = 1,$$

другими словами, многочлены $f(x)$ и $g(x)$ взаимно просты. □

Замечание 1.13.28. Многочлены $f(x)$ и $g(x)$ взаимно просты тогда и только тогда, когда

$$K[x]f(x) + K[x]g(x) = K[x]$$

(идеал кольца $K[x]$, порождённый многочленами $f(x)$ и $g(x)$, совпадает со всем кольцом многочленов $K[x]$).

Теорема 1.13.29 (основные свойства взаимно простых многочленов). *Пусть* $f(x), g(x), \varphi(x), \psi(x) \in K[x]$.

1) *Если* НОД$(f, \varphi) = 1$, НОД$(f, \psi) = 1$, *то* НОД$(f, \varphi\psi) = 1$.

2) *Если* fg *делится на* φ *и* НОД$(f, \varphi) = 1$, *то* g *делится на* φ.

3) *Если* f *делится на* φ *и делится на* ψ, НОД$(\varphi, \psi) = 1$, *то* f *делится на* $\varphi\psi$.

Доказательство.

1) Пусть

$$fu + \varphi v = 1$$

для $u(x), v(x) \in K[x]$. Умножая это равенство на ψ, получаем

$$f(u\psi) + (\varphi\psi)v = \psi.$$

Отсюда следует, что любой общий делитель многочленов f и $\varphi\psi$ является делителем многочлена ψ, но многочлены f и ψ взаимно просты. Таким образом, НОД$(f, \varphi\psi) = 1$.

2) Пусть для $u(x), v(x) \in K[x]$ имеем

$$fu + \varphi v = 1.$$

Умножив это равенство на $g(x)$, получим

$$(fg)u + \varphi(vg) = g,$$

и поэтому многочлен g делится на φ, поскольку оба слагаемых в левой части делятся на φ.

3) Пусть $f = \varphi q$, где $q(x) \in K[x]$. Так как $f = \varphi q$ делится на ψ и НОД$(\varphi, \psi) = 1$, то, в силу 2), $q = \psi\chi$, где $\chi(x) \in K[x]$. Итак:

$$f = \varphi q = (\varphi\psi)\chi. \quad \square$$

Замечание 1.13.30. Определив наибольший общий делитель

$$d(x) = \text{НОД}(f_1(x), \ldots, f_s(x))$$

многочленов

$$f_1(x), \ldots, f_s(x) \in K[x], \quad s \geqslant 1,$$

как такой делитель этих многочленов $f_1(x), \ldots, f_s(x)$, который делится на любой их общий делитель, получаем, проводя индукцию по s, что

$$d(x) = \text{НОД}(f_s(x), \text{НОД}(f_1(x), \ldots, f_{s-1}(x))).$$

Упражнение 1.13.31. Если

$$f(x) = x(x-1),\ g(x) = x(x-2),\ h(x) = (x-1)(x-2) \in \mathbb{R}[x],$$

то

$$\text{НОД}(f, g) = x, \quad \text{НОД}(f, h) = x - 1,$$
$$\text{НОД}(g, h) = x - 2, \quad \text{НОД}(f, g, h) = 1.$$

Замечание 1.13.32. В алгоритме Евклида можно для удобства делимое и делитель на каждом шаге умножать на любые ненулевые числа (при этом мы не заботимся о точном вычислении коэффициентов в частных $q_i(x)$).

Пример 1.13.33. Найти $\text{НОД}(f(x), g(x))$, где

$$f(x) = 2x^4 + 2x^3 + x^2 - x - 1,$$
$$g(x) = 3x^4 + 2x^2 - x + 2.$$

Решение. $3f(x) = g(x)q_1(x) + r_1(x)$, где $q_1(x) = 2$, $r_1(x) = 6x^3 - x^2 - x - 7$. Делим $2g(x)$ на $r_1(x)$:

$$\begin{array}{rl|l} 6x^4 + 0x^3 + 4x^2 - 2x + 4 & 6x^3 - x^2 - x - 7 & \\ \cline{1-1} 6x^4 - x^3 - x^2 - 7x & x + 1 & \\ \cline{1-1} x^3 + 5x^2 + 5x + 4 & & \\ \cdots\cdots\cdots\cdots\cdots & & \\ 6x^3 + 30x^2 + 30x + 24 & & \\ \cline{1-1} 6x^3 - x^2 - x - 7 & & \\ \cline{1-1} 31x^2 + 31x + 31 & & \end{array}$$

Многоточием ... отмечено место, в котором мы произвели домножение на 6 (соответственно многоточие ⋮ показывает, что мы не находим точные коэффициенты для $q_2(x)$). Таким образом,

$$g(x) = r_1(x)q_2(x) + r_2(x),$$

где с точностью до ненулевого множителя $r_2(x) = x^2 + x + 1$. Далее,

$$\begin{array}{r|l} 6x^3 - x^2 - x - 7 & x^2 + x + 1 \\ \hline 6x^3 + 6x^2 + 6x \quad\vdots & 6x - 7 \\ \hline -7x^2 - 7x - 7 & \\ -7x^2 - 7x - 7 & \\ \hline 0 & \end{array}$$

То есть $r_1(x)$ делится нацело на $r_2(x)$. Итак,

$$\text{НОД}(f(x), g(x)) = x^2 + x + 1.$$

Упражнение 1.13.34. Наибольший общий делитель $d(x)$ многочленов $f(x) = 3x^5 - 4x^4 + x^3 - 3x^2 + 4x - 1$ и $g(x) = 3x^5 + 5x^4 + x^3 - x^2 - 3x + 1$ представить в виде

$$d(x) = f(x)u(x) + g(x)v(x),$$

где $u(x)$, $v(x)$ — многочлены степеней, меньших чем степени многочленов $g(x)$ и $f(x)$ соответственно.

Решение. Сначала с помощью алгоритма Евклида находим

$$d(x) = 3x^3 + 2x^2 + 2x - 1,$$

при этом

$$f_1(x) = \frac{f(x)}{d(x)} = x^2 - 2x + 1,$$
$$g_1(x) = \frac{g(x)}{d(x)} = x^2 + x - 1.$$

Ищем многочлены $u(x)$ и $v(x)$ такие, что

$$1 = f_1(x)u(x) + g_1(x)v(x). \tag{1.1}$$

Так как степени многочленов $u(x)$ и $v(x)$ должны быть меньше двух, то $u(x) = ax + b$, $v(x) = cx + d$, где $a, b, c, d \in \mathbb{R}$. Приравнивая в (1.1) коэффициенты при одинаковых степенях переменной x, получаем систему линейных уравнений для a, b, c, d. Решая эту систему, получаем, что $a = 3$, $b = 5$, $c = -3$, $d = 4$. Итак,

$$d(x) = 3x^3 + 2x^2 + 2x - 1 = f(x)(3x + 5) + g(x)(-3x + 4).$$

Определение 1.13.35. Пусть K — поле,

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in K[x], \quad a_n, \ldots, a_0 \in K.$$

Если $c \in K$, то элемент

$$f(c) = a_n c^n + a_{n-1}c^{n-1} + \ldots + a_1 c + a_0 \in K$$

назовём *значением многочлена* $f(x)$ при $x = c$. Таким образом, получаем отображения:

$$f\colon K \to K, \quad c \mapsto f(c)$$

(*полиномиальная функция, определяемая многочленом* $f(x)$);

$$K[x] \to K, \quad f(x) \mapsto f(c)$$

(ясно, что если $f(x) = g(x)$ в $K[x]$, то $f(c) = g(c)$ для всех $c \in K$).

Лемма 1.13.36. *Если в* $K[x]$

$$\varphi(x) = f(x) + g(x), \quad \psi(x) = f(x)g(x)$$

и $c \in K$, *то*

$$\varphi(c) = f(c) + g(c), \quad \psi(c) = f(c)g(c).$$

Таким образом, отображение

$$\Delta_c\colon K[x] \to K, \quad f(x) \mapsto f(c),$$

является гомоморфизмом колец (*при этом* $\operatorname{Ker}\Delta_c = \{f(x) \in K[x] \mid f(c) = 0\}$).

Доказательство следует из определения сложения и умножения многочленов в кольце $K[x]$. □

Определение 1.13.37. Элемент $c \in K$ называется *корнем многочлена* $f(x) \in K[x]$, если $f(c) = 0$.

Теорема 1.13.38 (Безу). *Пусть $c \in K$. Остаток от деления многочлена $f(x)$ в кольце $K[x]$ на множитель $x - c$ равен значению $f(c)$ многочлена $f(x)$ при $x = c$.*

Доказательство. В силу алгоритма деления

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

где или $r(x) = 0$, или $\deg r(x) = 0$, и поэтому $r(x) = r \in K$. Итак, $f(x) = (x - c)q(x) + r$, следовательно, $f(c) = (c - c)q(c) + r = r$, и поэтому $f(x) = (x - c)q(x) + f(c)$. □

Следствие 1.13.39. *Элемент $c \in K$ является корнем многочлена $f(x) \in K[x]$ тогда и только тогда, когда многочлен $f(x)$ делится на $x - c$.* □

Замечание 1.13.40.

1) Если $a, b \in K$, $a \neq 0$, то делимость многочлена $f(x) \in K[x]$ на многочлен $ax + b = a\left(x - \left(-\dfrac{b}{a}\right)\right)$ равносильна делимости на многочлен $x - c$, $c = -\dfrac{b}{a}$, и поэтому нахождение корней многочлена $f(x) \in K[x]$ в поле K равносильно нахождению его линейных делителей в кольце $K[x]$.

2) Если $c \in K$, $\Delta_c \colon K[x] \to K$, $\Delta_c(f) = f(c)$, то

$$\operatorname{Ker} \Delta_c = \{f(x) \in K[x] \mid f(c) = 0\} = (x - c)K[x] = I_c$$

(главный идеал в кольце $K[x]$, порождённый многочленом $x - c$).

Замечание 1.13.41 (схема (алгоритм) Горнера деления многочлена $f(x) \in K[x]$ на линейный многочлен $x - c$, $c \in K$). Пусть $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in K[x]$,

$$f(x) = (x - c)q(x) + r, \quad r \in K,$$
$$q(x) = b_{n-1}x^{n-1} + \ldots + b_1 x + b_0 \in K[x].$$

Тогда, приравнивая коэффициенты при $x^n, x^{n-1}, \ldots, x, 1$, соответственно получаем

$$
\begin{aligned}
a_n &= b_{n-1};\\
a_{n-1} &= b_{n-2} - cb_{n-1};\\
a_{n-2} &= b_{n-3} - cb_{n-2};\\
&\ldots\\
a_k &= b_{k-1} - cb_k;\\
&\ldots\\
a_1 &= b_0 - cb_1;\\
a_0 &= r - cb_0.
\end{aligned}
$$

Пересчитывая, получаем

$$
\begin{aligned}
b_{n-1} &= a_n;\\
b_{n-2} &= cb_{n-1} + a_{n-1};\\
b_{n-3} &= cb_{n-2} + a_{n-2};\\
&\ldots\\
b_{k-1} &= cb_k + a_k;\\
&\ldots\\
b_0 &= cb_1 + a_1;\\
r &= cb_0 + a_0.
\end{aligned}
$$

Таким образом, коэффициенты частного $b_{n-1}, \ldots, b_1, b_0$ и остаток $r = f(c)$ последовательно вычисляются по коэффициентам $a_n, \ldots, a_1, a_0$ и элементу c, если использовать однотипную процедуру

	a_n	a_{n-1}	...	a_{k+1}	a_k	...	a_1	a_0
	$\Vert$				$\downarrow$			
c	b_{n-1}	$b_{n-2}=$	...	$b_k=$	$b_{k-1}=$	...	$b_0=$	$r=$
		$=cb_{n-1}+a_{n-1}$	...	$=cb_{k+1}+a_{k+1}$	$=cb_k+a_k$	...	$=cb_1+a_1$	$=cb_0+a_0$

Пример 1.13.42. Пусть $f(x) = 2x^4 - x^2 + 3x - 2$, $c = -2$. Тогда

	2	0	−1	3	−2
−2	2	−4	7	−11	20

поэтому $f(x) = (x+2)q(x) + 20$, где $q(x) = (x+2)(2x^3 - 4x^2 + 7x - 11)$.

Замечание 1.13.43.

1) Схема Горнера даёт быстрый алгоритм вычисления значения $r = f(c)$ многочлена $f(x) \in K[x]$ в точке c (минимизируя число умножений).

2) Последовательное применение схемы Горнера позволяет построить эффективный алгоритм записи многочлена $f(x)$ в виде формулы Тейлора по степеням $(x - c)$. А именно, при первом применении схемы Горнера крайний правый коэффициент равен $f(c)$, при втором применении крайний справа коэффициент равен $f'(c)$, при третьем — $\dfrac{f''(c)}{2!}$, и так далее. Таким образом, если $\deg f(x) = n$, то

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

(формула Тейлора).

Например, для

$$f(x) = x^4 - 6x^3 - 2x^2 + 5x - 4$$

и $c = 5$ имеем

	1	-6	-2	5	-4
5	1	-1	-7	-30	$-154 = f(5)$
5	1	4	13	$35 = f'(5)$	
5	1	9	$58 = \dfrac{f''(5)}{2!}$		
5	1	$14 = \dfrac{f^{(3)}(5)}{3!}$			
	$1 = \dfrac{f^{(4)}(5)}{4!}$				

Таким образом,

$$f(x) = (x-5)^4 + 14(x-5)^3 + 58(x-5)^2 + 35(x-5) - 154.$$

Определение 1.13.44. Пусть $f(x) \in K[x]$, $c \in K$, и c — корень многочлена $f(x)$, т. е. $f(c) = 0$. По теореме Безу многочлен $f(x)$ делится на $x-c$. Возможно, многочлен $f(x)$ делится на более высокие степени многочлена $x - c$. Пусть $k \in \mathbb{N}$ — такое натуральное число, что $f(x)$ делится на $(x-c)^k$, но не делится на $(x-c)^{k+1}$, поэтому

$$f(x) = (x-c)^k \varphi(x),$$

многочлен $\varphi(x) \in K[x]$ уже не делится на $x - c$ (это равносильно тому, что $\varphi(c) \neq 0$). В этом случае число k назовём *кратностью* корня c многочлена $f(x)$, а сам корень c — k-кратным корнем многочлена $f(x)$. Если $k = 1$, то корень c называется *простым корнем* многочлена $f(x)$.

Замечание 1.13.45. Понятие *абстрактного линейного пространства* мы детально рассмотрим в главе 9, после того как изучим ряд конкретных линейных пространств.

Понятие *алгебры над полем* (как кольца, являющегося к тому же и линейным пространством) будет рассмотрено в главе 8.

Глава 2

Поле $\mathbb{C}$ комплексных чисел

Понятие числа является одним из основных понятий в математических теориях. К основным числовым системам принадлежат:

- натуральные числа $\mathbb{N}$ (полукольцо);
- натуральные числа с нулём $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ (полукольцо с нулём);
- целые числа $\mathbb{Z}$ (кольцо);
- рациональные числа $\mathbb{Q}$ (поле);
- действительные числа $\mathbb{R}$ (поле).

При этом

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Отметим, что рациональные числа $\mathbb{Q}$ и действительные числа $\mathbb{R}$ с операциями сложения и умножения являются *полями*. Напомним, что множество K с операциями сложения и умножения, $(K, +, \cdot)$, называется полем, если:

1) операция сложения

коммутативна $(a + b = b + a \quad \forall a, b \in K)$;
ассоциативна $((a + b) + c = a + (b + c) \quad \forall a, b, c \in K)$;
существует нейтральный элемент 0 $(0 + a = a \quad \forall a \in K)$;
$\forall a \in K$ существует противоположный элемент $-a$ $(a + (-a) = 0)$

(кратко, $(K, +)$ — *коммутативная группа*):

2) операция умножения

коммутативна ($ab = ba \quad \forall a, b \in K$);

ассоциативна ($(ab)c = a(bc) \quad \forall a, b, c \in K$);

существует нейтральный элемент 1 ($1a = a \quad \forall a \in K$), $1 \neq 0$

(кратко, $(K, \cdot)$ — *коммутативный моноид*);

3) имеет место дистрибутивность, связывающая операции сложения и умножения ($(a + b)c = ac + bc \quad \forall a, b, c \in K$).

Условия 1), 2), 3) определяют *коммутативное кольцо.*

4) Имеет место обратимость ненулевых элементов ($\forall a \in K$, $a \neq 0$, $\exists b \in K \quad ab = 1$).

Поле действительных чисел $\mathbb{R}$, при всех его достоинствах, не является *алгебраически замкнутым полем* (т. е. многочлены с действительными коэффициентами могут не иметь действительных корней: например, многочлен $x^2 + 1$ не имеет действительного корня). Нашей целью является построение расширения $\mathbb{C}$ поля действительных чисел $\mathbb{R}$, $\mathbb{R} \subset \mathbb{C}$, в котором есть такой элемент $i \in \mathbb{C}$, что $i^2 = -1$ (уравнение $x^2 + 1 = 0$ имеет решение), при этом в некотором смысле это минимальное расширение с этим свойством. Построенное поле $\mathbb{C}$ окажется алгебраически замкнутым (*алгебраическим замыканием* поля $\mathbb{R}$).

2.1. Анализ ситуации

Допустим, что существует поле K, содержащее в качестве подполя поле действительных чисел, $\mathbb{R} \subset K$, и элемент $i \in K$ такой, что $i^2 = -1$. Тогда:

1) для $a, b, c, d \in \mathbb{R}$ равенство $a + bi = c + di$ выполнено тогда и только тогда, когда $a = c$, $b = d$.

Доказательство. Если $a + bi = c + di$, то $a - c = (d - b)i$, поэтому $(a - c)^2 = -(d - b)^2$, следовательно, $(a - c)^2 = 0 = (d - b)^2$, т. е. $a = c$, $b = d$. □

2) подмножество D всех элементов $a+bi$, $a,b\in\mathbb{R}$, замкнуто относительно операции сложения

$$(a+bi)+(c+di)=(a+c)+(b+d)i,$$

$0=0+0i\in D$ является в D нейтральным элементом, $-(a+bi)=(-a)+(-b)i$ — противоположный элемент для $a+bi$. Итак, D относительно сложения — коммутативная группа. □

3) подмножество D замкнуто относительно умножения

$$(a+bi)(c+di)=(ac-bd)+(ad+bc)i. \quad \square$$

4) $(a+bi)(a-bi)=a^2+b^2$. □

5) если $a+bi\neq 0$, то $a^2+b^2>0$, и

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

следовательно,

$$(a+bi)^{-1}=\frac{a}{a^2+b^2}-\frac{b}{a^2+b^2}i$$

для $a+bi\neq 0$. Итак, D является подполем поля K, $\mathbb{R}\subset D$, $i\in D$, D — наименьшее подполе в K, содержащее $\mathbb{R}$ и i. □

2.2. Построение поля комплексных чисел

На основе проведённого анализа положим

$$\mathbb{C}=\mathbb{R}^2=\{(a,b)\mid a,b\in\mathbb{R}\}-$$

совокупность упорядоченных пар действительных чисел.

Рассмотрим следующие операции сложения и умножения:

$$(a,b)+(c,d)=(a+c,b+d),$$
$$(a,b)(c,d)=(ac-bd,ad+bc).$$

Тогда:

1) $\mathbb{C} = (\mathbb{R}^2, +)$ — абелева группа (сложение ассоциативно и коммутативно; $(0,0)$ — нейтральный элемент; $(-a,-b)$ — противоположный элемент для (a,b));

2) умножение: ассоциативно

$$\begin{aligned}((a,b)(c,d))(e,f) &= (ac-bd, ad+bc)(e,f) = \\ &= ((ac-bd)e-(ad+bc)f, (ac-bd)f+(ad+bc)e) = \\ &= (ace-bde-adf-bcf, acf-bdf+ade+bce) = \\ &= (ace-adf-bcf-bde, acf+ade+bce-bdf) = \\ &= (a(ce-df)-b(cf+de), a(cf+de)+b(ce-df)) = \\ &= (a,b)(ce-df, cf+de) = (a,b)((c,d)(e,f));\end{aligned}$$

коммутативно

$$(a,b)(c,d) = (ac-bd, ad+bc) = (ca-db, cb+da) = (c,d)(a,b);$$

$(1,0)$ — нейтральный элемент, $(a,b)(1,0) = (a,b)$, $(1,0) \neq (0,0)$;

3) выполнено свойство дистрибутивности:

$$\begin{aligned}(a,b)((c,d)+(e,f)) &= (a,b)((c+e, d+f)) = \\ &= (a(c+e)-b(d+f), a(d+f)+b(c+e)) = \\ &= (ac+ae-bd-bf, ad+af+bc+be) = \\ &= (ac-bd+ae-bf, ad+bc+af+be) = \\ &= (ac-bd, ad+bc)+(ae-bf, af+be) = \\ &= (a,b)(c,d)+(a,b)(e,f).\end{aligned}$$

Итак, $\mathbb{C} = \mathbb{R}^2$ с этими операциями сложения и умножения является коммутативным кольцом с единицей $(1,0)$.

Если $(a,b) \neq (0,0)$, $a, b \in \mathbb{R}$, то $a^2+b^2 > 0$ и

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

таким образом, каждый элемент $(0,0) \neq (a,b) \in \mathbb{C} = \mathbb{R}^2$ имеет обратный.

Итак, $\mathbb{C} = \mathbb{R}^2$ с этими операциями сложения и умножения — поле.

Осуществим вложение поля действительных чисел $\mathbb{R}$ в построенное поле $\mathbb{C} = \mathbb{R}^2$, сопоставляя любому элементу $a \in \mathbb{R}$ пару $(a, 0) \in \in \mathbb{C} = \mathbb{R}^2$. Так как для $a, b \in \mathbb{R}$ имеем

$$(a + b, 0) = (a, 0) + (b, 0),$$
$$(ab, 0) = (a, 0)(b, 0),$$

то это отображение является изоморфизмом поля $\mathbb{R}$ на подполе $\{(a, 0) \mid a \in \mathbb{R}\}$ поля $\mathbb{C} = \mathbb{R}^2$. В дальнейшем мы будем отождествлять a и $(a, 0)$, полагая $a = (a, 0)$, в частности $1 = (1, 0)$.

Если $i = (0, 1)$, то

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

(элемент $i = (0, 1)$ в построенном расширении $\mathbb{C} = \mathbb{R}^2$ поля $\mathbb{R}$ является корнем уравнения $x^2 + 1 = 0$).

Для любых $a, b \in \mathbb{R}$ имеем

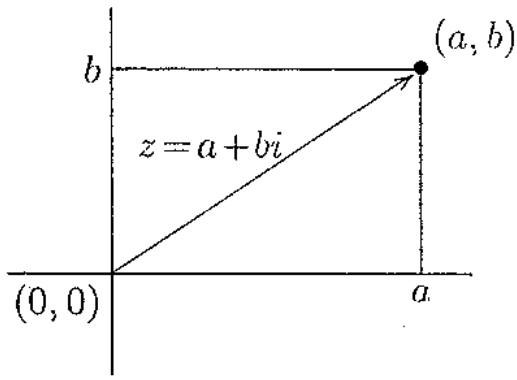
$$(a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0)(0, 1) = a + bi,$$

при этом это представление единственно (как из «анализа задачи», так и непосредственно: если $a+bi = c+di$, то $(a, b) = a+bi = c+di = = (c, d)$, следовательно, $a = c$, $b = d$).

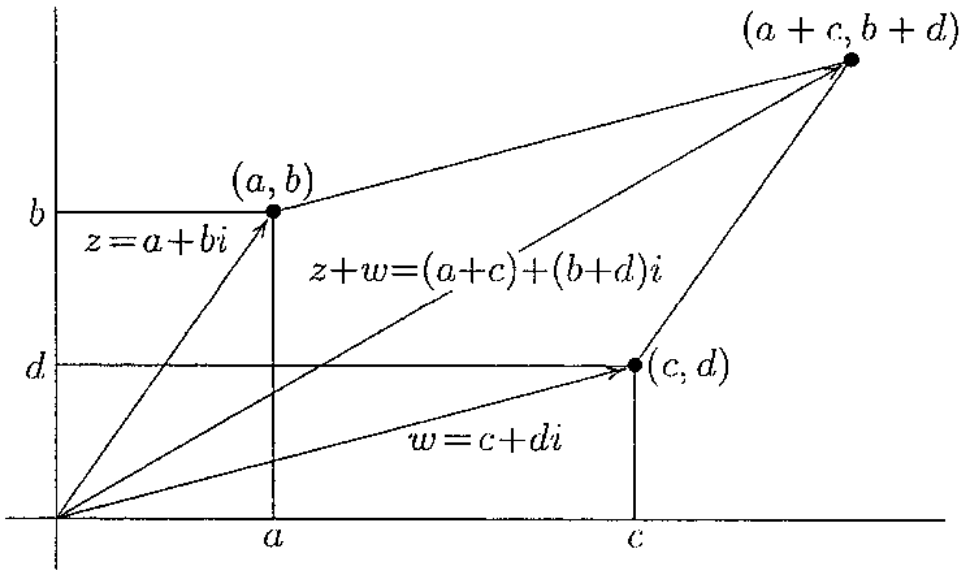
Элементы построенного поля $\mathbb{C} = \mathbb{R}^2$ называются *комплексными числами*. Форма записи комплексного числа в виде $a + bi$, $a, b \in \mathbb{R}$, называется *алгебраической формой записи*, в которой:

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i,$$
$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i \quad \text{для} \quad a + bi \neq 0.$$

В геометрической интерпретации комплексное число $z = a + bi$ изображается вектором в прямоугольной системе координат, выходящим из точки $(0, 0)$ в точку (a, b).



Сложение комплексных чисел соответствует сложению векторов:



Геометрическая интерпретация умножения и перехода к обратному элементу будет дана позже.

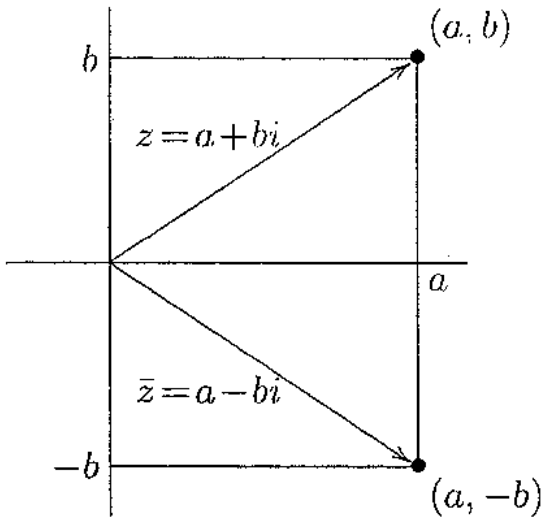
Для комплексного числа $z = a + bi \in \mathbb{C}$, $a, b \in \mathbb{R}$, $a = \operatorname{Re} z$ называется его вещественной частью, $b = \operatorname{Im} z$ — его мнимой частью.

Замечание 2.2.1. В построенном поле $\mathbb{C}$ уравнение $x^2 + 1 = 0$ имеет лишь два решения: $x = i$, $x = -i$. Действительно, если $(a + bi)^2 = -1$, то $a^2 - b^2 = -1$, $2ab = 0$. Так как $b \neq 0$ (иначе $a^2 = -1$), то $a = 0$ и $b^2 = 1$, поэтому $b = \pm 1$.

2.3. Сопряжение комплексных чисел

Каждому комплексному числу $z = x + iy \in \mathbb{C}$ сопоставим комплексное число $\bar{z} = x - iy \in \mathbb{C}$, называемое *комплексно сопряжённым*. Геометрическая интерпретация перехода от $z = a + bi$ к сопря-

жённому комплексному числу $\bar{z} = a - bi$ прозрачна: это отражение относительно вещественной оси:



Теорема 2.3.1.

1) *Операция комплексного сопряжения* $z \to \bar{z}$ *является автоморфизмом поля* $\mathbb{C}$ *комплексных чисел (т. е. биекцией, для которой* $\overline{z+w} = \bar{z}+\bar{w}$, $\overline{zw} = \bar{z}\bar{w}$ *для* $z, w \in \mathbb{C}$ *и, как следствие,* $\overline{\left(\frac{w}{z}\right)} = \frac{\bar{w}}{\bar{z}}$ *для* $z \neq 0$*), оставляющим все действительные числа и только их на месте* ($\bar{a} = a$ *для* $a \in \mathbb{R} \subseteq \mathbb{C}$*; если* $\bar{z} = z$*, то* $z \in \mathbb{R}$).

2) *Квадрат комплексного сопряжения равен тождественному отображению* ($\bar{\bar{z}} = z$).

3) *Если* $z = a+bi \in \mathbb{C}$, $a, b \in \mathbb{R}$, *то* $z+\bar{z} = 2a \in \mathbb{R}$, $z-\bar{z} = 2bi \in \mathbb{R}i$, $N(z) = z\bar{z} = a^2 + b^2 \in \mathbb{R}$, *при этом* $N(wz) = N(w)N(z)$ *для* $w, z \in \mathbb{C}$.

4) *Если* $f\colon \mathbb{C} \to \mathbb{C}$ *— такой автоморфизм поля* $\mathbb{C}$ *комплексных чисел, что* $f(a) = a$ *для всех* $a \in \mathbb{R} \subseteq \mathbb{C}$, *то либо* $f = 1_{\mathbb{C}}$, *либо* $f(z) = \bar{z}$ *для* $z \in \mathbb{C}$ *(тем самым показано, что группа Галуа расширения* $\mathbb{R} \subset \mathbb{C}$ *состоит из двух элементов*).

Доказательство.

1) Ясно, что соответствие

$$z = a + bi = (a, b) \to \bar{z} = a - bi = (a, -b)$$

является биекцией.

Если $z = a + bi$, $w = c + di$, то

$$\overline{z+w} = \overline{(a+b)+(c+d)i} = (a+b)-(c+d)i =$$
$$= (a-ci)+(b-di) = \bar{z}+\bar{w};$$
$$\overline{-z} = \overline{-a-bi} = -a+bi = -(a-bi) = -(\bar{z});$$
$$\overline{z-w} = \bar{z}+\overline{(-w)} = \bar{z}-\bar{w};$$
$$\overline{zw} = \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i =$$
$$= (a-bi)(c-di) = \bar{z}\bar{w}.$$

Если $z \neq 0$, то $1 = \overline{z \cdot z^{-1}} = \overline{z} \cdot \overline{z^{-1}}$, т. е. $\overline{z^{-1}} = (\bar{z})^{-1}$. Поэтому

$$\overline{\left(\frac{w}{z}\right)} = \overline{wz^{-1}} = \bar{w}\overline{(z^{-1})} = \bar{w}(\bar{z})^{-1} = \frac{\bar{w}}{\bar{z}}.$$

Если $z = a \in \mathbb{R}$, то $\bar{z} = a$. Если $z = a + bi$, то $\bar{z} = z$ означает, что $z = a + bi = a - bi = \bar{z}$, т. е. $b = -b$, поэтому $b = 0$ и $z = a \in \mathbb{R}$. Итак, $z = \bar{z}$ тогда и только тогда, когда $z \in \mathbb{R}$.

2) $\bar{\bar{z}} = \overline{a-bi} = a + bi = z$.

3) Если $z = a + bi$, то

$$z + \bar{z} = (a+bi)+(a-bi) = 2a \in \mathbb{R},$$
$$z - \bar{z} = (a+bi)-(a-bi) = 2bi \in \mathbb{R}i,$$
$$z \cdot \bar{z} = (a+bi)(a-bi) = a^2+b^2 \in \mathbb{R}.$$

Далее,

$$N(wz) = wz\overline{wz} = wz\bar{w}\bar{z} = w\bar{w}z\bar{z} = N(w)N(z).$$

4) Так как $i^2 = -1$, то $f(i)^2 = f(-1) = -1$, поэтому

либо $f(i) = i$, и тогда $f(a+bi) = f(a) + f(b)f(i) = a + bi$,
либо $f(i) = -i$, и тогда $f(a+bi) = f(a) + f(b)f(i) = a - bi$. □

Замечание 2.3.2.

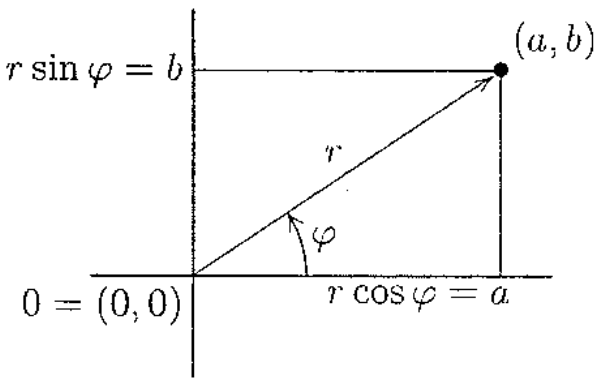
1) Если комплексное число α получено как выражение из комплексных чисел $\alpha_1, \ldots, \alpha_n$ с помощью операций сложения, вычитания, умножения и деления, то то же выражение из комплексных чисел $\bar{\alpha}_1, \ldots, \bar{\alpha}_n$ даёт $\bar{\alpha}$.

2) Правило деления комплексного числа $w = c + di$ на ненулевое комплексное число $0 \neq z = a + bi \in \mathbb{C}$ (в алгебраической форме):

$$\frac{w}{z} = \frac{c+di}{a+bi} = \frac{(c+di)(a-bi)}{(a+bi)(a-bi)} = \left(\frac{ca+db}{a^2+b^2}\right) + \left(\frac{-cb+da}{a^2+b^2}\right) i.$$

2.4. Полярные координаты точек плоскости (отличных от начала координат)

Точка плоскости (a, b), отличная от начала координат $(0, 0)$, однозначно задаётся своими *полярными координатами* r, φ, где r — расстояние от данной точки до начала координат, φ — угол между положительной полуосью абсцисс и радиусом-вектором точки (a, b), отсчитываемый против часовой стрелки (определённый с точностью до $2\pi k$, $k \in \mathbb{Z}$, и называемый *аргументом точки* (a, b)).



Аргумент точки $0 = (0, 0)$ не определён.

Формулы перехода от декартовых координат a и b точки (a, b) к полярным координатам и обратно:

$$r = \sqrt{a^2 + b^2},$$

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

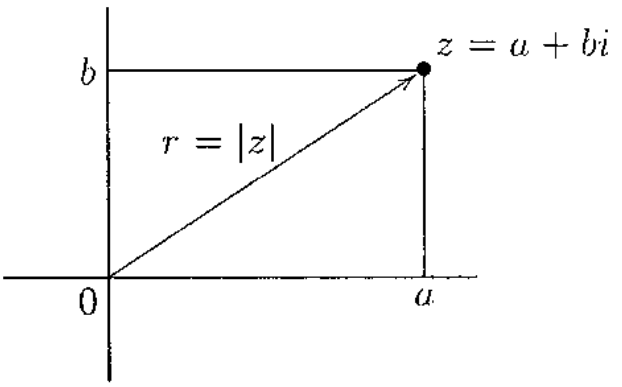
$$a = r \cos \varphi, \quad b = r \sin \varphi.$$

2.5. Свойства модуля комплексных чисел

Для комплексного числа $z = a + bi \in \mathbb{C}$ определим его *модуль* как

$$|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$$

(в геометрической интерпретации на плоскости $\mathbb{C} = \mathbb{R}^2$ модуль комплексного числа $r = |z| = \sqrt{a^2 + b^2}$ — это расстояние от точки $(0, 0)$ до точки (a, b), т. е. длина вектора z).



Пример 2.5.1.

1) Если $z = a \in \mathbb{R}$, то $|z| = \sqrt{a^2} = |a|$, т. е. функция модуль комплексного числа $|\ |\colon \mathbb{C} \to \mathbb{R}^+$ является продолжением функции модуль действительного числа $|\ |\colon \mathbb{R} \to \mathbb{R}^+ = \{r \in \mathbb{R},\ r \geqslant 0\}$.

2) $|i| = 1$, $|1 + i| = \sqrt{2}$.

3) $|z| = \sqrt{a^2 + b^2} = |\bar{z}|$ для $z = a + bi \in \mathbb{C}$.

4) $|a| = \sqrt{a^2} \leqslant \sqrt{a^2 + b^2} = |z|$ для $z = a + bi \in \mathbb{C}$.

Лемма 2.5.2. $|wz| = |w|\,|z|$ *для* $w, z \in \mathbb{C}$.

Доказательство.

$$|wz| = \sqrt{(wz)\overline{(wz)}} = \sqrt{(w\bar{w})(z\bar{z})} = \sqrt{w\bar{w}}\sqrt{z\bar{z}} = |w|\,|z|. \quad \square$$

Другое доказательство этого факта следует из свойств тригонометрической формы.

Следствие 2.5.3.

1) *Если* $w = z^{-1}$ *для* $0 \neq z \in \mathbb{C}$, *то* $1 = |1| = |z^{-1}z| = |z^{-1}|\,|z|$, *поэтому*

$$|w| = \left|\frac{1}{z}\right| = \frac{1}{|z|}, \quad \textit{или} \quad |z^{-1}| = |z|^{-1}.$$

2) *Для* $w, z \in \mathbb{C}$, $z \neq 0$:

$$\left|\frac{w}{z}\right| = |wz^{-1}| = |w|\,|z^{-1}| = |w|\,|z^{-1}| = \frac{|w|}{|z|}.$$

Лемма 2.5.4. $|w+z| \leqslant |w| + |z|$ *для* $w, z \in \mathbb{C}$.

Первое доказательство. Длина $|w+z|$ стороны треугольника не превосходит суммы длин $|w| + |z|$ двух других сторон. □

Второе доказательство. Если $w = 0$ или $z = 0$, то утверждение очевидно.

Пусть теперь $w \neq 0$ и $z \neq 0$. Так как для $z = a + bi$ имеем

$$|a| = \sqrt{a^2} \leqslant \sqrt{a^2 + b^2} = |z|,$$

то

$$|1+z|^2 = (1+z)(1+\bar{z}) = 1 + (z+\bar{z}) + z\bar{z} =$$
$$= 1 + 2a + |z|^2 \leqslant 1 + 2|z| + |z|^2 = (1+|z|)^2,$$

и поэтому, поскольку $|1+z| \geqslant 0$, $1+|z| \geqslant 0$,

$$|1+z| \leqslant 1 + |z|.$$

Далее,

$$|w+z| = |w(1+w^{-1}z)| = |w| \cdot |1+w^{-1}z| \leqslant$$
$$\leqslant |w|(1+|w^{-1}z|) = |w|(1+|w^{-1}|\,|z|) =$$
$$= |w| + |ww^{-1}|\,|z| = |w| + |z|. \quad \square$$

Следствие 2.5.5. $|w| - |z| \leqslant |w \pm z| \leqslant |w| + |z|$ *для* $w, z \in \mathbb{C}$.

Доказательство.
1) $|w - z| \leqslant |w| + |-z| = |w| + |z|$.
2) Так как $|w| = |(w-z)+z| \leqslant |w-z| + |z|$, то $|w| - |z| \leqslant |w-z|$.
3) $|w| - |z| = |w| - |-z| \leqslant |w+z|$. □

2.6. Тригонометрическая форма ненулевого комплексного числа

Используя полярные координаты, модуль $r = \sqrt{a^2 + b^2}$ и аргумент $\varphi = \arg z$, для комплексного числа $z = a + bi$ и принимая во внимание, что $a = r\cos\varphi$, $b = r\sin\varphi$, получаем *тригонометрическую форму*:

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

Примеры 2.6.1.

1) $1 = 1(\cos 0 + i\sin 0)$;

2) $i = 1\left(\cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}\right)$;

3) $z = -1 + \sqrt{3}i$, $r = |z| = \sqrt{1+3} = 2$, $\cos\varphi = -\dfrac{1}{2}$, $\sin\varphi = \dfrac{\sqrt{3}}{2}$, $\varphi = \dfrac{2\pi}{3}$, поэтому

$$z = -1 + \sqrt{3}i = 2\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right);$$

4) $\cos\varphi - i\sin\varphi = \cos(-\varphi) + i\sin(-\varphi)$.

Теорема 2.6.2 (о единственности тригонометрической формы). *Если* $0 \neq z = a + bi \in \mathbb{C}$ *и*

$$z = r_1(\cos\varphi_1 + i\sin\varphi_1) = r_2(\cos\varphi_2 + i\sin\varphi_2),$$

где $\mathbb{R} \ni r_1 > 0$, $\mathbb{R} \ni r_2 > 0$, *то*

$$r_1 = r_2 \quad \text{и} \quad \varphi_1 - \varphi_2 = 2\pi k, \quad k \in \mathbb{Z}.$$

Доказательство. Из единственности алгебраической формы имеем

$$a = r_1\cos\varphi_1 = r_2\cos\varphi_2, \quad b = r_1\sin\varphi_1 = r_2\sin\varphi_2.$$

Возводя в квадрат и складывая, получаем

$$r_1^2 = r_1^2(\cos^2\varphi_1 + \sin^2\varphi_1) = r_2^2(\cos^2\varphi_2 + \sin^2\varphi_2) = r_2^2.$$

Так как $r_1 > 0$, $r_2 > 0$, то $r_1 = r_2$. Поэтому $\cos\varphi_1 = \cos\varphi_2$, $\sin\varphi_1 = \sin\varphi_2$, следовательно, $\varphi_1 - \varphi_2 = 2\pi k$, $k \in \mathbb{Z}$. □

Следствие 2.6.3. *Если*

$$0 \neq z = a + bi \in \mathbb{C}, \quad z = r(\cos\varphi + i\sin\varphi), \quad \mathbb{R} \ni r > 0,$$

то

$$r = |z| = \sqrt{a^2 + b^2}, \quad \varphi = \arg z$$

(т. е. $\arg z = \varphi + 2\pi k$, $k \in \mathbb{Z}$).

Упражнение 2.6.4. Если $z = r(\cos\varphi + i\sin\varphi)$, $r > 0$, то

$$-z = r(\cos(\varphi + \pi) + i\sin(\varphi + \pi)),$$
$$\bar{z} = r(\cos(-\varphi) + i\sin(-\varphi)).$$

2.7. Умножение комплексных чисел в тригонометрической форме

Алгебраическая форма записи комплексных чисел удобна для операций сложения и разности. Как мы сейчас убедимся, тригонометрическая форма записи ненулевых комплексных чисел удобна для операции умножения (и как следствие — для деления, возведения в степень, извлечения корней).

Теорема 2.7.1. *Если*

$$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1), \quad z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2),$$

$r_1 > 0$, $r_2 > 0$, $r_1, r_2 \in \mathbb{R}$, *то*

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

т. е. $|z_1 z_2| = |z_1|\,|z_2|$, $\arg z_1 z_2 = (\varphi_1 + \varphi_2) + 2\pi k$ *(аргумент произведения равен сумме аргументов).*

Доказательство.

$$\begin{aligned} z_1 z_2 = (r_1 r_2)((&\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + \\ &+ i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2)) = \\ &= (r_1 r_2)(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)), \end{aligned}$$

$r_1 r_2 > 0$. Итак, это тригонометрическая форма для $z_1 z_2$, поэтому

$$|z_1 z_2| = r_1 r_2 = |z_1|\,|z_2|, \quad \arg z_1 z_2 = (\varphi_1 + \varphi_2) + 2\pi k. \quad \square$$

Следствие 2.7.2. $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}$ *для* $z_1, z_2 \in \mathbb{C}$, $z_2 \neq 0$, $\arg\left(\dfrac{z_1}{z_2}\right) = \arg z_1 - \arg z_2$. *В частности,* $|z^{-1}| = |z|^{-1}$, $\arg z^{-1} = -\arg z$.

Доказательство. Если $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$, $|z_1| = r_1$, $\arg z_1 = \varphi_1 + 2\pi k$, $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$, $|z_2| = r_2$, $\arg z_2 = \varphi_2 + 2\pi k$, то

$$r_2(\cos\varphi_2 + i\sin\varphi_2)\cdot\frac{r_1}{r_2}(\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)) = \\ = r_1(\cos\varphi_1 + i\sin\varphi_1),$$

следовательно,

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2)), \quad \frac{r_1}{r_2} > 0,$$

поэтому

$$\left|\frac{z_1}{z_2}\right| = \frac{r_1}{r_2} = \frac{|z_1|}{|z_2|}, \quad \arg\frac{z_1}{z_2} = (\varphi_1-\varphi_2) + 2\pi k, \quad k \in \mathbb{Z}.$$

Если

$$z = r(\cos\varphi + i\sin\varphi), \quad r > 0,$$

то

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

и поэтому

$$|z^{-1}| = \frac{1}{r} = |z|^{-1}, \quad \arg z^{-1} = -\varphi + 2\pi k, \quad k \in \mathbb{Z}. \quad \square$$

Следствие 2.7.3. *Умножение комплексного числа* z *на комплексное число* $r(\cos\varphi + i\sin\varphi)$, $r > 0$, *означает «растяжение» вектора* z *в* r *раз и поворот полученного вектора на угол* φ (*т. е. умножение модуля* $|z|$ *на* r, *а затем прибавление* φ *к* $\arg z$).

В частности, умножение комплексного числа на $\cos\varphi + i\sin\varphi$ *равносильно повороту на* φ (*умножение на* $i = \cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}$ *означает поворот плоскости вокруг начала координат на* $\dfrac{\pi}{2}$).

Упражнение 2.7.4 (экспоненциальная форма Эйлера записи комплексного числа). Рассмотрим последовательность

$$c_n = a_n + ib_n \in \mathbb{C},$$

где $a_n, b_n \in \mathbb{R}$. Если существуют

$$\lim_{n\to\infty} a_n = a \in \mathbb{R}, \quad \lim_{n\to\infty} b_n = b \in \mathbb{R},$$

то существует

$$\lim_{n\to\infty} (a_n + ib_n) = a + ib \in \mathbb{C}$$

(в метрике на $\mathbb{C} = \mathbb{R}^2$, определяемой $|z|$ для $z \in \mathbb{C}$).

Покажите, что

$$\lim_{n\to\infty} \left(1 + \frac{a+bi}{n}\right)^n = e^a(\cos b + i \sin b).$$

Это даёт основание (Эйлер) ввести обозначение

$$e^{a+bi} = e^a e^{bi},$$

где

$$e^{bi} = \cos b + i \sin b.$$

Если $z, w \in \mathbb{C}$, то $e^z \cdot e^w = e^{z+w}$.

2.8. Геометрическая интерпретация обратного элемента z^{-1} для $z = a + bi \in \mathbb{C}$

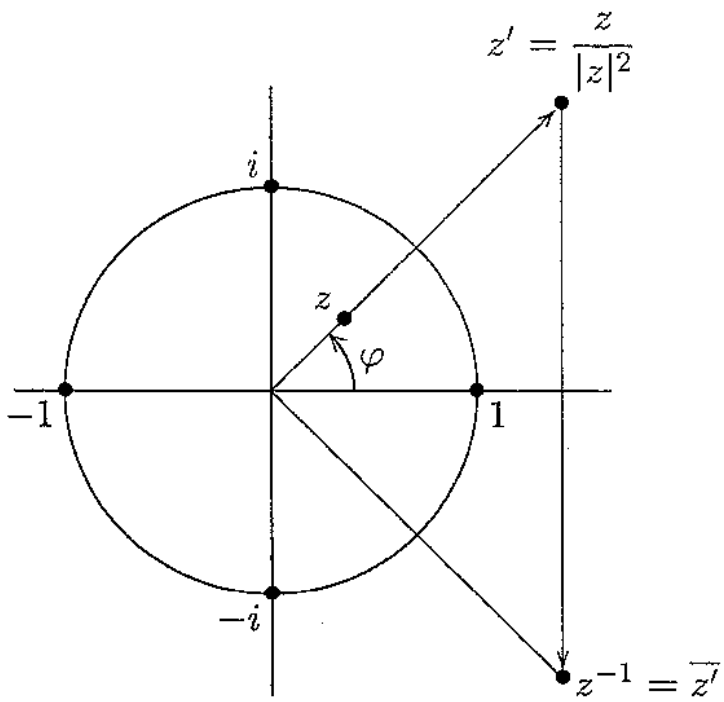
Если $0 \neq z = a + bi \in \mathbb{C}$, то, как мы видели, $z\bar{z} = N(z) = |z|^2 = a^2 + b^2$,

$$z^{-1} = \frac{\bar{z}}{|z|^2} = \overline{\left(\frac{z}{|z|^2}\right)}.$$

Таким образом, геометрическое построение комплексного числа z^{-1} можно провести двумя последовательными процедурами:

а) инверсия $z \to z' = \dfrac{z}{|z|^2}$ относительно окружности единичного радиуса ($|z'| = \dfrac{1}{|z|}$);

б) сопряжение $z' \to \overline{z'} = z^{-1}$.



Задача 2.8.1. Найти геометрическое множество точек z^{-1}, где z пробегает прямую $\{1 + bi \mid b \in \mathbb{R}\}$.

Упражнение 2.8.2.

а) Для $w \in \mathbb{C}$, $w \neq 0$, имеем

$$\left|\frac{w}{\bar{w}}\right| = \frac{|w|}{|\bar{w}|} = \frac{|w|}{|w|} = 1,$$

таким образом,

$$\frac{w}{\bar{w}} \in T = \{z \in \mathbb{C} \mid |z| = 1\}.$$

б) Если $z \in T$, т. е. $|z| = 1$, $z = \cos\varphi + i\sin\varphi$, то $z = \dfrac{w}{\bar{w}}$ для некоторого $0 \neq w \in \mathbb{C}$. Таким образом,

$$\left\{ z = \frac{1+it}{1-it} \middle| \, t \in \mathbb{R} \right\} = T.$$

Действительно, если $w = \cos\theta + i\sin\theta$, то $\bar{w} = \cos(-\theta) + i\sin(-\theta)$ и

$$\frac{w}{\bar{w}} = \cos 2\theta + i\sin 2\theta = \cos\varphi + i\sin\varphi.$$

Таким образом, если $2\theta = \varphi$, т. е. $\theta = \dfrac{\varphi}{2}$, то $w = \cos\dfrac{\varphi}{2} + i\sin\dfrac{\varphi}{2}$ является одним из решений этой задачи.

Упражнение 2.8.3.

1) Единичная окружность $T = \{z \in \mathbb{C} \mid |z| = 1\}$ с операцией умножения является группой (подгруппой мультипликативной группы $(\mathbb{C}^* = \mathbb{C} \setminus \{0\}, \cdot)$ поля $\mathbb{C}$ комплексных чисел).

2) $\{r \in \mathbb{R} \mid r < 0\} = \{z \in \mathbb{C} \mid \arg z = \pi + 2\pi k\}$.

3) Найти все $z \in \mathbb{C}$, для которых $\left|\dfrac{z-i}{z+i}\right| = 1$.

4) Найти все $z \in \mathbb{C}$, для которых $|z+i| + |z-i| = 2$.

5) Три различных комплексных числа $z_1, z_2, z_3 \in \mathbb{C} \in \mathbb{R}^2$ лежат на одной прямой в $\mathbb{R}^2$ тогда и только тогда, когда

$$\frac{z_1 - z_3}{z_2 - z_3} \in \mathbb{R}.$$

6) Четыре различных комплексных числа $z_1, z_2, z_3, z_4 \in \mathbb{C} = \mathbb{R}^2$, не лежащие на одной прямой в $\mathbb{R}^2$, лежат на одной окружности тогда и только тогда, когда их *двойное отношение* является вещественным числом:

$$\frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4} \in \mathbb{R}.$$

7) Рассмотрим отображение *инфлексии* $\mathbb{C} \to \mathbb{C}$,

$$z = x + yi \mapsto y + xi = \underline{z}, \quad x, y \in \mathbb{R}.$$

Показать, что:

(a) отображение $z \to \underline{z}$ является биекцией, при этом $\underline{(z_1 + z_2)} = \underline{z_1} + \underline{z_2}$, $|\underline{z}| = |z| = \sqrt{x^2 + y^2}$;

(b) $\underline{zw} = \dfrac{1}{i}\underline{z} \cdot \underline{w}$ для $z, w \in \mathbb{C}$;

(c) $|\underline{zw}| = |\underline{z} \cdot \underline{w}| = |\underline{z}|\,|\underline{w}|$, в частности $|z\underline{z}| = |z|^2$ для $z \in \mathbb{C}$.

Теорема 2.8.4 (формула Муавра о возведении в степень комплексного числа в тригонометрической форме). *Пусть* $0 \neq z \in \mathbb{C}$, $z = r(\cos\varphi + i\sin\varphi)$, $r > 0$, $n \in \mathbb{Z}$. *Тогда*

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

Доказательство. Утверждение теоремы — частный случай теоремы 2.7.1. □

Упражнение 2.8.5. Так как для $n \in \mathbb{N}$

$$(\cos n\varphi + i\sin n\varphi) = (\cos\varphi + i\sin\varphi)^n,$$

то, выражая правую часть с помощью формулы бинома Ньютона, получаем, приравнивая действительные и мнимые части:

$$\cos n\varphi = \cos^n\varphi - C_n^2\cos^{n-2}\varphi\sin^2\varphi + C_n^4\cos^{n-4}\varphi\sin^4\varphi - \ldots,$$
$$\sin n\varphi = n\cos^{n-1}\varphi\sin\varphi - C_n^3\cos^{n-3}\varphi\sin^3\varphi + C_n^5\cos^{n-5}\varphi\sin^5\varphi - \ldots.$$

Например:

$$\cos 2\varphi = \cos^2\varphi - \sin^2\varphi,$$
$$\cos 3\varphi = \cos^3\varphi - 3\cos\varphi\sin^2\varphi,$$
$$\cos 4\varphi = \cos^4\varphi - 6\cos^2\varphi\sin^2\varphi + \sin^4\varphi,$$
$$\sin 2\varphi = 2\cos\varphi\sin\varphi,$$
$$\sin 3\varphi = 3\cos^2\varphi\sin\varphi - \sin^3\varphi,$$
$$\sin 4\varphi = 4\cos^3\varphi\sin\varphi - 4\cos\varphi\sin^3\varphi.$$

Упражнение 2.8.6. Если

$$u = \cos\varphi + i\sin\varphi, \quad v = \bar{u} = \cos\varphi - i\sin\varphi,$$

то

$$u + v = 2\cos\varphi, \quad u - v = 2i\sin\varphi, \quad uv = 1,$$
$$u^m = \cos m\varphi + i\sin m\varphi,$$
$$v^m = (\bar{u})^m = \overline{(u^m)} = \cos m\varphi - i\sin m\varphi,$$

$$2^n \cos^n \varphi = (u+v)^n = \sum_{k=0}^{n} C_n^k u^{n-k} v^k =$$
$$= (u^n + v^n) + nuv(u^{n-2} + v^{n-2}) + C_n^2 u^2 v^2 (u^{n-4} + v^{n-4}) + \ldots.$$

Если $n = 2k$, то

$$(-1)^{n/2} 2^n \sin^n \varphi = (u-v)^n =$$
$$= (u^n + v^n) - nuv(u^{n-2} + v^{n-2}) + C_n^2 u^2 v^2 (u^{n-4} + v^{n-4}) - \ldots.$$

Если $n = 2k+1$, то

$$(-1)^{(n-1)/2} i 2^n \sin^n \varphi = (u-v)^n =$$
$$= (u^n - v^n) - nuv(u^{n-2} - v^{n-2}) + C_n^2 u^2 v^2 (u^{n-4} - v^{n-4}) - \ldots.$$

Отсюда: если $n = 2k$, то

$$2^n \cos^n \varphi =$$
$$= 2\cos n\varphi + 2n\cos(n-2)\varphi + 2C_n^2 \cos(n-4)\varphi + \ldots + C_n^{n/2},$$
$$(-1)^{n/2} 2^n \sin^n \varphi =$$
$$= 2\cos n\varphi - 2n\cos(n-2)\varphi + 2C_n^2 \cos(n-4)\varphi - \ldots + (-1)^{n/2} C_n^{n/2};$$

если $n = 2k+1$, то

$$2^n \cos^n \varphi =$$
$$= 2\cos n\varphi + 2n\cos(n-2)\varphi + 2C_n^2 \cos(n-4)\varphi + \ldots +$$
$$+ 2C_n^{(n-1)/2} \cos\varphi,$$
$$(-1)^{(n-1)/2} 2^n \sin^n \varphi =$$
$$= 2\sin n\varphi - 2n\sin(n-2)\varphi + 2C_n^2 \sin(n-4)\varphi - \ldots +$$
$$+ (-1)^{(n-1)/2} 2C_n^{(n-1)/2} \sin\varphi.$$

Упражнение 2.8.7. Если $u = \cos\varphi + i\sin\varphi$, $\varphi \neq 2\pi k$, то

$$u + \ldots + u^n = u\frac{u^n - 1}{u - 1}.$$

Приравнивая вещественные и мнимые части, получаем:

$$\sum_{k=1}^{n} \cos k\varphi = \frac{\sin\dfrac{n\varphi}{2}\cos\dfrac{(n+1)\varphi}{2}}{\sin\dfrac{\varphi}{2}};$$

$$\sum_{k=1}^{n} \sin k\varphi = \frac{\sin\dfrac{n\varphi}{2}\sin\dfrac{(n+1)\varphi}{2}}{\sin\dfrac{\varphi}{2}}.$$

Теорема 2.8.8 (извлечение корней n-й степени из комплексных чисел). *Пусть $n \geqslant 1$, $0 \neq z \in \mathbb{C}$, $z = r(\cos\varphi + i\sin\varphi)$, $r > 0$. Тогда существует ровно n различных корней n-й степени из z (таких $w \in \mathbb{C}$, что $w^n = z$):*

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

Они все лежат на окружности радиуса $\rho = \sqrt[n]{r}$, образуя вершины правильного n-угольника с аргументами

$$\frac{\varphi}{n}, \frac{\varphi+2\pi}{n} = \frac{\varphi}{n} + \frac{2\pi}{n}, \ldots, \frac{\varphi+2\pi(n-1)}{n} = \frac{\varphi}{n} + \frac{2\pi}{n}(n-1).$$

Доказательство. Будем искать решения w уравнения $w^n = z$ в тригонометрической форме:

$$w = \rho(\cos\theta + i\sin\theta), \quad \rho > 0.$$

Тогда по формуле Муавра

$$w^n = \rho^n(\cos n\theta + i\sin n\theta) = r(\cos\varphi + i\sin\varphi) = z,$$

т. е. $\rho^n = r$, и поэтому $\rho = \sqrt[n]{r}$, $n\theta = \varphi + 2\pi k$, $k \in \mathbb{Z}$. Различных корней будет ровно n при $k = 0, 1, 2, \ldots, n-1$:

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

Упражнение 2.8.9. Найдём корни уравнения

$$x^2 - (2+i)x + (-1+7i) = 0$$

(в алгебраической форме):

$$x_{1,2} = \frac{(2+i) \pm \sqrt{(2+i)^2 - 4(-1+7i)}}{2} = \frac{(2+i) \pm \sqrt{7-24i}}{2};$$

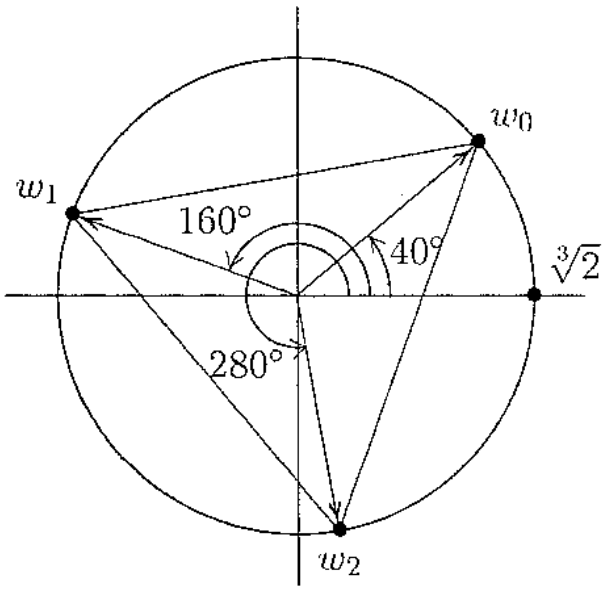
$\sqrt{7-24i} = 5\sqrt{\frac{7}{25} - \frac{24}{25}i}$, $z = \frac{7}{25} - \frac{24}{25}i = \cos\varphi + i\sin\varphi$, где $\cos\varphi = \frac{7}{25}$, $\sin\varphi = -\frac{24}{25}$, $\sqrt{z} = \pm\left(\cos\frac{\varphi}{2} + i\sin\frac{\varphi}{2}\right)$. Так как $\sin\varphi < 0$, $\cos\varphi > 0$, то $\frac{3\pi}{2} < \varphi < 2\pi$, и поэтому $\frac{3\pi}{4} < \frac{\varphi}{2} < \pi$, т. е. $\cos\frac{\varphi}{2} < 0$, $\sin\frac{\varphi}{2} > 0$, следовательно, $\cos\frac{\varphi}{2} = -\sqrt{\frac{1+\cos\varphi}{2}} = -\frac{4}{5}$, $\sin\frac{\varphi}{2} = +\frac{1-\cos\varphi}{2} = \frac{3}{5}$, $\sqrt{z} = \pm\left(-\frac{4}{5} + \frac{3}{5}i\right)$. Итак, $x_{1,2} = \frac{(2+i) \pm (-4+3i)}{2}$, $x_1 = -1+2i$, $x_2 = 3-i$.

Упражнение 2.8.10. Найти все корни третьей степени из $-1+\sqrt{3}i = 2\left(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right)$. По формуле из теоремы все три корня из $-1+\sqrt{3}i$ имеют следующий вид:

$$w_0 = \sqrt[3]{2}\left(\cos\frac{2\pi}{9} + i\sin\frac{2\pi}{9}\right);$$

$$w_1 = \sqrt[3]{2}\left(\cos\frac{8\pi}{9} + i\sin\frac{8\pi}{9}\right); \quad w_2 = \sqrt[3]{2}\left(\cos\frac{14\pi}{9} + i\sin\frac{14\pi}{9}\right).$$

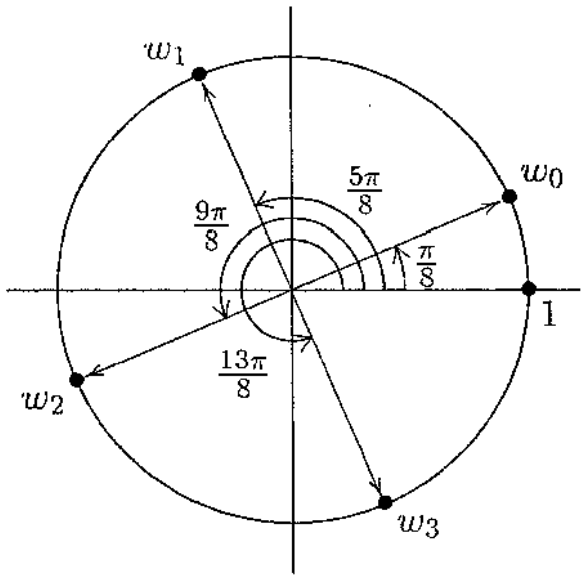
На рисунке:



Упражнение 2.8.11. Найти все корни четвёртой степени из i. Так как $i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$, по формуле из теоремы все четыре корня из i имеют следующий вид:

$$w_0 = \cos\frac{\pi}{8} + i\sin\frac{\pi}{8}; \quad w_1 = \cos\frac{5\pi}{8} + i\sin\frac{5\pi}{8};$$
$$w_2 = \cos\frac{9\pi}{8} + i\sin\frac{9\pi}{8}; \quad w_3 = \cos\frac{13\pi}{8} + i\sin\frac{13\pi}{8}.$$

На рисунке:



Упражнение 2.8.12. Извлеките все корни

$$\sqrt[6]{\frac{1-i}{\sqrt{3}+i}}.$$

Упражнение 2.8.13. Покажите, что

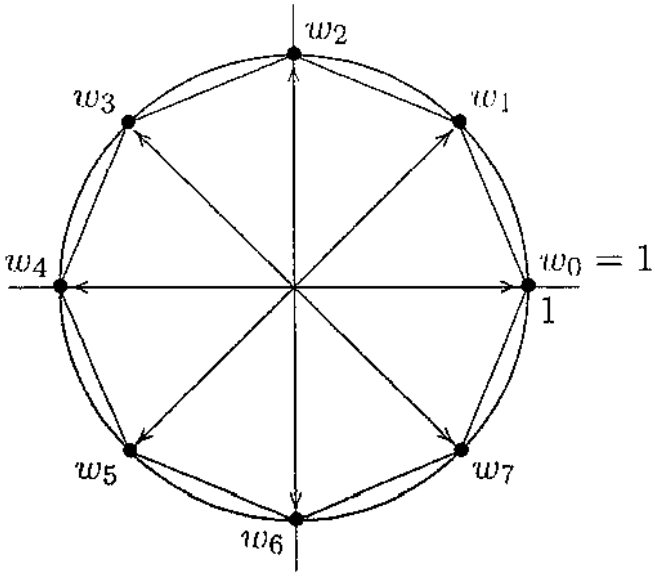
$$\sqrt[4]{-\frac{18}{1+i\sqrt{3}}} = \left\{\pm\left(\frac{3}{2} + i\frac{\sqrt{3}}{2}\right), \pm\left(\frac{\sqrt{3}}{2} - \frac{3}{2}i\right)\right\}.$$

2.9. Комплексные корни n-й степени из единицы

Так как $1 = 1(\cos 0 + i\sin 0)$, $r = 1$, $\varphi = 0$, то формула для корней n-й степени из 1 принимает вид

$$w_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}, \quad k = 0, 1, 2, \ldots, n-1.$$

Точки w_k являются вершинами правильного n-угольника, вписанного в окружность единичного радиуса с центром в начале координат, при этом одной из вершин этого многоугольника является 1. Например, при $n = 8$



Теорема 2.9.1. *Совокупность* $T_n = \{w \in \mathbb{C} \mid w^n = 1\}$ *всех* n *корней* n*-й степени из* 1 *с операцией умножения является коммутативной группой (подгруппой в* $T = \{z \mid |z| = 1\} \subset \mathbb{C}^* = \mathbb{C} \setminus \{0\}$*).*

Доказательство.

1) Если $w, z \in T_n$, т. е. $w^n = 1$, $z^n = 1$, то $(wz)^n = w^n z^n = 1 \cdot 1 = 1$, поэтому $wz \in T_n$. Таким образом, на T_n определена операция умножения (очевидно, коммутативная и ассоциативная).

2) Ясно, что $1^n = 1$, т. е. $1 \in T_n$, и 1 — нейтральный элемент в T_n.

3) Если $w \in T_n$, то $w^n = 1$,

$$\left(\frac{1}{w}\right)^n = \frac{1}{w^n} = \frac{1}{1} = 1,$$

и поэтому $w^{-1} \in T_n$. □

Замечание 2.9.2. Группа T_n является циклической, т. е. все её элементы являются степенями одного элемента, называемого циклическим образующим (в качестве одного из циклических образующих можно взять $w_1 = \cos\dfrac{2\pi}{n} + i\sin\dfrac{2\pi}{n}$, так как $w_k = (w_1)^k$ для $0 \leqslant k \leqslant n-1$, т. е. все элементы w_k группы T_n являются степенями корня w_1, такие корни называются *первообразными*). Покажите,

что $w_k = \cos\dfrac{2\pi k}{n} + i\sin\dfrac{2\pi k}{n}$ является первообразным корнем тогда и только тогда, когда наибольший общий делитель чисел k и n равен 1.

Упражнение 2.9.3. Доказать, что сумма всех k-х степеней корней уравнения $x^n = 1$ равна

$$\begin{cases} n, & \text{если } k \text{ делится на } n; \\ 0, & \text{если } k \text{ не делится на } n. \end{cases}$$

Задача 2.9.4. Если $z = \dfrac{2+i}{2-i}$, то $|z| = 1$, но z не является корнем из единицы (т. е. $z \in T \setminus T_n$ для любого $n \in \mathbb{N}$).

Задача 2.9.5. Доказать, что

а) $\sin\left(\dfrac{\pi}{2n}\right)\sin\left(\dfrac{2\pi}{2n}\right)\ldots\sin\left(\dfrac{(n-1)\pi}{2n}\right) = \dfrac{\sqrt{n}}{2^{n-1}}$;

б) $\prod\limits_{k=1}^{n}\sin\dfrac{\pi k}{2n+1} = \dfrac{\sqrt{2n+1}}{2^n}$.

Указание. Пусть

$$x_s = \varepsilon_s = \cos\frac{\pi s}{n} + i\sin\frac{\pi s}{n}, \quad s = 1, 2, \ldots, 2n$$

(все корни степени $2n$ из 1). Тогда

$$x^{2n} - 1 = \prod_{s=1}^{2n}(x - x_s) = \prod_{s=1}^{n-1}(x - x_s)\prod_{s=n+1}^{2n-1}(x - x_s)(x^2 - 1)$$

(так как $x_n = -1$, $x_{2n} = 1$). Но $x_{2n-s} = \bar{x}_s$, поэтому

$$x^{2n} - 1 = (x^2 - 1)\prod_{s=1}^{n-1}(x - x_s)(x - \bar{x}_s) = \\ = (x^2 - 1)\prod_{s=1}^{n-1}\left(x^2 - 2x\cos\frac{\pi s}{n} + 1\right).$$

Следовательно,

$$\frac{x^{2n}-1}{x^2-1} = x^{2(n-1)} + x^{2(n-2)} + \ldots + x^2 + 1 = \prod_{s=1}^{n-1}\left(x^2 - 2x\cos\frac{\pi s}{n} + 1\right).$$

Полагая $x = 1$, имеем

$$n = \prod_{s=1}^{n-1} \left(2 - 2\cos\left(\frac{\pi s}{n}\right)\right) = \prod_{s=1}^{n-1} 4\sin^2\left(\frac{\pi s}{2n}\right) =$$
$$= 2^{2(n-1)} \sin^2\left(\frac{\pi}{2n}\right) \sin^2\left(\frac{2\pi}{2n}\right) \ldots \sin^2\left(\frac{\pi(n-1)}{2n}\right).$$

Пункт б) доказывается аналогично.

2.10. Решение уравнений третьей и четвёртой степени

Любое уравнение

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0, \quad a_i \in \mathbb{C},$$

с помощью замены

$$x = y - \frac{a_{n-1}}{n}$$

(если $a_{n-1} \neq 0$) сводится к уравнению

$$y^n + b_{n-2}y^{n-2} + \ldots + b_1 y + b_0 = 0, \quad b_i \in \mathbb{C}.$$

Упражнение 2.10.1 (решение уравнений третьей степени, формула Кардано). Покажите, что для $n = 3$ все решения кубического уравнения $x^3 + px + q = 0$ $(p, q \in \mathbb{C})$ имеют вид $u + v$, где $uv = -\dfrac{p}{3}$, u^3 и v^3 — корни квадратного уравнения $z^2 + qz - \dfrac{p^3}{27} = 0$. Таким образом, для всех трёх решений имеем формулу Кардано

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

где кубические корни u и v связаны соотношением $uv = -\dfrac{p}{3}$.

Если u_1 и v_1 — какие-либо значения корней

$$\sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \quad \text{и} \quad \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

соответственно и $u_1v_1 = -\dfrac{p}{3}$, то корни находятся по правилу

$$x_1 = u_1 + v_1, \quad x_2 = u_1\omega + v_1\omega^2, \quad x_3 = u_1\omega^2 + v_1\omega,$$

где $\omega = -\dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i = \sqrt[3]{1}$.

Величина $D = -27q^2 - 4p^3$ называется *дискриминантом* многочлена $x^3 + px + q$. Условие $D = 0$ равносильно существованию кратного корня (при $D = 0$ и $p \neq 0$ имеем $x_1 = \dfrac{3q}{p}$, $x_2 = x_3 = -\dfrac{3q}{2p}$, при этом если $\dfrac{3q}{p} = -\dfrac{3q}{2p}$, то имеется корень кратности 3; если $D = 0$ и $p = 0$, то $q = 0$, а уравнение принимает вид $x^3 = 0$).

Если $p, q \in \mathbb{R}$, то: при $D > 0$ имеется три различных действительных корня; при $D < 0$ имеется один действительный и два мнимых сопряжённых корня; при $D = 0$ все корни действительные, из них хотя бы два совпадают.

Примеры 2.10.2.

1) $x^3 + 5x^2 + 2x - 8 = 0$, $x_1 = 1$, $x_2 = -2$, $x_3 = -4$.

2) $x^3 - 6ix + 4(1 - i) = 0$, $x_1 = -1 - i$, $x_2 = -1 - i$, $x_3 = 2 + 2i$.

3) $x^3 + 9x^2 + 18x + 28 = 0$, $x_1 = -7$, $x_2 = -1 - i\sqrt{3}$, $x_3 = -1 + i\sqrt{3}$.

Упражнение 2.10.3 (решение уравнений четвёртой степени; Феррари, Эйлер). Для решения уравнения

$$x^4 + px^2 + qx + r = 0 \quad (p, q, r \in \mathbb{C})$$

рассматривается соответствующее кубическое уравнение

$$y^3 + 2py^2 + (p^2 - 4r)y - q^2 = 0.$$

Если y_1, y_2, y_3 — корни этого уравнения, то все корни исходного уравнения находятся по правилу

$$x = \frac{1}{2}\left(\sqrt{y_1} + \sqrt{y_2} + \sqrt{y_3}\right),$$

где выбор квадратных корней подчинён условию

$$\sqrt{y_1}\sqrt{y_2}\sqrt{y_3} = -q.$$

Задача 2.10.4. Решить уравнения

а) $x^4 + 2x^3 + x^2 - 1 = 0$,

Ответ: $-\frac{1}{2}(1 \pm \sqrt{5})$, $-\frac{1}{2}(1 \pm i\sqrt{3})$;

б) $x^4 + 2x^3 + 2x^2 + x - 7 = 0$,

Ответ: $-\frac{1}{2}\left(1 \pm i\sqrt{2\sqrt{29}+1}\right)$, $-\frac{1}{2}\left(1 \pm \sqrt{2\sqrt{29}-1}\right)$.

Замечание 2.10.5. Отметим, что общее уравнение пятой степени неразрешимо в радикалах, при этом существует критерий разрешимости в радикалах уравнения любой степени (Абель, Галуа).

2.11. Основная теорема алгебры комплексных чисел (теорема Гаусса, 1799 г.)

Теорема 2.11.1. *Если* $f(x) \in \mathbb{C}[x]$, $\deg f(x) \geqslant 1$, *то существует корень* $c \in \mathbb{C}$ *многочлена* $f(x)$, *т. е.* $f(c) = 0$.

Доказательство.

Шаг 1 (существование абсолютного минимума вещественнозначной функции $|f(x)|$ на комплексных числах $\mathbb{C}$). Напомним, что

$$|z_1 z_2| = |z_1|\,|z_2|$$

и

$$|z_1| - |z_2| \leqslant |z_1 + z_2| \leqslant |z_1| + |z_2|$$

для $z_1, z_2 \in \mathbb{C}$.

Лемма 2.11.2. *Если* $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, $a_i \in \mathbb{C}$, $n \geqslant 1$, *то найдётся радиус* $0 < A \in \mathbb{R}$ *такой, что*

$$|f(z)| > |f(0)| \ (= |a_0|) \quad \textit{для всех} \quad z \in \mathbb{C}, \quad |z| > A$$

(это означает, что вне круга радиуса A *с центром в* 0 *значение функции* $|f(x)|$ *превосходит* $|f(0)| = |a_0|$*).*

Доказательство. Пусть $0 \neq z \in \mathbb{C}$. Тогда

$$f(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 = z^n \left(1 + \frac{a_{n-1}}{z} + \ldots + \frac{a_0}{z^n}\right),$$

и поэтому

$$\begin{aligned} |f(z)| &= |z|^n \left|1 + \left(\frac{a_{n-1}}{z} + \ldots + \frac{a_0}{z^n}\right)\right| \geqslant \\ &\geqslant |z|^n \left(1 - \left|\frac{a_{n-1}}{z} + \ldots + \frac{a_0}{z^n}\right|\right) \geqslant \\ &\geqslant |z|^n \left(1 - \frac{|a_{n-1}|}{|z|} - \ldots - \frac{|a_0|}{|z|^n}\right) = \varphi(|z|), \end{aligned}$$

где

$$\varphi(t) = t^n \left(1 - \frac{|a_{n-1}|}{t} - \ldots - \frac{|a_0|}{t^n}\right) \quad \text{для} \quad t \in \mathbb{R}.$$

Ясно, что $\lim\limits_{t \to +\infty} \varphi(t) = +\infty$, и поэтому для любого C (например, для $C = |f(0)| = |a_0|$) найдётся $\mathbb{R} \ni A > 0$ такое, что для $t > A$ имеем $\varphi(t) > C$. Итак, если $|z| = t > A$, то

$$|f(z)| \geqslant \varphi(|z|) = \varphi(t) > C = |f(0)| = |a_0|. \quad \square$$

Так как функция $|f(z)| \colon \mathbb{C} \to \mathbb{R}$ непрерывна как композиция двух непрерывных функций $\mathbb{C} \to \mathbb{C}$, $z \mapsto f(z)$, $\mathbb{C} \to \mathbb{R}$, $w \mapsto |w|$ (или если $z = u + vi$, $(u, v) \in \mathbb{R}^2$, то $f(z) = \psi_1(u, v) + \psi_2(u, v)i$, где $\psi_1(u, v)$ и $\psi_2(u, v)$ — многочлены с действительными коэффициентами от u, v, и поэтому $|f(z)| = \sqrt{\psi_1(u, v)^2 + \psi_2(u, v)^2}$ — непрерывная функция от (u, v)), то на замкнутом ограниченном множестве (компакте)

$$K = \{z \in \mathbb{C} \mid |z| \leqslant A\}$$

непрерывная функция $|f(z)|$ достигает своего минимума в точке $z_0 \in K$. В частности, $|f(z_0)| \leqslant |f(0)| = |a_0|$. Если $z \in \mathbb{C} \setminus K$, т. е. $|z| > A$, то, как мы видели,

$$|f(z_0)| \leqslant |f(0)| \leqslant |f(z)|.$$

Таким образом, в точке z_0 достигается абсолютный минимум функции $|f(z)|$ на $\mathbb{C}$.

Шаг 2. Мы покажем, что $f(z_0) = 0$, т. е. $c = z_0$ является корнем многочлена $f(x)$. Действительно, если $f(z_0) \neq 0$, то $|f(z_0)| > 0$ и, как показывает следующая лемма Даламбера, это допущение противоречит тому, что z_0 — абсолютный минимум функции $|f(x)|$.

Лемма 2.11.3 (лемма Даламбера). *Пусть* $f(x) \in \mathbb{C}[x]$, $\deg f(x) \geqslant 1$, $f(z_0) \neq 0$ *для* $z_0 \in \mathbb{C}$. *Тогда для любого* $\varepsilon > 0$ *найдётся такой элемент* $y \in \mathbb{C}$, *что* $|y| < \varepsilon$ *и* $|f(z_0 + y)| < |f(z_0)|$.

Доказательство. Если $z = z_0 + y$, т. е. $y = z - z_0$, то

$$f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n = \\ = c_0 + c_1 y + \ldots + c_{n-1} y^{n-1} + c_n y^n,$$

где $c_0 = f(z_0) \neq 0$ (при $y = 0$ имеем $z = z_0$), $c_n = 1$ (как коэффициент при y^n в $(z_0 + y)^n$).

Пусть $k > 0$ — наименьший номер слагаемого, для которого $c_k \neq 0$. Итак,

$$f(z) = c_0 + c_k y^k + c_{k+1} y^{k+1} + \ldots + c_n y^n.$$

Основное соображение заключается в том, что в окрестности точки z_0 (т. е. $y = 0$) поведение многочлена определяется первыми двумя членами $c_0 + c_k y^k$.

Сначала пусть y_0 — одно из решений уравнения $c_0 + c_k y^k = 0$ (т. е. $y_0^k = -\dfrac{c_0}{c_k}$, y_0 — один из k корней из комплексного числа $-\dfrac{c_0}{c_k}$). Если, далее, $t \in (0, 1) \subseteq \mathbb{R}$, то $c_k y_0^k = -c_0$, и поэтому

$$f(z_0 + t y_0) = c_0 + c_k t^k y_0^k + c_{k+1} t^{k+1} y_0^{k+1} + \ldots + c_n t^n y_0^n = \\ = c_0(1 - t^k) + (c_{k+1} y_0^{k+1} + \ldots + c_n t^{n-(k+1)}) t^{k+1}.$$

Если $|c_{k+1}|\,|y_0|^{k+1} + \ldots + |c_n| = M$, то

$$|f(z_0 + t y_0)| \leqslant |c_0|(1 - t^k) + M t^{k+1} = |c_0| \left(1 - t^k \left(1 - \frac{Mt}{|c_0|}\right)\right).$$

Выберем $t \in (0, 1)$ достаточно малым, так что $Mt < |c_0|$, $t|y_0| = |t y_0| < \varepsilon$. Тогда $0 < 1 - \dfrac{Mt}{|c_0|} < 1$, и поэтому

$$|f(z_0 + t y_0)| < |c_0| = |f(z_0)|, \quad |t y_0| < \varepsilon.$$

Таким образом, $y = t y_0$ удовлетворяет утверждению леммы. □

Теорема 2.11.4 (о разложении многочлена с комплексными коэффициентами в произведение линейных множителей). *Пусть* $f(x) \in \mathbb{C}[x]$, $\deg f(x) = n \geqslant 1$. *Тогда*

$$f(x) = a(x - \alpha_1) \dots (x - \alpha_n), \quad a, \alpha_1, \dots, \alpha_n \in \mathbb{C},$$

при этом это разложение единственное (*с точностью до порядка сомножителей*).

Доказательство. В силу теоремы Гаусса найдётся такое $c \in \mathbb{C}$, что $f(c) = 0$. По теореме Безу

$$f(x) = (x - c)q(x), \quad q(x) \in \mathbb{C}[x], \quad \deg q(x) = n - 1.$$

Применим далее теорему Гаусса к $q(x)$, если $n - 1 \geqslant 1$. Продолжая этот процесс, убеждаемся в существовании разложения на линейные множители.

Пусть теперь

$$f(x) = a(x - \alpha_1) \dots (x - \alpha_n) = \\ = b(x - \beta_1) \dots (x - \beta_n), \quad a, b, \alpha_i, \beta_i \in \mathbb{C}, \ a \neq 0, \ b \neq 0.$$

Ясно, что $a = b$. Если $\alpha_i \neq \beta_j$ для всех $j = 1, \dots, n$, то

$$f(\alpha_i) = 0 = b(\alpha_i - \beta_1) \dots (\alpha_i - \beta_j) \neq 0.$$

Поэтому в оба разложения входит одинаковое множество различных корней. Убедимся в совпадении кратностей вхождения каждого корня в оба разложения. Действительно, если

$$f(x) = (x - \alpha)^r q_1(x) = (x - \alpha)^s q_2(x), \quad q_1(\alpha) \neq 0, \quad q_2(\alpha) \neq 0, \quad r < s,$$

то, сокращая в $\mathbb{C}[x]$ на $(x - \alpha)^r$, получаем $q_1(x) = (x - \alpha)^{s-r} q_2(x)$, и поэтому $q_1(\alpha) = 0$, что противоречит $q_1(\alpha) \neq 0$. □

Следствие 2.11.5. *Если* $\alpha_1, \dots, \alpha_r$ *— различные корни многочлена* $f(x) \in \mathbb{C}[x]$, $k_1, \dots, k_r$ *— их кратности,* $n = \deg f(x)$, *то* $n = k_1 + \dots + k_n$ (*таким образом, многочлен степени* $n = \deg f$ *имеет ровно* n *корней с учётом их кратности*).

Замечание 2.11.6 (о неприводимых многочленах над полем комплексных чисел). По аналогии с определением простых чисел в кольце целых чисел $\mathbb{Z}$ многочлен $f(x) \in K[x]$, $\deg f(x) \geqslant 1$, называется *неприводимым*, если $f(x)$ нельзя представить в виде $f(x) = \varphi(x)\psi(x)$, $\deg \varphi(x) \geqslant 1$, $\deg \psi(x) \geqslant 1$ (иными словами, если $\varphi(x)$ — делитель многочлена $f(x)$, $\deg \varphi(x) \geqslant 1$, то $\deg \varphi(x) = n = \deg f(x)$).

Таким образом, мы установили, что *неприводимые многочлены над полем $\mathbb{C}$ комплексных чисел — это в точности многочлены первой степени*. Из единственности разложения на линейные множители над $\mathbb{C}$ получаем существование и единственность разложения на неприводимые многочлены над $\mathbb{C}$.

Лемма 2.11.7. *Если K — поле, $f(x), g(x) \in K[x]$, $\deg f(x) \leqslant n$, $\deg g(x) \leqslant n$, $f(x)$ и $g(x)$ совпадает в $(n+1)$-й различных точках $\alpha_1, \ldots, \alpha_{n+1} \in K$, то $f(x) = g(x)$.*

Доказательство. Пусть $h(x) = f(x) - g(x)$. Тогда если $h(x) \neq 0$, то $\deg h(x) \leqslant n$ и $h(\alpha_i) = f(\alpha_i) - g(\alpha_i) = 0$ для $i = 1, \ldots, n+1$. Но это противоречит тому, что число различных корней не превосходит степени многочлена. □

Следствие 2.11.8. *Если $|K| = \infty$ (в частности, для $K = \mathbb{Q}, \mathbb{R}$ или $\mathbb{C}$), то формальное и функциональное определение равенства многочленов совпадают.*

Замечание 2.11.9. Для конечного поля $\mathbb{Z}_2$ разные многочлены x и x^2 в точках 0 и 1 принимают одинаковые значения, т. е. равны как функции их $\mathbb{Z}_2$ в $\mathbb{Z}_2$.

Теорема 2.11.10 (формулы Виета). *Если K — поле, $\alpha_1, \ldots, \alpha_n \in K$,*

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = (x - \alpha_1)\ldots(x - \alpha_n),$$

то

$$\begin{aligned}
a_{n-1} &= -(\alpha_1 + \alpha_2 + \ldots + \alpha_n),\\
a_{n-2} &= \alpha_1\alpha_2 + \ldots + \alpha_{n-1}\alpha_n,\\
&\ldots\\
a_1 &= (-1)^{n-1}(\alpha_1\alpha_2\ldots\alpha_{n-1} + \ldots + \alpha_2\alpha_3\ldots\alpha_n),\\
a_0 &= (-1)^n\alpha_1\alpha_2\ldots\alpha_n.
\end{aligned}$$

Доказательство. В силу закона дистрибутивности умножение на $(x-\alpha)$ сводится к умножениям на x и на $-\alpha$. Формулы Виета получаются подсчётом коэффициента при x^k (т. е. надо при указанных раскрытиях скобок k раз выбрать x и, следовательно, $(n-k)$ раз корни). □

Упражнение 2.11.11. Пусть сумма корней многочлена с комплексными коэффициентами (считая кратность) равна нулю. Докажите, что сумма корней производной этого многочлена также равна нулю.

Упражнение 2.11.12. Пусть $x_1, \ldots, x_n$ — корни многочлена $1+x+x^2+\ldots+x^n \in \mathbb{C}[x]$. Тогда:

1) многочлен $(1+x)^{n+1}-x^{n+1}$ имеет корни $\dfrac{1}{x_1-1}, \ldots, \dfrac{1}{x_n-1}$;

2) $\dfrac{1}{x_1-1}+\dfrac{1}{x_2-1}+\ldots+\dfrac{1}{x_n-1}=-\dfrac{n}{2}$.

Глава 3

Системы линейных уравнений

В средней школе рассматривались линейные уравнения $ax = b$ и системы линейных уравнений

$$\begin{cases} ax + by = e, \\ cx + dy = f, \end{cases}$$

где $a, b, c, d, e, f \in \mathbb{R}$ — действительные числа.

В излагаемой теории систем линейных уравнений мы будем совершать с коэффициентами операции сложения и умножения, а также делить (т. е. умножать на обратный элемент) на ненулевой элемент. Таким образом, естественно рассматривать системы линейных уравнений с коэффициентами из произвольного поля K. Для понимания основных моментов теории систем линейных уравнений можно считать, что K — поле $\mathbb{R}$ действительных чисел.

Наша ближайшая цель — исследовать системы m линейных уравнений общего вида от n переменных $x_1, x_2, x_3, \ldots, x_n$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{3.1}$$

где $m, n \in \mathbb{N}$, $a_{ij}, b_i \in K$.

Таким образом, i-е уравнение, $1 \leqslant i \leqslant m$, нашей системы записывается в виде

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i$$

(a_{ij} — коэффициент при переменной x_j в i-м уравнении, b_i — свободный член i-го уравнения), или, кратко,

$$\sum_{j=1}^{n} a_{ij}x_j = b_i.$$

Прямоугольная $(m \times n)$-таблица коэффициентов $a_{ij} \in K$ (m строк, n столбцов)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}$$

называется *матрицей коэффициентов* системы линейных уравнений (3.1), а прямоугольная $(m \times (n+1))$-матрица (m строк, $n+1$ столбец)

$$\bar{A} = (a_{ij}|b_i) = \left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} & b_m \end{array}\right)$$

называется *расширенной* матрицей системы линейных уравнений (3.1) (уже полностью её определяющей).

Если $m = n$ (число уравнений равно числу переменных), то система линейных уравнений (и матрица $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$ её коэффициентов при переменных) называется *квадратной*.

В квадратной матрице

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

можно определить *диагональ* и *побочную диагональ*:

$$\begin{pmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}; \quad \begin{pmatrix} & & & a_{1n} \\ & & a_{2(n-1)} & \\ & \iddots & & \\ a_{n1} & & & \end{pmatrix}.$$

Если в системе линейных уравнений (3.1) $b_1 = \ldots = b_m = 0$, то система называется *однородной*.

3.1. Совокупность решений системы линейных уравнений

Определение 3.1.1. *Решением* системы линейных уравнений (3.1) называется строчка n элементов поля K $(l_1, \ldots, l_n)$, $l_i \in K$, такая, что при подстановке в i-е уравнение, $1 \leqslant i \leqslant m$, l_1 вместо x_1, l_2 вместо $x_2, \ldots$, l_i вместо $x_i, \ldots$, l_n вместо x_n получаем b_i (свободный член i-го уравнения), т. е.

$$\sum_{j=1}^{n} a_{ij} l_j = b_i.$$

Таким образом, строчка $(l_1, \ldots, l_n)$ является решением, если значения $l_1, \ldots, l_n$ соответственно для $x_1, \ldots, x_n$ удовлетворяют *всем* m уравнениям системы (3.1).

Через X обозначим *совокупность всех решений* системы линейных уравнений (3.1).

Замечание 3.1.2.

1) $X \subseteq K^n$ (т. е. совокупность всех решений является подмножеством в множестве K^n всех строк длины n элементов из поля K).

2) Возможно, что $X = \varnothing$ (т. е. система линейных уравнений не имеет решений), в этом случае система называется *несовместной*.

3) Если $X \neq \varnothing$ (т. е. система имеет решение), то система (3.1) называется *совместной*. Например, однородная система линейных уравнений всегда имеет нулевое решение, $(0, \ldots, 0) \in X \subseteq K^n$. Если система имеет только одно решение ($|X| = 1$), то система называется *определённой*. Если $|X| > 1$, то совместная система называется *неопределённой*. Итак, для числа решений имеются следующие возможности:

Число решений		
0	1	> 1
Система несовместная, $X = \varnothing$	Система определённая, $\lvert X\rvert = 1$	Система неопределённая, $\lvert X\rvert > 1$
Примеры		
$\begin{cases} x_1 + x_2 = 0, \\ x_1 + x_2 = 1 \end{cases}$	$\begin{cases} x_1 + x_2 = 1, \\ x_1 - x_2 = 0 \end{cases}$ $X = \left\{\left(\frac{1}{2}, \frac{1}{2}\right)\right\}$	$\begin{cases} x_1 + x_2 = 1 \end{cases}$ $x_2 = c \in K,$ $x_1 = 1 - c$ $X = \{(1-c, c) \mid c \in K\}$
$X = \varnothing$ несовместная с. л. у.	$\lvert X\rvert = 1$ определённая с. л. у.	$\lvert X\rvert = \lvert K\rvert > 1$ неопределённая с. л. у.

Основная задача исследования систем линейных уравнений (3.1) заключается в описании (нахождении) множества решений $X \subseteq K^n$ (в частности, определения, к какому типу принадлежит система (3.1): несовместная, определённая, неопределённая).

3.2. Эквивалентные системы линейных уравнений

Две системы линейных уравнений от *одного* набора $x_1, \ldots, x_n$ неизвестных и соответственно из m и p уравнений

$$(\mathbf{I}) \quad \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

$$(\mathbf{II}) \quad \begin{cases} a'_{11}x_1 + \ldots + a'_{1n}x_n = b'_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{p1}x_1 + \ldots + a'_{pn}x_n = b'_p \end{cases}$$

называются *эквивалентными*, если их множества решений $X_{\mathbf{I}}$ и $X_{\mathbf{II}}$ совпадают (т. е. подмножества $X_{\mathbf{I}}$ и $X_{\mathbf{II}}$ в K^n совпадают, $X_{\mathbf{I}} = X_{\mathbf{II}}$). Это означает, что: либо они одновременно являются пустыми подмножествами $X_{\mathbf{I}} = \varnothing = X_{\mathbf{II}}$ (т. е. обе системы **(I)** и **(II)** несовместны), либо они одновременно непустые $X_{\mathbf{I}} \neq \varnothing$, $X_{\mathbf{II}} \neq \varnothing$ и $X_{\mathbf{I}} = X_{\mathbf{II}}$ (т. е. каждое решение системы **I** является решением системы **II** и каждое решение системы **II** является решением системы **I**).

Пример 3.2.1.

1) Любые две несовместные системы от неизвестных $x_1, \ldots, x_n$ эквивалентны (в этом случае $X_1 = \varnothing = X_2$).

2) Системы

$$\textbf{(I)} \quad \begin{cases} x_1 + x_2 = 1, \\ x_1 - x_2 = 0, \end{cases} \qquad \textbf{(II)} \quad \begin{cases} 2x_1 = 1, \\ 2x_2 = 1 \end{cases}$$

эквивалентны (при $K = \mathbb{R}$), поскольку

$$X_{\mathbf{I}} = \left\{ \left(\frac{1}{2}, \frac{1}{2} \right) \right\} = X_{\mathbf{II}}.$$

3.3. Метод Гаусса

План алгоритма, предложенного Гауссом, был весьма прост:

1) применять к системе линейных уравнений последовательно преобразования, не меняющие множество решений (таким образом мы сохраняем множество решений исходной системы), и перейти к эквивалентной системе, имеющей «простой вид» (так называемую ступенчатую форму);

2) для «простого вида» системы (со ступенчатой матрицей) описать множество решений, которое совпадает с множеством решений исходной системы.

Отметим, что близкий метод «фан-чен» был известен уже в древнекитайской математике.

3.4. Элементарные преобразования систем линейных уравнений (строк матриц)

Определение 3.4.1 (элементарное преобразование 1-го типа). При $i \neq k$ к i-му уравнению системы прибавляется k-е уравнение, умноженное на число $c \in K$ (обозначение: $(i)' = (i) + c(k)$; т. е. лишь одно i-е уравнение (i) заменяется на новое уравнение $(i)' = (i)+c(k)$). Новое i-е уравнение имеет вид

$$(a_{i1} + ca_{k1})x_1 + \ldots + (a_{in} + ca_{kn})x_n = b_i + cb_k,$$

или, кратко,

$$\sum_{j=1}^{n} a'_{ij}x_j = \sum_{j=1}^{n}(a_{ij} + ca_{kj})x_j = b_i + cb_k = b'_i,$$

т. е. в новом i-м уравнении

$$a'_{ij} = a_{ij} + ca_{kj}, \quad b'_i = b_i + cb_k.$$

Определение 3.4.2 (элементарное преобразование 2-го типа). При $i \neq k$ i-е и k-е уравнение меняются местами, остальные уравнения не изменяются (обозначение: $(i)' = (k)$, $(k)' = (i)$; для коэффициентов это означает следующее: для $j = 1, \ldots, n$

$$a'_{ij} = a_{kj}, \quad b'_i = b_k; \qquad a'_{kj} = a_{ij}, \quad b'_k = b_i).$$

Замечание 3.4.3. Для удобства в конкретных вычислениях можно применять *элементарное преобразование 3-го типа:* i-е уравнение умножается на ненулевое число $0 \neq c \in K$, $(i)' = c(i)$.

Предложение 3.4.4. *Если от системы* **(I)** *мы перешли к системе* **(II)** *при помощи конечного числа элементарных преобразований 1-го и 2-го типа, то от системы* **(II)** *можно вернуться к системе* **(I)** *также элементарными преобразованиями 1-го и 2-го типа.*

Доказательство.

1) Если $i \neq k$ и $(i)' = (i) + c(k)$, то $(k)' = (k)$, $(i) = (i)' - c(k) = (i)' - c(k)'$.

2) Если $i \neq k$ и $(i)' = (k)$, $(k)' = (i)$, то $(i) = (k)'$, $(k) = (i)'$. □

Замечание 3.4.5. Утверждение верно и с включением в число элементарных преобразований элементарного преобразования 3-го типа. Если $0 \neq c \in K$ и $(i)' = c(i)$, то $0 \neq c^{-1} \in K$ и $(i) = c^{-1}(i)'$.

Теорема 3.4.6. *После последовательного применения конечного числа элементарных преобразований 1-го или 2-го типа к системе линейных уравнений получается система линейных уравнений, эквивалентная первоначальной.*

Доказательство. Заметим, что достаточно рассмотреть случай перехода от системы **I** к системе **II** при помощи одного элементарного преобразования и доказать для множеств решений включение $X_{\mathbf{I}} \subseteq X_{\mathbf{II}}$ (поскольку в силу доказанного предложения от системы **II** можно вернуться к системе **I** и поэтому будем иметь включение $X_{\mathbf{II}} \subseteq X_{\mathbf{I}}$, т. е. будет доказано равенство $X_{\mathbf{I}} = X_{\mathbf{II}}$).

Случай 1, *элементарное преобразование 1-го типа:* $(i)' = (i) + c(k)$, $i \neq k$. Пусть $l = (l_1, \ldots, l_n) \in X_{\mathbf{I}}$ — решение первой системы. Проверим, что оно удовлетворяет новому i-му уравнению:

$$\sum_{j=1}^{n} (a_{ij} + ca_{kj}) l_j = b_i + cb_k.$$

Действительно,

$$\sum_{j=1}^{n} (a_{ij} + ca_{kj}) l_j = \sum_{j=1}^{n} a_{ij} l_j + c \sum_{j=1}^{n} a_{kj} l_j = b_i + cb_k.$$

Случай 2, *элементарное преобразование 2-го типа:* $(i)' = (k)$, $(k)' = (i)$, $i \neq k$. Утверждение очевидно. □

Замечание 3.4.7. Утверждение верно и для элементарного преобразования 3-го типа: $(i)' = c(i)$, $c \neq 0$. Действительно, подставляя решение $(l_1, \ldots, l_n) \in X_{\mathbf{I}}$ в новое i-е уравнение

$$\sum_{j=1}^{n} ca_{ij} x_j = cb_i,$$

получаем

$$\sum_{j=1}^{n} ca_{ij} l_j = c\left(\sum_{j=1}^{n} a_{ij} l_j\right) = cb_i. \quad \square$$

3.5. Приведение системы линейных уравнений с помощью элементарных преобразований к ступенчатому виду

Определение 3.5.1 (определение ступенчатой матрицы (системы)). Под *ступенчатой* системой линейных уравнений понимается система линейных уравнений со *ступенчатой матрицей* коэффициентов, т. е.:

1) все нулевые строки находятся в матрице ниже ненулевых строк;

2) если $(0, \ldots, 0, a_{ik}, \ldots, a_{in})$, $a_{ik} \neq 0$ — первый ненулевой элемент в i-й строке (называемый *лидером* i-й строки), то $a_{rs} = 0$ для всех $i < r \leqslant m$, $1 \leqslant s \leqslant k$ (элементы $a_{rs} = 0$ для всех мест (r, s), расположенных в строчках, ниже i-й, и в столбцах $s = 1, 2, \ldots, k$). Другими словами, лидер строки с бóльшим номером стоит строго правее.

Определение 3.5.2. Ненулевая матрица $A \in \mathrm{M}_{m,n}(K)$ имеет главный ступенчатый вид, если матрица A имеет ступенчатый вид, все лидеры ненулевых строк $a_{1l_1}, a_{2l_2}, \ldots, a_{rl_r}$ $(1 \leqslant l_1 < \ldots < l_r \leqslant n)$ равны 1 и для каждого j, $1 \leqslant j \leqslant r$, в l_j-м столбце матрицы A единственный ненулевой элемент — это $a_{jl_j} = 1$.

Примеры 3.5.3. Матрица

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

имеет ступенчатый вид (выделены лидеры строк), но не главный ступенчатый вид.

Матрица

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

имеет главный ступенчатый вид.

Нулевая матрица имеет ступенчатый вид.

Матрица

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

не является ступенчатой (нулевая строка находится выше ненулевых строк).

Матрицы

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & \mathbf{1} \end{pmatrix}$$

не являются ступенчатыми (лидер третьей строки находится не строго правее, чем лидер второй строки).

Замечание 3.5.4. Свойство быть ступенчатой матрицей алгоритмически (с помощью компьютера) распознаваемо.

Лемма 3.5.5. *Пусть* $\alpha = (a_1, a_2, \ldots, a_n), \beta = (b_1, b_2, \ldots, b_n) \in K^n$, a_k *— лидер строки* α, b_l *— лидер строки* β, $k \leqslant l$, c_m *— лидер строки* $\alpha + \beta = (c_1, \ldots, c_n)$, $c_i = a_i + b_i$, $1 \leqslant i \leqslant n$. *Тогда:*

1) $k \leqslant m$;

2) *если* $k < m$, *то* $k = l$.

Доказательство.

1) Так как

$$\alpha = (0, \ldots, 0, a_k, \ldots, a_n), \quad a_k \neq 0,$$
$$\beta = (0, \ldots, 0, b_l, \ldots, b_n), \quad b_l \neq 0,$$

$k \leqslant l$, то

$$\alpha + \beta = (0, \ldots, 0, a_k + b_k, \ldots, a_n + b_n),$$

и поэтому $k \leqslant m$ (если $b_k = -a_k$, то $a_k + b_k = 0$, и тогда $k < m$).

2) Пусть $k < m$. Если $k < l$, то $b_k = 0$,

$$\alpha + \beta = (0, \ldots, 0, a_k, \ldots, a_n + b_n), \quad a_k \neq 0,$$

и поэтому $k = m$, что противоречит $k < m$.

Итак, $k = l$. □

Следствие 3.5.6. *Пусть* $\alpha, \beta_1, \ldots, \beta_m \in K^n$, $\alpha = (a_1, \ldots, a_n)$, $\beta_i = (b_{i1}, \ldots, b_{in})$, $1 \leqslant i \leqslant m$, $\alpha = \sum\limits_{j=1}^{m} \lambda_j \beta_j$, $\lambda_j \in K$, a_k — *лидер строки* α, b_{l_i} — *лидер строки* β_i. *Тогда* $k \geqslant \min\{l_i\}$. □

Теорема 3.5.7 (алгоритм Гаусса). *Всякую систему линейных уравнений конечным числом элементарных преобразований 1-го и 2-го типов можно привести к ступенчатому виду (т. е. к системе линейных уравнений, матрица коэффициентов которой является ступенчатой матрицей).*

Доказательство. Можно считать, что не все коэффициенты a_{ij} равны нулю и, более того, что при x_1 (т. е. в первом столбце матрицы коэффициентов) есть ненулевой элемент $a_{j1} \neq 0$ (в противном случае можно перейти к системе от переменных $x_2, \ldots, x_n$). Если $a_{11} = 0$, то, переставляя 1-е и j-е уравнения (строки расширенной матрицы) (т. е. совершая преобразование 2-го типа), приходим к случаю, когда $a'_{11} \neq 0$.

Для $i = 2, 3, \ldots, m$ последовательно проведём преобразования 1-го типа

$$(i)' = (i) - \frac{a_{i1}}{a_{11}}(1)$$

(здесь $c = -\dfrac{a_{i1}}{a_{11}}$). Тогда

$$a'_{i1} = a_{i1} - \frac{a_{i1}}{a_{11}} a_{11} = 0.$$

Рассматривая получившиеся 2-е, $\ldots$, m-е уравнения, если среди коэффициентов есть ненулевые (пусть k — первый столбец с ненулевым элементом a'_{lk} среди $a'_{2k}, \ldots, a'_{mk}$), повторим нашу процедуру: переставим второе уравнение (строку) с l-м уравнением (строкой) и обеспечим нули ниже коэффициента a'_{2k}.

Этот процесс остановится в том случае, когда все коэффициенты при переменных в оставшихся уравнениях равны нулю.

Итак, окончательная получившаяся система линейных уравнений будет иметь ступенчатый вид (т. е. матрица коэффициентов при переменных $x_1, x_2 \ldots, x_n$ будет иметь ступенчатый вид).

$$\begin{cases} \bar{a}_{11}x_1 + \ldots + \ldots + \ldots + \ldots + \ldots + \bar{a}_{1n}x_n = \bar{b}_1, \\ 0x_1 + \ldots + \bar{a}_{2k}x_k + \ldots + \ldots + \ldots + \bar{a}_{2n}x_n = \bar{b}_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0x_1 + \ldots + \ldots + \bar{a}_{sl}x_l + \ldots + \ldots + \bar{a}_{sn}x_n = \bar{b}_s, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0x_1 + \ldots + \ldots + \ldots + \bar{a}_{rt}x_t + \ldots + \bar{a}_{rn}x_n = \bar{b}_r, \\ \hline 0x_1 + \ldots + \ldots + \ldots + \ldots + \ldots + 0x_n = \bar{b}_{r+1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0x_1 + \ldots + \ldots + \ldots + \ldots + \ldots + 0x_n = \bar{b}_m. \end{cases} \tag{3.2}$$

□

Замечания 3.5.8.

1) Важный инвариант — число r уравнений в ступенчатом виде с ненулевыми коэффициентами при переменных, т. е. число «ступенек», $r \leqslant m$. Возможен случай $r = m$ (т. е. блок уравнений с нулевыми коэффициентами при $x_1, \ldots, x_n$ отсутствует). Независимость числа r от способа приведения к ступенчатому виду будет установлена позже (это — ранг матрицы коэффициентов).

2) Можно было бы продолжить процесс приведения к ступенчатому виду для расширенной матрицы системы линейных уравнений.

Следствие 3.5.9. *Всякая система линейных уравнений эквивалентна некоторой ступенчатой системе линейных уравнений.*

Следствие 3.5.10. *Каждую матрицу элементарными преобразованиями строк 1-го и 2-го типа можно привести к ступенчатому виду.*

3.6. Исследование ступенчатых систем линейных уравнений

Лемма 3.6.1. *Однородная система линейных уравнений всегда совместна.*

Доказательство. Решением системы является нулевая строчка $(0, \ldots, 0) \in K^n$. □

Лемма 3.6.2. *Если система линейных уравнений содержит уравнение*

$$0x_1 + \ldots + 0x_n = b \neq 0$$

(назовём его «экзотическим» уравнением), то система несовместна.

Доказательство. Для любой строчки $(k_1, \ldots, k_n) \in K^n$ $0 \cdot k_1 + \ldots + 0 \cdot k_n = 0 \neq b$. □

Замечание 3.6.3. Если матрица коэффициентов системы линейных уравнений нулевая (т. е. все коэффициенты равны нулю), то её совместность равносильна тому, что все свободные члены нулевые (при этом $X = K^n$). □

По ненулевой ступенчатой матрице переменные $x_1, \ldots, x_n$ разобьём на две группы: *главные* $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$, «проходящие» через уголки ступенек (их r штук), и *свободные* — все остальные $n - r$ переменных (их может и не быть совсем при $r = n$).

Лемма 3.6.4. *Если в ступенчатой системе линейных уравнений нет «экзотических» уравнений (т. е. если $r = m$ или $r < m$ и $\bar{b}_{r+1} = \ldots = \bar{b}_m = 0$), то для любого набора значений для свободных неизвестных существует (и единственный) набор значений для главных неизвестных и эти наборы дают в совокупности решение системы линейных уравнений.*

Доказательство. Так как значения для свободных неизвестных заданы, то, рассматривая r-е уравнение и перенося в правую часть уравнения члены со значениями свободных неизвестных, расположенных правее места (r, t) (если они есть), получаем уравнение (см. (3.2))

$$\bar{a}_{rt}x_t = c_r; \quad c_r \in K, \quad 0 \neq \bar{a}_{rt} \in K,$$

имеющее единственное решение для главного неизвестного

$$x_t = c_r \bar{a}_{rt}^{-1}.$$

Поднимаясь в $(r-1)$-е уравнение, повторяем этот же приём и однозначно определяем значение главного неизвестного в «уголке» $(r-1)$-го уравнения. Продолжая процесс, доходим до 1-го уравнения и определяем однозначно значение для первой главной переменной x_{i1} (в (3.2) $i_1 = 1$). Тем самым заданные значения свободных

неизвестных оказались однозначно дополнены найденными значениями главных до решения системы линейных уравнений. □

Теорема 3.6.5 (критерий совместности системы линейных уравнений по её ступенчатому виду).

1) *Система линейных уравнений* $(a_{ij}|b_i)$ *из* m *уравнений с неизвестными* $x_1, \ldots, x_n$ *совместна тогда и только тогда, когда в её ступенчатом виде нет «экзотических» уравнений* (*т. е. или* $r = m$, *или* $r < m$ *и* $\bar{b}_{r+1} = \ldots = \bar{b}_m = 0$).

2) *Для совместной системы свободным неизвестным можно придавать произвольные значения, при этом главные неизвестные однозначно определяются* (*при заданных значениях свободных неизвестных*), *тем самым мы получаем все решения системы линейных уравнений.*

Доказательство. Отметим, что исходная система и её ступенчатая системы эквивалентны.

1) а) Ясно, что совместная система не может содержать «экзотическое» уравнение (лемма 3.6.2). Таким образом, при первом появлении «экзотического» уравнения в методе Гаусса процесс надо остановить: система несовместна.

б) Если в ступенчатом виде нет «экзотических» уравнений, то утверждение следует из леммы 3.6.4.

2) Алгоритм нахождения всех решений в случае отсутствия «экзотических» уравнений рассмотрен в лемме 3.6.4. □

Следствие 3.6.6. *Система линейных уравнений несовместна тогда и только тогда, когда в её ступенчатом виде найдётся «экзотическое» уравнение.*

Теорема 3.6.7 (критерий определённости системы линейных уравнений по её ступенчатому виду). *Система линейных уравнений является определённой тогда и только тогда, когда в её ступенчатом виде:*

а) *нет «экзотических» уравнений* (*критерий совместности*);

б) $r = n$ (*т. е. все неизвестные главные, другим словами — отсутствуют свободные неизвестные*).

Доказательство.

1) При условии совместности, если $r < n$, т. е. имеется хотя бы одно свободное неизвестное, то ему можно придать как минимум два различных значения из поля K. После дополнения значений свободных переменных значениями главных переменных до решения системы мы получаем заведомо два различных решения системы, т. е. $|X| > 1$, система является неопределённой.

2) Если же при условии совместности $r = n$, т. е. нет свободных неизвестных, то главные неизвестные определяются в методе Гаусса однозначно (через свободные члены системы), таким образом, система линейных уравнений является определённой. □

Упражнение 3.6.8. Процесс приведения к ступенчатому виду можно продолжить на расширенную матрицу системы $(a_{ij}|b_i)$. Покажите, что система совместна тогда и только тогда, когда ступенчатый вид расширенной матрицы системы $(a_{ij}|b_i)$ содержит столько же ненулевых строк, сколько и ступенчатый вид матрицы (a_{ij}) (все лидеры строк ступенчатого вида расширенной матрицы находятся среди столбцов матрицы коэффициентов (a_{ij})).

Замечание 3.6.9. Любая ненулевая матрица $A \in \mathrm{M}_{m,n}(K)$ с помощью элементарных преобразований строк 1-го, 2-го и 3-го типа может быть приведена к главному ступенчатому виду. Действительно, вначале приведём матрицу A к ступенчатому виду. С помощью элементарных преобразований 3-го типа сделаем все лидеры ненулевых строк $a_{1l_1}, a_{2l_2}, \ldots, a_{rl_r}$, $1 \leqslant l_1 < l_2 < \ldots < l_r \leqslant n$, равными единице. После этого, применяя элементарные преобразования строк 1-го типа, добьёмся того, что в l_r-м столбце единственный ненулевой элемент — это $a_{rl_r} = 1$, затем аналогично добьёмся с использованием элементарных преобразований строк 1-го типа того, что единственный ненулевой элемент в l_{r-1}-м столбце — это $a_{r-1,l_{r-1}} = 1, \ldots$, в l_1-м столбце — это $a_{1l_1} = 1$ (эта процедура часто называется *обратным ходом метода Гаусса*). Таким образом, мы привели матрицу A к главному ступенчатому виду. Позже (см. 9.5.1) будет доказано, что главный ступенчатый вид матрицы определён однозначно.

Если совместная система линейных уравнений (в частности, однородная система) приведена к главному ступенчатому виду, то мы

сразу (без последовательной подстановки уже полученных выражений в предыдущие уравнения) получаем единственное выражение главных неизвестных через свободные: l-е уравнение ($1 \leqslant l \leqslant r$) главного ступенчатого вида имеет вид

$$x_{j_l} + \sum_{\substack{s=j_l+1 \\ s\neq j_{l+1},\dots,j_r}}^{n} a_{ls}x_s = \tilde{b}_l \in K,$$

и поэтому

$$x_{j_l} = \tilde{b}_l - \sum_{\substack{s=j_l+1 \\ s\neq j_{l+1},\dots,j_r}}^{n} a_{ls}x_s \tag{3.3}$$

(для однородной системы $\tilde{b}_l = 0$), в правой части присутствуют лишь свободные переменные. Таким образом, главный ступенчатый вид однородной системы *равносилен* (с заменой знака) выражению главных неизвестных через свободные (по этому ступенчатому виду).

В частном случае, при $r = n$, главный ступенчатый вид определённой системы линейных уравнений имеет форму

$$\left(\begin{array}{ccc|c} 1 & & 0 & \tilde{b}_1 \\ & \ddots & & \vdots \\ 0 & & 1 & \tilde{b}_n \\ \hline 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 \end{array}\right),$$

где $(\tilde{b}_1, \dots, \tilde{b}_n)$ — единственное решение.

3.7. Некоторые следствия из метода Гаусса

Следствие 3.7.1. *Над полем действительных чисел $K = \mathbb{R}$ (и над любым бесконечным полем) число решений системы линейных уравнений может быть равно 0 (несовместная система), 1 (определённая система) и ∞ (неопределённая система).*

Замечание 3.7.2. Над конечным полем $\mathbb{Z}_2 = \{0, 1\}$ из двух элементов система $x_1 + x_2 = 0$ имеет ровно два решения.

Следствие 3.7.3 (квадратные системы линейных уравнений).

1) *Пусть $m = n$ (т. е. число уравнений равно числу неизвестных). Тогда следующие условия эквивалентны:*

   а) *система определённая (т. е. имеет единственное решение);*

   б) *$r = n$ в ступенчатом виде (т. е. нет свободных неизвестных);*

   в) *соответствующая однородная система имеет только одно решение $(0, \ldots, 0)$.*

2) *Альтернатива Фредгольма: при $m = n$ либо система линейных уравнений определённая, либо соответствующая ей однородная система имеет ненулевое решение.*

Доказательство.

1) Если в ступенчатом виде $r = n$, то, учитывая, что $m = n$, получаем $r = n = m$. Следовательно, нет «экзотических» уравнений, и поэтому система совместна. Из критерия определённости с этим замечанием получаем, что утверждения а) и б) эквивалентны (и для однородной системы эквивалентны утверждения б) и в)).

2) С учётом 1) альтернатива Фредгольма соответствует для $r \leqslant n$ следующей альтернативе: либо $r = n$, либо $r < n$. □

3.8. Примеры применения метода Гаусса

1)

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

$$\left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 & -1 \\ 1 & -2 & 1 & 5 & 5 \end{array}\right) \to$$

$$\to \left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 4 & 4 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

В ступенчатом виде нет «экзотических» уравнений, следовательно, система совместна. Главные неизвестные — x_1, x_4, свободные неизвестные — x_2, x_3. Если $x_2 = a$, $x_3 = b$, то $x_4 = 1$, $x_1 = 1 + 2a - b - 1 = 2a - b$. Таким образом, множество решений имеет вид

$$X = \{(2a - b, a, b, 1) \mid a, b \in \mathbb{R}\}.$$

2)

$$\begin{cases} 0x_1 + x_2 + x_3 = 1, \\ 0x_1 + x_2 - x_3 = 0. \end{cases}$$

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

Система совместна, главные неизвестные — x_2, x_3, свободная неизвестная — x_1. Ясно, что $x_2 = x_3 = \dfrac{1}{2}$. Если $x_1 = a$, то множество решений имеет вид

$$X = \left\{ \left(a, \frac{1}{2}, \frac{1}{2}\right) \;\middle|\; a \in \mathbb{R} \right\}.$$

3)

$$\begin{cases} 0x_1 + x_2 - 8x_3 = -17, \\ x_1 + 0x_2 + x_3 = 10, \\ x_1 - x_2 + 0x_3 = 0. \end{cases}$$

$$\left(\begin{array}{ccc|c} 0 & 1 & -8 & -17 \\ 1 & 0 & 1 & 10 \\ 1 & -1 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 1 & 10 \\ 0 & 1 & -8 & -17 \\ 1 & -1 & 0 & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 0 & 1 & 10 \\ 0 & 1 & -8 & -17 \\ 0 & -1 & -1 & -10 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 1 & 10 \\ 0 & 1 & -8 & -17 \\ 0 & 0 & -9 & -27 \end{array}\right).$$

Система совместна (нет «экзотических уравнений»), все неизвестные x_1, x_2, x_3 главные, $x_3 = 3$, $x_2 = 7$, $x_1 = 7$. Система определённая, имеет единственное решение $(7, 7, 3)$.

4)

$$\begin{cases} x_1 + 2x_2 - 3x_3 = -2, \\ 3x_1 - x_2 + 2x_3 = 7, \\ 5x_1 + 3x_2 - 4x_3 = 2. \end{cases}$$

$$\left(\begin{array}{rrr|r} 1 & 2 & -3 & -2 \\ 3 & -1 & 2 & 7 \\ 5 & 3 & -4 & 2 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & -3 & -2 \\ 0 & -7 & 11 & 13 \\ 5 & 3 & -4 & 2 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 2 & -3 & -2 \\ 0 & -7 & 11 & 13 \\ 0 & -7 & 11 & 12 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & -3 & -2 \\ 0 & -7 & 11 & 13 \\ 0 & 0 & 0 & -1 \end{array}\right).$$

Возникло «экзотическое уравнение». Значит, система несовместна.

Глава 4

Линейное пространство строк над полем

Систематическое рассмотрение строки коэффициентов $(a_{i1}, \ldots, a_{in}) \in K^n$ i-го уравнения $a_{i1}x_1 + \ldots + a_{in}x_n = b_i$ (i-я строка матрицы $A = (a_{ij})$ коэффициентов системы линейных уравнений), строки $(a_{i1}, \ldots, a_{in}, b_i) \in K^{n+1}$ всех коэффициентов i-го уравнения (включая свободный член b_i i-й строки расширенной матрицы $\bar{A} = (a_{ij}, b_i)$ системы линейных уравнений), строки $\alpha = (\alpha_1, \ldots, \alpha_n) \in X \subseteq K^n$, являющейся решением системы линейных уравнений, с операциями сложения и умножения на элементы из поля K естественно подвело нас к определению *линейного пространства* строк K^n.

Пусть K — поле (например, $K = \mathbb{R}$ — поле действительных чисел). Рассмотрим

$$K^n = \{(\alpha_1, \ldots, \alpha_n) \mid \alpha_i \in K\} -$$

совокупность всех упорядоченных строк $\alpha = (\alpha_1, \ldots, \alpha_n)$ длины n элементов α_i, $i = 1, \ldots, n$, поля K. На множестве K^n определены следующие операции.

1) *Сложение строк* (бинарная операция): если

$$\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in K^n,$$

то

$$\alpha + \beta = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n).$$

2) Для каждого элемента $\lambda \in K$ (унарная) операция *умножение строк на элемент* $\lambda \in K$: если

$$\alpha = (\alpha_1, \ldots, \alpha_n),$$

то

$$\lambda\alpha = (\lambda\alpha_1, \ldots, \lambda\alpha_n).$$

4.1. Свойства операций

(1.1) *Ассоциативность сложения строк:* если $\alpha, \beta, \gamma \in K^n$, то $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

Действительно, на i-м месте в $(\alpha + \beta) + \gamma$ и в $\alpha + (\beta + \gamma)$ имеем $(\alpha_i + \beta_i) + \gamma_i = \alpha_i + (\beta_i + \gamma_i)$ (ассоциативность сложения в поле K). □

(1.2) *Коммутативность сложения строк:* если $\alpha, \beta \in K^n$, то $\alpha + \beta = \beta + \alpha$.

Действительно, на i-м месте в $\alpha + \beta$ и в $\beta + \alpha$ имеем $\alpha_i + \beta_i = \beta_i + \alpha_i$ (коммутативность сложения в поле K). □

(1.3) Нулевая строка $(0, \ldots, 0)$ в K^n является *нейтральным элементом* для операции сложения в K^n, поскольку $(\alpha_1, \ldots, \alpha_n) + (0, \ldots, 0) = (\alpha_1, \ldots, \alpha_n)$ для любой строки $(\alpha_1, \ldots, \alpha_n) \in K^n$. □

(1.4) Для любой строки $\alpha \in K^n$ существует противоположная строка δ такая, что $\alpha + \delta = (0, \ldots, 0)$.

Действительно, если $\alpha = (\alpha_1, \ldots, \alpha_n)$, то для $\delta = (-\alpha_1, \ldots, -\alpha_n)$ $(= (-1)\alpha)$ имеем $\alpha + \delta = (0, \ldots, 0)$. □

Таким образом, свойства (1.1)—(1.4) означают, что множество строк K^n с операцией сложения строк является *коммутативной группой*.

(2.1) Если $1 \in K$, $\alpha \in K^n$, то $\mathbf{1} \cdot \alpha = \alpha$.

Действительно, для $\alpha = (\alpha_1, \ldots, \alpha_n)$ имеем $1 \cdot \alpha = (1\alpha_1, \ldots, 1\alpha_n) = (\alpha_1, \ldots, \alpha_n) = \alpha$. □

(2.2) Если $\lambda_1, \lambda_2 \in K$, $\alpha = (\alpha_1, \ldots, \alpha_n) \in K^n$, то $\lambda_1(\lambda_2\alpha) = (\lambda_1\lambda_2)\alpha$.

Действительно, для $\alpha = (\alpha_1, \ldots, \alpha_n) \in K^n$ на i-м месте в $\lambda_1(\lambda_2\alpha)$ и в $(\lambda_1\lambda_2)\alpha$ имеем $\lambda_1(\lambda_2\alpha_i) = (\lambda_1\lambda_2)\alpha_i$ (ассоциативность умножения в поле K). □

(3.1) Если $\lambda \in K$, $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in K^n$, то $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$.

Действительно, на i-м месте в $\lambda(\alpha+\beta)$ и в $\lambda\alpha+\lambda\beta$ имеем $\lambda(\alpha_i+\beta_i)=\lambda\alpha_i+\lambda\beta_i$ (дистрибутивность в поле K). □

(3.2) Если $\lambda_1,\lambda_2\in K$, $\alpha=(\alpha_1,\ldots,\alpha_n)\in K^n$, то $(\lambda_1+\lambda_2)\alpha=$ $=\lambda_1\alpha+\lambda_2\alpha$.

Действительно, на i-м месте в $(\lambda_1+\lambda_2)\alpha$ и в $\lambda_1\alpha+\lambda_2\alpha$ имеем $(\lambda_1+\lambda_2)\alpha_i=\lambda_1\alpha_i+\lambda_2\alpha_i$ (дистрибутивность в поле K). □

Определение 4.1.1. Множество V с операцией сложения и операциями умножения на элементы λ поля K, удовлетворяющее свойствам (1.1)—(1.4), (2.1), (2.2), (3.1), (3.2), называется *линейным пространством над полем K*.

Итогом наших проверок является

Теорема 4.1.2. *Множество K^n строк длины n элементов поля K с операцией сложения и с операциями умножения на элементы λ поля K является линейным пространством над полем K.*

Определение 4.1.3. Если

$$\alpha_1=(a_{11},\ldots,a_{1n}),\ldots,\alpha_m=(a_{m1},\ldots,a_{mn})\in K^n,$$

то совокупность всех линейных комбинаций строк $\alpha_1,\ldots,\alpha_m$

$$\langle\alpha_1,\ldots,\alpha_m\rangle=\left\{\sum_{i=1}^{m}\lambda_i\alpha_i \mid \lambda_i\in K\right\}\subseteq K^n$$

называется *линейной оболочкой строк $\alpha_1,\ldots,\alpha_m$*.

Лемма 4.1.4. *Если $\alpha_1,\ldots,\alpha_m\in K^n$, то линейная оболочка $\langle\alpha_1,\ldots,\alpha_m\rangle$ является линейным пространством (подпространством в линейном пространстве строк K^n).*

Доказательство. Для $\lambda_i,\gamma_i\in K$ имеем:

$$\sum_{i=1}^{m}\lambda_i\alpha_i+\sum_{i=1}^{m}\gamma_i\alpha_i=\sum_{i=1}^{m}(\lambda_i+\gamma_i)\alpha_i\in\langle\alpha_1,\ldots,\alpha_m\rangle;$$

$$0=\sum_{i=1}^{m}0\cdot\alpha_i\in\langle\alpha_1,\ldots,\alpha_m\rangle;$$

$$-\left(\sum_{i=1}^{m}\lambda_i\alpha_i\right)=\sum_{i=1}^{m}(-\lambda_i)\alpha_i\in\langle\alpha_1,\ldots,\alpha_m\rangle.\quad\square$$

4.2. Связь решений неоднородной системы линейных уравнений с решениями соответствующей однородной системы

Пусть

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

(вообще говоря, неоднородная) совместная система линейных уравнений с множеством решений $X \subseteq K^n$,

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0 \end{cases} \quad -$$

соответствующая ей однородная система линейных уравнений с множеством решений $X_{\text{одн}} \subseteq K^n$.

Теорема 4.2.1. *Пусть* $\lambda \in K$, $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in$ $\in X_{\text{одн}}$, $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in X$. *Тогда:*

1) *если* $\alpha, \beta \in X_{\text{одн}}$, *то* $\alpha + \beta \in X_{\text{одн}}$,

2) *если* $\alpha \in X_{\text{одн}}$, *то* $\lambda\alpha \in X_{\text{одн}}$,

3) *если* $\alpha \in X_{\text{одн}}$, $u \in X$, *то* $\alpha + u \in X$,

4) *если* $u, v \in X$, *то* $v - u \in X_{\text{одн}}$.

Доказательство. Для любого i, $1 \leqslant i \leqslant n$:

1) $$\sum_{j=1}^{n} a_{ij}(\alpha_j + \beta_j) = \sum_{j=1}^{n} a_{ij}\alpha_j + \sum_{j=1}^{n} a_{ij}\beta_j = 0 + 0 = 0,$$

2) $$\sum_{j=1}^{n} a_{ij}(\lambda\alpha_j) = \lambda \sum_{j=1}^{n} a_{ij}\alpha_j = \lambda \cdot 0 = 0,$$

3) $$\sum_{j=1}^{n} a_{ij}(\alpha_j + u_j) = \sum_{j=1}^{n} a_{ij}\alpha_j + \sum_{j=1}^{n} a_{ij}u_j = 0 + b_i = b_i,$$

4) $$\sum_{j=1}^{n} a_{ij}(v_j - u_j) = \sum_{j=1}^{n} a_{ij}v_j - \sum_{j=1}^{n} a_{ij}u_j = b_i - b_i = 0. \quad \square$$

Следствие 4.2.2.

1) *Множество решений однородной системы* $X_{\text{одн}}$ *является линейным пространством (подпространством линейного пространства* K^n*).*

2) *Если* $u \in X$ *(любое частное решение совместной неоднородной системы), то*

$$X = X_{\text{одн}} + u = \{\alpha + u \mid \alpha \in X_{\text{одн}}\},$$

т. е. множество решений неоднородной системы X *является сдвигом подпространства решений однородной системы* $X_{\text{одн}}$ *на любое частное решение* $u \in X$*.*

Доказательство. Если $\alpha \in X_{\text{одн}}$, $u \in X$, то в силу 3) $\alpha + u \in X$, т. е. $X_{\text{одн}} + u \subseteq X$.

Если $v \in X$, то $v = (v - u) + u$, при этом в силу 4) $v - u \in X_{\text{одн}}$, таким образом, $X \subseteq X_{\text{одн}} + u$.

Итак, $X = X_{\text{одн}} + u$. □

Замечание 4.2.3. Позже мы покажем, что для любого линейного подпространства U линейного пространства строк K^n над полем K существует однородная система линейных уравнений, для которой $X_{\text{одн}} = U$, таким образом, любое подпространство в K^n может быть задано как пространство решений однородной системы.

Глава 5

Подстановки, перестановки

Теорема 5.0.4. *Множество* $\mathrm{S}(U)$ *всех биекций*

$$f\colon U \to U$$

с операцией произведения (композиции) отображений gf *для* $U \xrightarrow{f} U \xrightarrow{g} U$, $f, g \in \mathrm{S}(U)$, *обладает следующими свойствами:*

1) *операция произведения ассоциативна* ($h(gf) = (hg)f$ *для всех* $f, g, h \in \mathrm{S}(U)$),

2) *нейтральным элементом для этой операции является тождественное отображение* 1_U ($1_U f = f = f 1_U$ *для всех* $f \in \mathrm{S}(U)$),

3) *для всякой биекции* $f\colon U \to U$ *существует обратный элемент — биекция* $g = f^{-1}$ ($fg = 1_U = gf$).

(*Другими словами,* $\mathrm{S}(U)$ *— группа относительно операции произведения отображений;* $\mathrm{S}(U)$ *— подгруппа моноида* $\mathrm{T}(U)$: $\mathrm{S}(U) \subseteq \mathrm{T}(U)$.)

Доказательство следует из теоремы 1.6.4 и леммы 1.8.4. □

Биекции $f\colon U \to U$ множества U часто называются *подстановками*. Наиболее важный для нас случай $U = \{1, 2, \ldots, n\}$, в этом случае группу $\mathrm{S}_n = \mathrm{S}(\{1, 2, \ldots, n\})$ называют *группой подстановок* множества $\{1, 2, \ldots, n\}$ из n элементов (иногда называемой *симметрической группой*).

5.1. Запись подстановок. Перестановки

Если $f \in \mathrm{S}_n$ — подстановка, то рассмотрим её *каноническую запись*

$$\begin{pmatrix} 1 & 2 & \dots & n \\ f(1) & f(2) & \dots & f(n) \end{pmatrix}.$$

В нижней строчке $(f(1), f(2), \dots, f(n))$, поскольку f — биекция, встречаются все элементы i, $1 \leqslant i \leqslant n$, при этом только по одному разу. Такие строчки элементов $(i_1, \dots, i_n)$, $1 \leqslant i_j \leqslant n$, где каждый элемент i_j, $1 \leqslant i_j \leqslant n$, встречается один и только один раз, называются *перестановками* элементов $1, 2, \dots, n$.

Лемма 5.1.1. *Число всех перестановок $(i_1, \dots, i_n)$ из n элементов равно $n! = 1 \cdot 2 \cdot \ldots \cdot n$.*

Доказательство. Для i_1 имеем n возможностей. При выбранном i_1 для i_2 имеем $(n-1)$ возможность. Таким образом, число различных перестановок равно $n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n!$. □

Лемма 5.1.2. *Число различных подстановок множества $\{1, 2, \dots, n\}$ равно $n!$ (т. е. $|\mathrm{S}_n| = n!$).*

Доказательство. Для $f \in \mathrm{S}_n$ рассмотрим *каноническую запись*

$$f = \begin{pmatrix} 1 & 2 & \dots & n \\ f(1) & f(2) & \dots & f(n) \end{pmatrix}.$$

Таким образом, различных подстановок столько же, сколько различных перестановок n элементов, т. е. $n!$. □

Во многих случаях удобно рассматривать записи подстановки $f \in \mathrm{S}_n$, располагая в верхней строчке произвольную перестановку $(i_1, i_2, \dots, i_n)$:

$$f = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ f(i_1) & f(i_2) & \dots & f(i_n) \end{pmatrix}.$$

Каждый столбец этой таблицы имеет вид

$$\begin{pmatrix} i \\ f(i) \end{pmatrix}.$$

Пример 5.1.3.

1) Для тождественной подстановки в S_2 имеем

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

Для биекции $f\colon \{1,2\} \to \{1,2\}$, $f(1) = 2$, $f(2) = 1$, имеем

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

2) Если

$$f = \begin{pmatrix} i_1 & \dots & i_n \\ j_1 & \dots & j_n \end{pmatrix} \in S_n,$$

то

$$f^{-1} = \begin{pmatrix} j_1 & \dots & j_n \\ i_1 & \dots & i_n \end{pmatrix}. \quad \square$$

3) Так как $(\sigma\tau)(i) = \sigma(\tau(i))$, то

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

В частности,

$$\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}. \quad \square$$

4) Обозначим через $(i_1\ i_2\ \dots\ i_r)$ *цикл длины* r в группе подстановок S_n: подстановку, переводящую i_k в i_{k+1} для $1 \leqslant k \leqslant r-1$, i_r в i_1, и оставляющую все элементы из $\{1, 2, \dots, n\}$, отличные от $i_1, \dots, i_r$, на месте. Тогда в S_3 имеем шесть подстановок:

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}; \qquad (1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix};$$

$$(1\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}; \qquad (2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix};$$

$$(1\ 2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}; \qquad (1\ 3\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

При этом в S_n для $n \geqslant 3$ имеем

$$(1\ 2)(1\ 3) = (1\ 3\ 2) \neq (1\ 2\ 3) = (1\ 3)(1\ 2),$$

следовательно, группа S_3 и любая группа S_n при $n \geqslant 3$ некоммутативны. Так как $S_1 = \{e\}$ и $S_2 = \{e, (1\ 2)\}$ — коммутативные группы, то получаем, что группа S_n коммутативна тогда и только тогда, когда $n = 1$ или $n = 2$.

5.2. Перестановки и транспозиции

Рассмотрим перестановку двух элементов i и j, $i \neq j$, в перестановке $(i_1, \ldots, i_n)$ (все остальные элементы, отличные от i, j, остаются на своих местах). Эта процедура называется *транспозицией* перестановки $(i_1, \ldots, i_n)$.

Лемма 5.2.1.

1) *Умножение слева* $(i\ j)f$ *подстановки*

$$f = \begin{pmatrix} i_1 & \ldots & i_n \\ j_1 & \ldots & j_n \end{pmatrix}$$

на цикл $(i\ j)$ *длины* 2 *приводит к транспозиции элементов* i *и* j *в нижней строке* (*перестановке*) $(j_1, \ldots, j_n)$.

2) *Умножение справа* $f(i\ j)$ *подстановки*

$$f = \begin{pmatrix} i_1 & \ldots & i_n \\ j_1 & \ldots & j_n \end{pmatrix}$$

на цикл $(i\ j)$ *длины* 2 *приводит к транспозиции элементов* i *и* j *в верхней строке* (*перестановке*) $(i_1, \ldots, i_n)$.

Доказательство.
1)

$$\begin{pmatrix} \ldots & i & \ldots & j & \ldots \\ \ldots & j & \ldots & i & \ldots \end{pmatrix} \begin{pmatrix} i_1 & \ldots & i_k & \ldots & i_l & \ldots & i_n \\ j_1 & \ldots & j_k = i & \ldots & j_l = j & \ldots & j_n \end{pmatrix} =$$

$$= \begin{pmatrix} i_1 & \ldots & i_k & \ldots & i_l & \ldots & i_n \\ j_1 & \ldots & j = j_l & \ldots & i = j_k & \ldots & j_n \end{pmatrix}.$$

2)

$$\begin{pmatrix} i_1 & \dots & i_r = i & \dots & i_s = j & \dots & i_n \\ j_1 & \dots & j_r & \dots & j_s & \dots & j_n \end{pmatrix} \begin{pmatrix} \dots & i & \dots & j & \dots \\ \dots & j & \dots & i & \dots \end{pmatrix} =$$

$$= \begin{pmatrix} i_1 & \dots & j = i_s & \dots & i = i_r & \dots & i_n \\ j_1 & \dots & j_r & \dots & j_s & \dots & j_n \end{pmatrix}. \quad \square$$

Лемма 5.2.2 (о списке перестановок). *Все $n!$ перестановок из n элементов $\{1, 2, \dots, n\}$ можно расположить в список, начиная с произвольной перестановки $(i_1, i_2, \dots, i_n)$, так, что каждая следующая перестановка в этом списке получается из предыдущей с помощью некоторой транспозиции двух элементов.*

Доказательство. Проведём индукцию по n. Начало индукции $n = 2$, $n! = 2$, наши списки:

$$\begin{matrix} (1,2) \\ (2,1) \end{matrix} \quad \begin{matrix} (2,1) \\ (1,2) \end{matrix}$$

Пусть наше утверждение верно для всех k, $k < n$. Пользуясь этим, создадим первый блок из различных $(n-1)!$ перестановок с i_1 на первом месте (т. е. перестановок из элементов $\{i_2, \dots, i_n\}$), при этом каждая следующая перестановка получается из предыдущей с помощью транспозиции:

$$(n-1)! \begin{cases} (i_1, i_2, \dots, i_n), \\ \quad \vdots \\ (i_1, \dots, i_2, \dots). \end{cases}$$

Совершая транспозицию i_1 и i_2 в последней перестановке первого блока и повторяя наше рассуждение, построим второй блок из различных $(n-1)!$ перестановок с i_2 на первом месте (т. е. перестановок элементов $\{i_1, i_3, \dots, i_n\}$), при этом каждая следующая перестановка получается из предыдущей применением транспозиции:

$$(n-1)! \begin{cases} (i_2, \dots), \\ \quad \vdots \\ (i_2, \dots, i_3, \dots). \end{cases}$$

Продолжая этот процесс, получим n блоков из $(n-1)!$ перестановок каждый, всего $n!$ перестановок. Они все различны: в одном блоке по индуктивному предположению, в разных блоках перестановки различаются на первом месте. Таким образом, в этом списке присутствуют все $n!$ перестановок из n элементов, при этом каждая следующая получается из предыдущей с помощью одной транспозиции. □

Следствие 5.2.3. *От любой перестановки $(i_1, \ldots, i_n)$ можно перейти к любой другой перестановке $(j_1, \ldots, j_n)$ с помощью конечного числа транспозиций.*

Доказательство. В списке с началом $(i_1, \ldots, i_n)$ надо найти перестановку $(j_1, \ldots, j_n)$. □

Следствие 5.2.4. *Каждая подстановка*

$$\tau = \begin{pmatrix} 1 & 2 & \ldots & n \\ k_1 & k_2 & \ldots & k_n \end{pmatrix} \in \mathrm{S}_n$$

является произведением $\tau = \tau_r \ldots \tau_1$ конечного числа циклов τ_i длины два (называемых также транспозициями). Таким образом, циклы длины два (транспозиции) дают одну из систем образующих группы S_n.

Доказательство. Составим список перестановок, начинающийся с перестановки $(1, 2, \ldots, n)$, в котором каждая l-я перестановка получается из $(l-1)$-й транспозицией элементов i_{l-1} и j_{l-1}, и найдём в нём нашу перестановку $(k_1, \ldots, k_n)$ из канонической записи подстановки τ (пусть она занимает $(r+1)$-е место). Тогда (по лемме об умножении слева на цикл длины два)

$$(i_r\ j_r) \ldots (i_1\ j_1) \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \ldots & n \\ k_1 & k_2 & \ldots & k_n \end{pmatrix} = \tau,$$

т. е. $\tau = \tau_r \ldots \tau_1$, где $\tau_r = (i_r\ j_r), \ldots, \tau_1 = (i_1\ j_1)$. □

Замечание 5.2.5. Ясно, что представление подстановки $\tau = \tau_1 \ldots \tau_r$ в виде произведения транспозиций возможно разными способами (например, $(1\ 2) = (1\ 2)^3$).

5.3. Разложение подстановок в произведение циклов с непересекающимися орбитами

Орбитой цикла $(i_1\ i_2\ \dots\ i_r)$ назовём множество $\{i_1, \dots, i_r\}$.

Если $\sigma \in \mathrm{S}_n$ — подстановка символов $\{1, 2, \dots, n\}$ и $a \in \mathbb{N}$, $1 \leqslant a \leqslant n$, то рассмотрим последовательность

$$a, \sigma(a), \sigma^2(a) = \sigma(\sigma(a)), \dots, \sigma^r(a), \dots$$

(орбиту элемента a). Из конечности множества $\{1, 2, \dots, n\}$ следует, что найдутся такие натуральные числа t и s, $t < s$, что $\sigma^t a = \sigma^s a$. В группе S_n рассмотрим σ^{-1}. Применяя $(\sigma^{-1})^t$ к этому равенству, получим $a = \sigma^{s-t}(a)$, $r = s - t > 0$. Рассмотрим самое маленькое такое натуральное число r (со свойством $a = \sigma^r(a)$, при этом все r элементов $\{a, \sigma(a), \dots, \sigma^{r-1}(a)\}$ различны). Итак, получили цикл $(a\ \sigma(a)\ \dots\ \sigma^{r-1}(a))$ длины r. Выбирая элемент b вне этого цикла (если $r < n$), получаем цикл $(b\ \sigma(b)\ \dots\ \sigma^{r'-1}(b))$ длины r', при этом орбиты этих циклов не пересекаются. Продолжим этот процесс. Заметим, что циклы с непересекающимися орбитами перестановочны. Единственность этого разложения следует из инвариантности определения орбиты. Итак, получаем следующее утверждение.

Теорема 5.3.1. *Каждая подстановка $\tau \in \mathrm{S}_n$ разлагается (и притом единственным образом) в произведение циклов с непересекающимися орбитами (поэтому эти циклы перестановочны друг с другом).*

Замечание 5.3.2.

1) В практических задачах удобно начинать с $a = 1$, затем число b выбирать как наименьшее число, не вошедшее в $\{a, \sigma(a), \dots, \sigma^{r-1}(a)\}$, и т. д.

2) Как правило, циклы длины 1 (т. е. неподвижные элементы) опускают в записи циклового разложения подстановки.

Упражнение 5.3.3.

1) Пусть $\sigma, \tau \in \mathrm{S}_n$. Подстановка $\tau\sigma\tau^{-1}$ называется подстановкой, *сопряжённой* с подстановкой σ (с помощью подстановки τ). Проверьте, что *отношение сопряжённости* является от-

ношением эквивалентности. Соответствующее разбиение множества S_n на классы эквивалентных подстановок называется *разбиением* на *классы сопряжённых элементов*.

2) Доказать, что подстановки $\gamma, \sigma \in S_n$ сопряжены тогда и только тогда, когда γ и σ имеют *одинаковое цикловое разложение* (т. е. одинаковое число циклов каждой длины в своих разложениях в произведение циклов с непересекающимися орбитами).

Указания.

а) $\tau(\sigma_1\sigma_2)\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1})$.

б) Если $\sigma = (i_1, \ldots, i_r)$ — цикл длины r, то $\tau\sigma\tau^{-1} = (\tau(i_1), \ldots, \tau(i_r))$.

Пример 5.3.4. Пусть

$$\delta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 8 & 1 & 2 & 4 & 7 & 5 & 3 & 6 & 10 \end{pmatrix},$$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 1 & 6 & 10 & 2 & 4 & 9 & 7 & 3 & 8 \end{pmatrix}.$$

Требуется найти $(\delta\sigma)^{100}$.

Сначала находим

$$\delta\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 9 & 7 & 10 & 8 & 2 & 6 & 5 & 1 & 3 \end{pmatrix} = (5\ 8)(1\ 4\ 10\ 3\ 7\ 6\ 2\ 9)$$

(разложение в произведение циклов с непересекающимися орбитами). Поэтому

$$(\delta\sigma)^{100} = (5\ 8)^{100}(1\ 4\ 10\ 3\ 7\ 6\ 2\ 9)^{100}.$$

Так как $(5\ 8)^2$ и $(1\ 4\ 10\ 3\ 7\ 6\ 2\ 9)^8$ — тождественные подстановки, $100 = 12 \cdot 8 + 4$, то

$$(\delta\sigma)^{100} = (1\ 4\ 10\ 3\ 7\ 6\ 2\ 9)^4 =$$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 10 & 9 & 6 & 5 & 4 & 1 & 8 & 3 & 2 \end{pmatrix} = (4\ 6)(3\ 9)(2\ 10)(1\ 7).$$

Задача 5.3.5. Найти разбиение на классы сопряжённых элементов для групп S_3, S_4, S_5.

Задача 5.3.6.

1) Группа S_n порождается транспозициями

$$(1\ 2), (1\ 3), \ldots, (1\ n)$$

(т. е. любой элемент группы S_n является произведением этих транспозиций).

Указание. Если $i, j \neq 1$, то

$$(i\ j) = (1\ i)(1\ j)(1\ i).$$

2) Группа S_n, $n \geqslant 3$, порождается транспозицией $(1\ 2)$ и циклом $(1\ 2\ \ldots\ n)$.

5.4. Чётность перестановок и подстановок

Будем говорить, что числа i и j в перестановке $(\ldots, i, \ldots, j, \ldots)$ образуют *инверсию*, если число i расположено левее, чем j, но $i > j$ (в противном случае будем говорить, что числа i и j расположены в *правильном порядке*). Ясно, что сумма числа всех инверсией и числа всех порядков в любой перестановке из n чисел $1, 2, \ldots, n$ равна $C_n^2 = \dfrac{n(n-1)}{2}$.

Пример 5.4.1. Число инверсий в перестановке $(1, 2, \ldots, n)$ равно нулю, в перестановке $(n, n-1, \ldots, 2, 1)$ равно $\dfrac{n(n-1)}{2}$.

Удобный *алгоритм подсчёта числа инверсий:* считаем, сколько инверсий образует 1 (все числа, находящиеся левее), после чего вычёркиваем 1 и переходим к 2 и т. д.

Теорема 5.4.2. *Транспозиция в перестановке меняет чётность числа инверсий.*

Доказательство. Рассмотрим транспозицию элементов i и j:

$$(\ldots, i, \ldots, j, \ldots) \mapsto (\ldots, j, \ldots, i, \ldots).$$

Сначала рассмотрим случай «соседей»:

$$(\ldots, i, j, \ldots) \mapsto (\ldots, j, i, \ldots).$$

Так как при перестановке чисел i и j их отношение с числами, расположенными левее (как и правее) не изменяется, то число инверсий изменяется на единицу (т. е. ± 1), следовательно, чётность числа инверсий изменяется.

Если же между числами i и j находится k элементов, то последовательно переставляя i с правыми соседними элементами k раз, потом с j, затем переставляя k раз элемент j с левыми соседними элементами, мы, проведя $k + 1 + k = 2k + 1$ транспозиций соседних элементов, осуществим транспозицию чисел i и j. Таким образом, чётность изменилась. □

Следствие 5.4.3. *Число чётных перестановок при $n \geqslant 2$ равно числу нечётных перестановок и равно $\dfrac{n!}{2}$.*

Доказательство. Расположив все $n!$ перестановок, начиная, например, с $(1, 2, \ldots, n)$, в список, в котором каждая следующая перестановка получается из предыдущей одной транспозицией, мы видим, что чётные перестановки чередуются с нечётными, поэтому число чётных перестановок равно числу нечётных и равно $\dfrac{n!}{2}$. □

Чётность подстановки

$$\begin{pmatrix} i_1 & \ldots & i_n \\ j_1 & \ldots & j_n \end{pmatrix}$$

определяется как чётность суммы числа инверсий в верхней строчке и числа инверсий в нижней строчке.

Предложение 5.4.4. *Чётность подстановки $\sigma \in \mathrm{S}_n$ не зависит от её записи.*

Доказательство. Если

$$\sigma = \begin{pmatrix} i_1 & \dots & i_n \\ \sigma(i_1) & \dots & \sigma(i_n) \end{pmatrix} = \begin{pmatrix} i'_1 & \dots & i'_n \\ \sigma(i'_1) & \dots & \sigma(i'_n) \end{pmatrix} -$$

две записи подстановки $\sigma \in \mathrm{S}_n$, то, переходя конечным числом транспозиций от перестановки $(i_1, \dots, i_n)$ к перестановке $(i'_1, \dots, i'_n)$, переставляя при этом соответствующие «столбики»

$$\begin{pmatrix} i \\ \sigma(i) \end{pmatrix},$$

приходим от нижней строчки $(\sigma(i_1), \dots, \sigma(i_n))$ к строчке $(\sigma(i'_1), \dots, \sigma(i'_n))$. Перестановка двух «столбиков» является транспозицией в верхней и в нижней строчках, следовательно, меняется чётность в верхней и в нижней строчках, в итоге чётность суммы числа транспозиций в верхней и в нижней строчке при перестановке двух «столбиков» не изменится. □

Замечание 5.4.5. Подстановка, обратная к чётной подстановке, чётная. Действительно, если

$$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$$

чётная, то

$$\sigma^{-1} = \begin{pmatrix} j_1 & \dots & j_n \\ i_1 & \dots & i_n \end{pmatrix}$$

чётная. □

5.5. Чётность произведения подстановок

Возможность использовать произвольную запись подстановки удобна для рассмотрения произведения:

$$\begin{pmatrix} i_1 & \dots & i_n \\ j_1 & \dots & j_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix},$$

откуда следует

Лемма 5.5.1 (о чётности произведения).

σ	τ	$\sigma\tau$
Ч	Ч	Ч
Н	Ч	Н
Ч	Н	Н
Н	Н	Ч

. □

Рассмотрим отображение

$$\varepsilon\colon \mathrm{S}_n \to \{1, -1\},$$

$$\varepsilon(\sigma) = \begin{cases} 1, & \text{если } \sigma \text{ — чётная подстановка,} \\ -1, & \text{если } \sigma \text{ — нечётная подстановка.} \end{cases}$$

Замечание 5.5.2. Напомним, что $\{1, -1\}$ — коммутативная группа относительно операции произведения. Действительно, произведение является операцией на $\{1, -1\}$; эта операция ассоциативна и коммутативна; 1 — нейтральный элемент; $(1)^{-1} = 1$, $(-1)^{-1} = -1$.

Следствие 5.5.3. *Если $\sigma, \tau \in \mathrm{S}_n$, то:*

$$\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$$

(т. е. $\sigma\colon \mathrm{S}_n \to \{1, -1\}$ — гомоморфизм групп);

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

Следствие 5.5.4. *Если $\sigma = \tau_1 \dots \tau_k$ — разложение подстановки $\sigma \in \mathrm{S}_n$ в произведение транспозиций $\tau_1, \dots, \tau_k$, то $\varepsilon(\sigma) = (-1)^k$.*

Доказательство. Отметим только, что если $\tau = (i\ j)$ — транспозиция, то $\varepsilon((i\ j)) = -1$. □

Упражнение 5.5.5. $\varepsilon((i_1\ \dots\ i_r)) = (-1)^{r-1}$ для цикла $(i_1\ \dots\ i_r)$ длины r, $r \geqslant 2$.

Теорема 5.5.6. *Чётные подстановки A_n являются группой (подгруппой в группе подстановок S_n); $|\mathrm{A}_n| = \dfrac{n!}{2}$ при $n \geqslant 2$.*

Доказательство. Так как произведение $\sigma\tau$ чётных подстановок $\sigma, \tau \in \mathrm{A}_n$ является чётной подстановкой, то имеем операцию произведения на множестве A_n, которая ассоциативна. Тождественная подстановка чётная и является нейтральным элементом в A_n. Если $\sigma \in \mathrm{A}_n$, то мы уже отметили, что $\sigma^{-1} \in \mathrm{A}_n$. □

Задача 5.5.7. Найти разбиение в классы сопряжённых элементов групп A_4, A_5.

Задача 5.5.8. Группа A_n, $n \geqslant 3$, порождается тройными циклами (любой элемент группы A_n является произведением тройных циклов и обратных к ним; обратный элемент к тройному циклу сам является тройным циклом).

Указание. Чётная подстановка может быть представлена в виде произведения чётного числа транспозиций, при различных i, j, k

$$(i\ k)(i\ j) = (i\ j\ k),$$

при различных i, j, k, l

$$(i\ j)(k\ l) = (j\ k\ l)(i\ l\ j). \quad \square$$

Глава 6

Определители квадратных матриц

6.1. Определители малых порядков

Рассматривая систему линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{cases}$$

для вычисления x_1 умножим первое уравнение на a_{22}, второе уравнение на $-a_{12}$ и сложим их. Получим

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}.$$

Аналогично, для вычисления x_2 умножим первое уравнение на $-a_{21}$, второе уравнение на a_{11} и сложим их. Получим

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1.$$

Если мы *определителем (2×2)-матрицы*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

назовём число

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

то в этом частном случае мы получим следующее утверждение (правило Крамера для $n = 2$): если определитель квадратной системы отличен от нуля, т. е.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \neq 0,$$

то система является определённой и для её единственного решения справедливы формулы

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Непосредственная проверка показывает, что (x_1, x_2) — решение.

Упражнение 6.1.1. Проделать аналогичную процедуру в случае $n = 3$.

Замечание 6.1.2. Очевидно, что определители второго порядка обладают следующими свойствами:

1) $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1;$

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

3) $\begin{vmatrix} ca_{11} & ca_{12} \\ a_{21} & a_{22} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$

аналогично для второй строки;

4) если $(a_{11}, a_{12}) = (b_1, b_2) + (c_1, c_2)$, то

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} b_1 & b_2 \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} c_1 & c_2 \\ a_{21} & a_{22} \end{vmatrix};$$

аналогично для второй строки;

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

Наша ближайшая цель — построить общую теорию определителей квадратных $(n \times n)$-матриц и привести многочисленные приложения определителей, в частности в системах линейных уравнений.

Отметим, что на начальном периоде теория определителей формировалась параллельно с аксиоматической теорией площадей и объёмов. Например, в декартовой системе координат на плоскости определитель

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

равен (ориентированной) площади параллелограмма, построенного на векторах (a_{11}, a_{12}) и (a_{21}, a_{22}).

6.2. Определители квадратных $(n \times n)$-матриц

Пусть

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots\dots\dots\dots & & \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathrm{M}_n(K) -$$

квадратная $(n \times n)$-матрица, $a_{ij} \in K$, где K — любое поле (например, $K = \mathbb{R}$).

При $n = 1$: $|a| = a \in K$.

При $n = 2$ мы имеем

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

т. е. определитель (2×2)-матрицы является суммой двух слагаемых, каждое из которых является произведением элементов матрицы, взятых по одному (и только одному) из каждой строки (столбца), при этом знак определяется чётностью соответствующей подстановки индексов:

$$+ a_{11}a_{22}, \quad \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \text{ — чётная подстановка;}$$

$$- a_{12}a_{21}, \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \text{ — нечётная подстановка.}$$

С этой «подсказкой» определим *определитель* квадратной матрицы A как

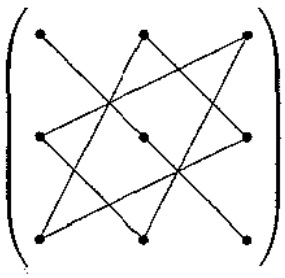
$$|A| = \sum_{\alpha \in \mathrm{S}_n} \varepsilon(\alpha) a_{1\alpha(1)} \dots a_{n\alpha(n)},$$

т. е. как сумму всех произведений элементов матрицы A, взятых по одному (и только одному) из каждой строки и каждого столбца ($a_{1\alpha(1)}$ — из 1-й строки и $\alpha(1)$-го столбца; ...; $a_{n\alpha(n)}$ — из n-й строки и $\alpha(n)$-го столбца), т. е. тех произведений, индексы которых дают подстановку $\alpha \in \mathrm{S}_n$, при этом эти произведения берутся со знаком $+$ ($\varepsilon(\alpha) = 1$), если подстановка α чётная, и со знаком $-$ ($\varepsilon(\alpha) = -1$), если подстановка α нечётная.

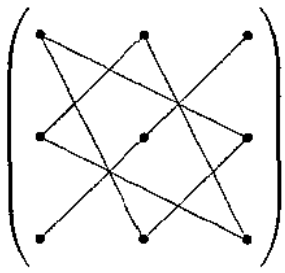
Упражнение 6.2.1. Если $n = 3$, $A = (a_{ij}) \in \mathrm{M}_3(K)$, то

$$|A| = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - \\ - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Мнемоническое правило: три произведения



входят со знаком $+$; три произведения



входят со знаком $-$.

Упражнение 6.2.2. При $n = 3$, $A = (a_{ij}) \in \mathrm{M}_3(\mathbb{R})$ в декартовой системе координат в $\mathbb{R}^3$ определитель $|A|$ матрицы A равен ориентированному объёму параллелепипеда, построенного на векторах (a_{11}, a_{12}, a_{13}), (a_{21}, a_{22}, a_{23}) и (a_{31}, a_{32}, a_{33}).

Упражнение 6.2.3. Если $A = (a_{ij}) \in \mathrm{M}_3(\mathbb{R})$, то все шесть слагаемых в разложении определителя третьего порядка $|A|$ одновременно не могут быть положительны.

6.3. Свойства определителя. Базовые свойства 1—4

Свойство 1. *Если*

$$E = E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

то

$$|E| = 1.$$

Доказательство следует из следующего утверждения. □

Лемма 6.3.1.

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11}a_{22}\dots a_{nn}.$$

Доказательство. Следует рассмотреть только те произведения, входящие в определитель, которые из первого столбца содержат сомножителем a_{11} (остальные равны нулю). Вхождение сомножителя a_{11} занимает первую строку и первый столбец. Из второго столбца (при уже занятой первой строчке) остаётся включить в произведение a_{22} (остальные произведения равны нулю). Повторяя это рассуждение, приходим к произведению $a_{11}a_{22}\dots a_{nn}$ (остальные из $n!$ произведений все равны нулю).

Так как подстановка

$$\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$$

чётная, то это произведение входит со знаком $+$. □

Следствие 6.3.2.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix} \neq 0$$

тогда и только тогда, когда все $a_{ii} \neq 0$, $1 \leqslant i \leqslant n$.

Задача 6.3.3. Чему равен определитель

$$\begin{pmatrix} a_{11} & \dots & a_{1\,n-1} & a_{1n} \\ a_{21} & \dots & a_{2\,n-1} & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & 0 & 0 \end{pmatrix}$$

(т. е. чему равен определитель, в котором все элементы ниже побочной диагонали равны нулю)?

Свойство 2. *При перестановке двух строк* A_i *и* A_j, $i \neq j$, *матрицы* A *определитель меняет знак* ($|A'| = -|A|$, *где* A' *— матрица, полученная из матрицы* A *перестановкой двух строк*).

Доказательство. Произведение $\varepsilon(\alpha)a_{1\alpha(1)} \dots a_{n\alpha(n)}$ из $|A|$ входит также в $|A'|$, при этом новая подстановка индексов α' отличается от α одной транспозицией i и j в верхней строке (номера строк), таким образом, в α' имеем

$$\begin{pmatrix} i \\ \alpha(j) \end{pmatrix} \text{ вместо } \begin{pmatrix} j \\ \alpha(j) \end{pmatrix}, \quad \begin{pmatrix} j \\ \alpha(i) \end{pmatrix} \text{ вместо } \begin{pmatrix} i \\ \alpha(i) \end{pmatrix},$$

поэтому $\varepsilon(\alpha') = -\varepsilon(\alpha)$. □

Свойство 3. *Если* $A'_i = cA_i$ (*т. е. i-я строка матрицы A умножена на число c*), *то* $|A'| = c|A|$.

Доказательство. В каждое произведение, входящее в $|A'|$, из i-й строки входит только один сомножитель $ca_{i\alpha(i)}$, таким образом, $|A'| = c|A|$. □

Упражнение 6.3.4. Если $A \in \mathrm{M}_n(K)$, то $|cA| = c^n|A|$.

Свойство 4. *Если*

$$A_i = (a_{i1}, \ldots, a_{in}) = (b_1, \ldots, b_n) + (c_1, \ldots, c_n) = B + C$$

(т. е. i-я строка в матрице A представлена суммой двух строк), то $|A|$ равен сумме двух определителей $|A'|+|A''|$ матриц A' и A'', в которых вместо i-й строки A_i в A стоят соответственно строки B и C.

Доказательство. В каждое произведение

$$\varepsilon(\alpha)a_{1\alpha(1)} \ldots a_{n\alpha(n)}$$

входит из i-й строки A_i только один сомножитель

$$a_{i\alpha(i)} = b_{\alpha(i)} + c_{\alpha(i)}.$$

Таким образом,

$$\begin{aligned}
|A| &= \sum_{\alpha\in S_n} \varepsilon(\alpha)a_{1\alpha(1)} \ldots a_{i\alpha(i)} \ldots a_{n\alpha(n)} = \\
&= \sum_{\alpha\in S_n} \varepsilon(\alpha)a_{1\alpha(1)} \ldots (b_{\alpha(i)} + c_{\alpha(i)}) \ldots a_{n\alpha(n)} = \\
&= \sum_{\alpha\in S_n} \varepsilon(\alpha)a_{1\alpha(1)} \ldots b_{\alpha(i)} \ldots a_{n\alpha(n)} + \\
&+ \sum_{\alpha\in S_n} \varepsilon(\alpha)a_{1\alpha(1)} \ldots c_{\alpha(i)} \ldots a_{n\alpha(n)} = \\
&= |A'| + |A''|. \quad \square
\end{aligned}$$

6.4. Вывод следствий из свойств 1—4

Нам удобно следующие далее свойства выводить из «базовых» свойств 1—4.

Свойство 5. *Если $A_i = (0, \ldots, 0)$, то $|A| = 0$.*

Так как $A_i = 0 \cdot A_i$, то $|A| = 0 \cdot |A| = 0$.

Свойство 6. *Пусть $K = \mathbb{R}$ (или K — любое поле). Если $i \neq j$ и $A_i = A_j$, то $|A| = 0$.*

1) Сначала приведём доказательство для случая $K = \mathbb{R}$ (или для поля K, $\operatorname{char} K \neq 2$: из $2a = 0$ следует $a = 0$). Действительно, переставляя строки A_i и A_j, получаем $|A| = -|A|$, $2|A| = 0$, и поэтому $|A| = 0$.

2) Приведём общее доказательство в случае любого поля K при $n \geqslant 2$. Пусть $i < j$. Для каждой подстановки α, участвующей в выражении определителя

$$\sum_{\alpha \in \mathrm{S}_n} \varepsilon(\alpha) a_{1\alpha(1)} \dots a_{i\alpha(i)} \dots a_{j\alpha(j)} \dots a_{n\alpha(n)},$$

рассмотрим подстановку

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

полученную из α переменой местами чисел $\alpha(i)$ и $\alpha(j)$ в нижней строке канонической записи. Ясно, что $\varepsilon(\alpha') = -\varepsilon(\alpha)$. Так как $A_i = A_j$, то $a_{ik} = a_{jk}$ для $k = 1, \dots, n$, $a_{i\alpha(j)} = a_{j\alpha(j)}$, $a_{j\alpha(i)} = a_{i\alpha(i)}$. Поэтому

$$\varepsilon(\alpha) a_{1\alpha(1)} \dots a_{n\alpha(n)} = -\varepsilon(\alpha') a_{1\alpha'(1)} \dots a_{n\alpha'(n)}.$$

Если $\tau = (\alpha(i)\ \alpha(j)) \in \mathrm{S}_n$, то $\tau^2 = 1$ и отношение $\alpha \sim \beta$ для $\alpha, \beta \in \mathrm{S}_n$, где $\alpha \sim \beta$ означает, что $\alpha = \beta$ или $\alpha = \tau\beta$, является отношением эквивалентности. Действительно,

а) $\alpha \sim \alpha$;

б) $\alpha \sim \beta \implies \beta \sim \alpha$

(если $\alpha = \beta$, это очевидно; если $\alpha = \tau\beta$, то $\beta = \tau\alpha$, так как $\tau^2 = 1$);

в) $\alpha \sim \beta,\ \beta \sim \gamma \implies \alpha \sim \gamma$

(имеем четыре случая

1) $\alpha = \beta$, $\beta = \gamma$, поэтому $\alpha = \gamma$;

2) $\alpha = \tau\beta$, $\beta = \gamma$, поэтому $\alpha = \tau\gamma$;

3) $\alpha = \beta$, $\beta = \tau\gamma$, поэтому $\alpha = \tau\gamma$;

4) $\alpha = \tau\beta$, $\beta = \tau\gamma$, поэтому $\alpha = \tau^2\gamma = \gamma$;

и поэтому $\alpha \sim \gamma$).

Таким образом, разбиение на классы эквивалентных элементов приводит к разбиению на непересекающиеся классы $\{\alpha,\ \alpha' = \tau\alpha\}$. При $n \geqslant 2$ сумма $n!$ чётного числа слагаемых разбивается на суммы пар слагаемых по подстановке α и по подстановке α', равные нулю, поскольку эти два слагаемые отличаются знаком. □

Свойство 7. *Если от квадратной матрицы A переходим к матрице A' с помощью элементарного преобразования 1-го типа $A'_i = A_i + cA_j$, $i \neq j$, $c \in K$, то $|A'| = |A|$.*

Действительно, разлагая определитель $|A'|$ в сумму двух определителей (по i-й строке), мы получаем $|A|$ и нулевой определитель, в котором после вынесения из i-й строки числа c имеем две одинаковые строки (A_j на месте i-й строки и A_j на своём j-м месте).

6.5. Линейная комбинация строк в линейном пространстве строк K^n

Если

$$a_1 = (a_{11}, \ldots, a_{1n}), \ldots, a_r = (a_{r1}, \ldots, a_{rn}) \in K^n$$

и

$$k_1, \ldots, k_r \in K,$$

то можно образовать *линейную комбинацию строк*

$$\sum_{i=1}^{r} k_i a_i = k_1 a_1 + \ldots + k_r a_r = \left(\sum_{i=1}^{r} k_i a_{i1}, \ldots, \sum_{i=1}^{r} k_i a_{in} \right) \in K^n,$$

здесь на j-м месте стоит элемент

$$\sum_{i=1}^{r} k_i a_{ij}.$$

Свойство 8. *Если найдётся строка A_i, являющаяся линейной комбинацией остальных строк квадратной матрицы A, то $|A| = 0$.*

Действительно, если

$$A_i = \sum_{\substack{l=1 \\ l \neq i}}^{n} k_l A_l,$$

то, разлагая определитель $|A|$ в сумму $(n-1)$ определителей и вынося в каждом из слагаемых-определителей из l-й строки число k_l, получаем определители с двумя одинаковыми строчками (на месте i-й строки стоит строка A_j, на месте j-й строки стоит строка A_j). □

Определение 6.5.1. Если $A = (a_{ij})$ — квадратная $(n \times n)$-матрица, то $(n \times n)$-матрица $A^* = (b_{ij})$, $b_{ij} = a_{ji}$, называется матрицей, полученной *транспонированием* из матрицы A (т. е. симметрией относительно диагонали).

Теорема 6.5.2. $|A^*| = |A|$ (*определитель квадратной матрицы не меняется при транспонировании*).

Доказательство. Каждый член $a_{1\alpha(1)} \dots a_{n\alpha(n)}$ определителя

$$|A| = \sum_{\alpha \in S_n} \varepsilon(\alpha) a_{1\alpha(1)} \dots a_{n\alpha(n)}$$

входит в определитель $|A^*|$ транспонированной матрицы A^*, при этом со знаком, определяемым подстановкой

$$\alpha' = \begin{pmatrix} \alpha(1) & \dots & \alpha(n) \\ 1 & \dots & n \end{pmatrix}.$$

Так как $\varepsilon(\alpha) = \varepsilon(\alpha')$, то в итоге мы имеем $|A^*| = |A|$. □

Следствие 6.5.3. *Свойства* 1—8 *выполняются и для столбцов определителя* $|A|$ *квадратной* $(n \times n)$*-матрицы* A.

Действительно, при переходе от матрицы A к транспонированной матрице A^* строки превращаются в столбцы, а столбцы — в строки. Преобразования строк транспонированной матрицы A^* соответствуют преобразованиям столбцов матрицы A.

6.6. Вычисление определителей

Определение определителя

$$|A| = \sum_{\alpha \in S_n} \varepsilon(\alpha) a_{1\alpha(n)} \dots a_{n\alpha(n)}$$

как суммы $n!$ слагаемых-произведений плохо пригодно для реальных вычислений при больших n. В теоретическом плане важно отметить, что определитель $|A|$ является многочленом от n^2 переменных a_{ij}, в котором мономы входят с коэффициентами ± 1. Отметим лишь одно из следствий этого факта: если $a_{ij} = a_{ij}(x)$ являются дифференцируемыми функциями от переменной x, то определитель $|A|$ также является дифференцируемой функцией от x, поскольку суммы и произведения дифференцируемых функций являются дифференцируемыми функциями.

Теорема 6.6.1. *Пусть от квадратной $(n \times n)$-матрицы $A = (a_{ij})$ элементарными преобразованиями 1-го и 2-го типа (t преобразований 2-го типа) мы пришли к треугольной матрице*

$$\bar{A} = \begin{pmatrix} \bar{a}_{11} & & & \\ 0 & \bar{a}_{22} & & * \\ \vdots & & \ddots & \\ 0 & \dots\dots & & \bar{a}_{nn} \end{pmatrix}$$

(все элементы ниже диагонали равны нулю; любая ступенчатая матрица, очевидно, является треугольной). Тогда

$$|A| = (-1)^t \bar{a}_{11} \dots \bar{a}_{nn}.$$

Доказательство. Так как

$$|\bar{A}| = (-1)^t |A|,$$

то

$$|A| = (-1)^t |\bar{A}| = (-1)^t \bar{a}_{11} \dots \bar{a}_{nn}. \quad \square$$

6.7. Характеризация функции определителя матрицы базовыми свойствами

Теорема 6.7.1 (о единственности функции с базовыми свойствами 1—4 определителя). *Пусть функция F, сопоставляющая каждой квадратной $(n \times n)$-матрице $A \in \mathrm{M}_n(\mathbb{R})$ число $F(A) \in \mathbb{R}$ удовлетворяет базовым свойствам* 1—4 *функций определителя. Тогда $F(A) = |A|$, т. е. функция определителя $|A|$ однозначно определяется свойствами* 1—4.

Доказательство. Приведём $(n \times n)$-матрицу A к треугольному виду

$$\bar{A} = \begin{pmatrix} \bar{a}_{11} & & & * \\ 0 & \bar{a}_{22} & & \\ \vdots & & \ddots & \\ 0 & \ldots\ldots & & \bar{a}_{nn} \end{pmatrix}$$

элементарными преобразованиями строк 1-го и 2-го типа (t преобразований 2-го типа). Тогда

$$F(\bar{A}) = (-1)^t F(A),$$

следовательно,

$$F(A) = (-1)^t F(\bar{A}).$$

Далее, вынося элемент $\bar{a}_{nn}$ из n-й строки и создавая 0 над ним, получаем

$$F(\bar{A}) = \bar{a}_{nn} F \begin{pmatrix} \bar{a}_{11} & & & * \\ \vdots & \ddots & & \\ 0 & \ldots & \bar{a}_{n-1\,n-1} & \\ 0 & \ldots & 0 & 1 \end{pmatrix} = \bar{a}_{nn} F \begin{pmatrix} \bar{a}_{11} & & * & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \ldots & \bar{a}_{n-1\,n-1} & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}.$$

Продолжая это рассуждение, получаем

$$F(\bar{A}) = \bar{a}_{11} \ldots \bar{a}_{nn} F \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \bar{a}_{11} \ldots \bar{a}_{nn}.$$

Итак,

$$F(A) = (-1)^t F(\bar{A}) = (-1)^t \bar{a}_{11} \ldots \bar{a}_{nn} = |A|. \quad \square$$

6.8. Сведение вычисления определителя к определителям меньшего порядка

Определение 6.8.1 (дополняющие миноры и алгебраические дополнения). Зафиксируем элемент a_{ij} квадратной $(n \times n)$-матрицы $A = (a_{ij})$. Вычёркивая в определителе $|A|$ i-ю строку и j-й столбец (проходящие через a_{ij}), получаем определитель M_{ij} матрицы порядка $(n-1) \times (n-1)$, называемый (*дополняющим*) *минором* элемента a_{ij}. *Алгебраическим дополнением* элемента a_{ij} называется число $A_{ij} = (-1)^{i+j} M_{ij}$.

Замечание 6.8.2. Имеем n^2 (дополняющих) миноров M_{ij}.

Лемма 6.8.3.

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} M_{11} = a_{11} A_{11}.$$

Доказательство. Каждый член определителя вида

$$a_{11} a_{2\alpha(2)} \dots a_{n\alpha(n)}$$

(все остальные заведомо равны нулю) входит в правую часть доказываемого равенства, при этом с тем же знаком:

$$\varepsilon \begin{pmatrix} 1 & 2 & \dots & n \\ 1 = \alpha(1) & \alpha(2) & \dots & \alpha(n) \end{pmatrix} = \\ = \varepsilon \begin{pmatrix} 2 & \dots & n \\ \alpha(2) & \dots & \alpha(n) \end{pmatrix} = \varepsilon \begin{pmatrix} 1 & \dots & n-1 \\ \alpha(2)-1 & \dots & \alpha(n)-1 \end{pmatrix}. \quad \square$$

Следствие 6.8.4.

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} A_{11}.$$

Лемма 6.8.5.

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

Доказательство. Переставляя последовательно i-ю строку $(i-1)$ раз с $(i-1)$ строками, стоящими над ней, а затем переставляя последовательно k-й столбец $(k-1)$ раз с $(k-1)$ столбцами, стоящими левее его, получаем

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

$$\overset{\text{лемма 6.8.3}}{=} (-1)^{i+k} a_{ik} M_{ik} = a_{ik}A_{ik}. \quad \square$$

Теорема 6.8.6 (разложение определителя по i-й строке и по j-му столбцу, $1 \leqslant i, j \leqslant n$).

$$1) \quad |A| = a_{i1}A_{i1} + \dots + a_{in}A_{in} \left(= \sum_{j=1}^{n} a_{ij}A_{ij} \right);$$

$$2) \quad |A| = a_{1j}A_{1j} + \dots + a_{nj}A_{nj} \left(= \sum_{i=1}^{n} a_{ij}A_{ij} \right).$$

Доказательство.

1) Поскольку

$$(a_{i1}, \dots, a_{in}) = (a_{i1}, 0, \dots, 0) + \dots + (0, \dots, 0, a_{in}),$$

то, применяя лемму 6.8.5, получаем

$$|A| = \sum_{j=1}^{n} \begin{vmatrix} & & * & & \\ 0 & \dots & a_{ij} & \dots & 0 \\ & & * & & \end{vmatrix} = \sum_{j=1}^{n} a_{ij}A_{ij}.$$

2) Так как $|A| = |A^*|$, то разложение по j-й строке для $|A^*|$ является разложением по j-му столбцу для $|A|$. $\square$

Теорема 6.8.7 (о фальшивом разложении по i-й строке и по j-му столбцу).

1) *При* $i \neq k$

$$\sum_{j=1}^{n} a_{ij}A_{kj} = a_{i1}A_{k1} + \ldots + a_{in}A_{kn} = 0$$

(*сумма произведений элементов* a_{ij} i-*й строки на алгебраические дополнения* A_{kj} *элементов «чужой»* k-*й строки при* $i \neq k$ *равна нулю*);

2) *при* $j \neq k$

$$\sum_{i=1}^{n} a_{ij}A_{ik} = a_{1j}A_{1k} + \ldots + a_{nj}A_{nk} = 0$$

(*сумма произведений элементов* a_{ij} j-*го столбца на алгебраические дополнения* A_{ik} *элементов «чужого»* k-*го столбца при* $j \neq k$ *равна нулю*).

Доказательство.
1)

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{vmatrix} & * & \\ a_{i1} & \ldots & a_{in} \\ & * & \\ a_{i1} & \ldots & a_{in} \\ & * & \end{vmatrix} \begin{matrix} \\ i \\ \\ k \\ \\ \end{matrix} = 0$$

(разложение по k-й строке определителя, полученного из исходного заменой k-й строки на i-ю и равного 0, поскольку в нём имеется две одинаковые строки, i-я и k-я).

2) Применяя 1) к фальшивому разложению по строке для $|A^*|$, $|A^*| = |A|$, получаем фальшивое разложение по столбцу для $|A|$. □

Пример 6.8.8. Найти определитель

$$\Delta = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix}.$$

а) По определению,

$$\Delta = 1\cdot 3\cdot 2 + 2\cdot 3\cdot 1 + 2\cdot 1\cdot 3 - 3\cdot 3\cdot 3 - 1\cdot 1\cdot 1 - 2\cdot 2\cdot 2 = -18.$$

б) Разлагая по первой строке, получаем

$$\Delta = 1\cdot\begin{vmatrix}3 & 1\\ 1 & 2\end{vmatrix} + 2\cdot(-1)\cdot\begin{vmatrix}2 & 1\\ 3 & 2\end{vmatrix} + 3\cdot\begin{vmatrix}2 & 3\\ 3 & 1\end{vmatrix} = -18.$$

в) Используя элементарные преобразования строк, имеем

$$\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\\ 3 & 1 & 2\end{pmatrix} \to \begin{pmatrix}1 & 2 & 3\\ 0 & -1 & -5\\ 3 & 1 & 2\end{pmatrix} \to \begin{pmatrix}1 & 2 & 3\\ 0 & -1 & -5\\ 0 & -5 & -7\end{pmatrix} \to \begin{pmatrix}1 & 2 & 3\\ 0 & -1 & -5\\ 0 & 0 & 18\end{pmatrix},$$

и мы пришли к треугольному виду. При этом мы применяли только преобразования 1-го типа, не меняющие определитель. Следовательно, $\Delta = -18$.

Пример 6.8.9. Найти определитель

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

Используем элементарные преобразования строк, оставляя неизменной третью строку:

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

Мы применяли только преобразования 1-го типа, не меняющие определитель. Применяя разложение последнего определителя по первому столбцу, имеем

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

Пример 6.8.10 (вычисление определителя n-го порядка с помощью рекуррентного соотношения). Найти определитель

$$\Delta_n = \begin{vmatrix} 9 & 5 & 0 & 0 & \dots & 0 & 0 \\ 4 & 9 & 5 & 0 & \dots & 0 & 0 \\ 0 & 4 & 9 & 5 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 4 & 9 \end{vmatrix}.$$

Разложим определитель по первой строке:

$$\Delta_n = 9 \cdot \Delta_{n-1} + (-1) \cdot 5 \cdot \Delta = 9 \cdot \Delta_{n-1} + (-5) \cdot 4 \cdot \Delta_{n-2}$$

(в соответствующем миноре Δ мы применим разложение по первому столбцу). Если учесть, что $\Delta_1 = 9$ и $\Delta_2 = 61$, полученная рекуррентная формула позволяет вычислить Δ_n для любого n. Нетрудно убедиться, что $\Delta_n = 5^{n+1} - 4^{n+1}$ (это можно доказать, например, индукцией по n).

Задача 6.8.11. Вычислить определители порядка n:

а)

$$\begin{vmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \iddots & & \vdots \\ 1 & 0 & \dots & 0 \end{vmatrix}$$

(все элементы вне побочной диагонали равны 0, а на побочной диагонали стоят 1).

б)

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

Упражнение 6.8.12 (игра в определитель). Играют два участника, расставляя по очереди числа 1, 2, 3, 4, 5, 6, 7, 8, 9 без повторений в качестве элементов матрицы 3×3. Один из участников (I)

стремится в итоге получить положительный определитель, а другой (II) — отрицательный. Чтобы уравнять шансы, играется две партии: в первой партии первый ход делает участник I, а во второй — участник II. После этих двух партий значения полученных определителей складываются. Если получилось положительное число, то выиграл участник I, если отрицательное число, то выиграл участник II, если нуль, то ничья. Покажите, что сумма всех 9! определителей, возможных в этой игре, равна нулю.

Теорема 6.8.13 (об определителе с углом нулей).

$$|C| = \begin{vmatrix} A & U \\ 0 & B \end{vmatrix} = |A|\,|B|,$$

где $A \in \mathrm{M}_n(K)$, $U \in \mathrm{M}_{n,m}(K)$, $0 \in \mathrm{M}_{m,n}(K)$ *— нулевая* $(m \times n)$*-матрица,* $B \in \mathrm{M}_m(K)$.

Доказательство. Проведём индукцию по n. Начало индукции $n = 1$ рассмотрено в лемме 6.8.3. Пусть $n \geqslant 2$ и утверждение верно для всех $n' < n$. Разложим наш определитель $|C|$ по первому столбцу:

$$|C| = a_{11}C_{11} + \ldots + a_{n1}C_{n1}.$$

Так как по индуктивному предположению для M_{i1}

$$C_{i1} = (-1)^{i+1}M_{i1} = (-1)^{i+1}M'_{i1} \cdot |B|,$$

где

$$M_{i1} = \begin{vmatrix} M'_{i1} & U' \\ 0 & B \end{vmatrix}, \quad 1 \leqslant i \leqslant n,$$

M'_{i1} — дополняющий минор элемента a_{i1} в матрице A, то

$$|C| = \sum_{i=1}^{n} a_{i1}C_{i1} = \sum_{i=1}^{n} a_{i1}(-1)^{i+1}M'_{i1}|B| = \\ = \left(\sum_{i=1}^{n} a_{i1}A_{i1}\right)|B| = |A|\,|B|. \quad \square$$

Следствие 6.8.14. *Пусть A_i, $1 \leqslant i \leqslant r$, — квадратные матрицы. Тогда*

$$\begin{vmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_r \end{vmatrix} = |A_1|\,|A_2| \ldots |A_r|.$$

Упражнение 6.8.15 (теорема Лапласа). Если M — минор (т. е. определитель матрицы), проходящий через k строк с номерами $i_1, \ldots, i_k$ и k столбцов с номерами $j_1, \ldots, j_k$, $k \geqslant 1$, то дополнительный минор $\bar{M}$ определяется как определитель, получаемый вычёркиванием строк $i_1, \ldots, i_k$ и столбцов $j_1, \ldots, j_k$. Алгебраическое дополнение минора M определяется следующим образом:

$$A(M) = (-1)^{(i_1+\ldots+i_k)+(j_1+\ldots+j_k)}\bar{M}.$$

Если $A = (a_{ij}) \in \mathrm{M}_n(K)$, $1 \leqslant k \in \mathbb{N}$, $i_1, \ldots, i_k$ — зафиксированные номера k строк, то определитель $|A|$ равен сумме всех произведений $MA(M)$, где M пробегает все C_n^k миноров, проходящих через строки с номерами $i_1, \ldots, i_k$.

Частными случаями теоремы Лапласа являются теорема о разложении по строке ($k = 1$) и теорема об определителе с углом нулей.

Теорема 6.8.16 (правило Крамера). *Для квадратной системы линейных уравнений $(a_{ij} \mid b_i)$ с $(n \times n)$-матрицей $A = (a_{ij})$ имеем:*

1) *система является определённой тогда и только тогда, когда $|A| \neq 0$;*

2) *в этом случае (т. е. если $|A| \neq 0$) это единственное решение $(k_1, \ldots, k_n)$ имеет следующий вид для $j = 1, \ldots, n$:*

$$k_j = \frac{D_j}{D},$$

где

$$D = |A|, \quad D_j = \begin{vmatrix} a_{11} & \ldots & \boxed{b_1} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{vmatrix} \text{ —}$$

определитель, полученный из определителя $|A|$ путём замены j-го столбца на столбец

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

свободных членов системы.

Доказательство.

1) Приведя элементарными преобразованиями систему к ступенчатому виду $(\bar{a}_{ij}, \bar{b}_i)$ со ступенчатой матрицей $\bar{A} = (\bar{a}_{ij})$, из критерия определённости квадратной системы имеем: система (a_{ij}, b_i) является определённой тогда и только тогда, когда ступенчатая матрица

$$\bar{A} = \begin{pmatrix} \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1n} \\ 0 & \bar{a}_{22} & \dots & \bar{a}_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \bar{a}_{nn} \end{pmatrix}$$

треугольная с ненулевыми элементами по диагонали, $\bar{a}_{11} \neq 0$, $\bar{a}_{22} \neq 0, \dots, \bar{a}_{nn} \neq 0$, т. е.

$$|A| = (-1)^t |\bar{A}| = (-1)^t \bar{a}_{11} \dots \bar{a}_{nn} \neq 0.$$

2) Если $(k_1, \dots, k_n)$ — единственное решение нашей системы,

$$\begin{cases} a_{11}k_1 + \dots + a_{1j}k_j + \dots + a_{1n}k_n = b_1, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}k_1 + \dots + a_{nj}k_j + \dots + a_{nn}k_n = b_n, \end{cases}$$

то, умножая 1-е уравнение на A_{1j}, i-е — на A_{ij}, n-е — на A_{nj} и складывая, получаем

$$0 \cdot k_1 + \dots + Dk_j + \dots + 0 \cdot k_n = b_1 A_{1j} + \dots + b_n A_{nj} = D_j.$$

Итак, $Dk_j = D_j$, $D \neq 0$, поэтому $k_j = \dfrac{D_j}{D}$.

Второе доказательство утверждения 2). Покажем, что $(k_1, \dots, k_n)$, где $k_j = \dfrac{D_j}{D}$, является решением.

Действительно, подставим строчку $(k_1, \ldots, k_n)$ в i-е уравнение $\sum\limits_{j=1}^{n} a_{ij}x_j = b_i$:

$$\sum_{j=1}^{n} a_{ij}k_j = \sum_{j=1}^{n} \frac{a_{ij}D_j}{D} =$$

$$= \frac{\sum\limits_{j=1}^{n} a_{ij} \sum\limits_{k=1}^{n} b_k A_{kj}}{D} = \frac{\sum\limits_{k=1}^{n} b_k \Big(\sum\limits_{j=1}^{n} a_{ij}A_{kj} \Big)}{D} = \frac{b_i D}{D} = b_i.$$

Мы использовали разложение определителя D_j по j-му столбцу $D_j = \sum\limits_{k=1}^{n} b_k A_{kj}$, а также при $k = i$ разложение $\sum\limits_{j=1}^{n} a_{ij}A_{ij} = D$ и при $k \neq i$ фальшивое разложение $\sum\limits_{j=1}^{n} a_{ij}A_{kj} = 0$. □

Из теоремы Крамера можно вывести полезные следствия.

Следствие 6.8.17. *Если квадратная система линейных уравнений (n уравнений с n неизвестными) не имеет решения, то определитель матрицы её коэффициентов равен нулю.*

Доказательство. Если $|A| = |(a_{ij})| \neq 0$, то по правилу Крамера система имеет решение. □

Следствие 6.8.18. *Если квадратная система линейных уравнений (n уравнений с n неизвестными) имеет более чем одно решение, то определитель матрицы её коэффициентов равен нулю.*

Доказательство. Если $|A| = |(a_{ij})| \neq 0$, то по правилу Крамера система имеет единственное решение. □

Следствие 6.8.19. *Однородная квадратная система линейных уравнений (n уравнений с n неизвестными) имеет ненулевое решение тогда и только тогда, когда определитель матрицы её коэффициентов равен нулю.*

Следствие 6.8.20. *Если коэффициенты квадратной системы $a_{ij}(t)$ и свободные члены $b_i(t)$ являются непрерывными функциями от t, то в силу правила Крамера компоненты k_j решения $(k_1, \ldots, k_n)$*

являются рациональными дробями от переменных $\{a_{ij}, b_i\}$ *с целыми коэффициентами и поэтому являются непрерывными функциями от* t *в некоторой окрестности точки* $t_0 \in \mathbb{R}$, *где* $|a_{ij}(t_0)| \neq 0$.

Задача 6.8.21. Пусть $A, B, C, D \in \mathrm{M}_n(K)$. Тогда

$$1) \quad \begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A|\,|D| - |B|\,|C|;$$

$$2) \quad \begin{vmatrix} A & B \\ B & A \end{vmatrix} = |A+B| \cdot |A-B|.$$

Задача 6.8.22. Показать (разлагая по последнему столбцу), что

$$\begin{vmatrix} x & 0 & 0 & \ldots & 0 & a_0 \\ -1 & x & 0 & \ldots & 0 & a_1 \\ 0 & -1 & x & \ldots & 0 & a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & x & a_{n-1} \\ 0 & 0 & 0 & \ldots & -1 & a_n \end{vmatrix} = a_0 + a_1 x + \ldots + a_n x^n.$$

Задача 6.8.23. Пусть $f(x) = (c_1 - x)(c_2 - x)\ldots(c_n - x)$, $a \neq b$. Тогда

$$\begin{vmatrix} c_1 & a & a & \ldots & a \\ b & c_2 & a & \ldots & a \\ b & b & c_3 & \ldots & a \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ b & b & b & \ldots & c_n \end{vmatrix} = \frac{af(b) - bf(a)}{a-b}.$$

Задача 6.8.24. Вычислить определитель порядка n

$$\begin{vmatrix} n & 1 & \ldots & 1 \\ 1 & n & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & \ldots & n \end{vmatrix}$$

(элементы на главной диагонали равны n, все остальные элементы равны 1).

Ответ. $(2n-1)(n-1)^{n-1}$.

Задача 6.8.25. Доказать (разлагая по строке и получая рекуррентное соотношение), что

$$\begin{vmatrix} a & 0 & \dots & 0 & b \\ 0 & a & \dots & b & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & b & \dots & a & 0 \\ b & 0 & \dots & 0 & a \end{vmatrix} = (a^2 - b^2)^k,$$

где $n = 2k$ — размер матрицы.

6.9. Определитель Вандермонда

Теорема 6.9.1.

$$V(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & \dots & a_1^{n-1} \\ 1 & a_2 & \dots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \dots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j), \quad a_1, \dots, a_n \in K.$$

Доказательство. Проведём индукцию по n (начало индукции $n = 2$). Пусть утверждение верно для $n' < n$. Тогда, применяя элементарные преобразования столбцов $\hat{A}_n - a_1 \hat{A}_{n-1}$, $\hat{A}_{n-1} - a_1 \hat{A}_{n-2}, \dots$, $\hat{A}_2 - a_1 \hat{A}_1$ и предположение индукции, получаем

$$V(a_1, \dots, a_n) = \begin{vmatrix} 1 & a_1 & \dots & a_1^{n-1} \\ 1 & a_2 & \dots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \dots & a_n^{n-1} \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 0 & \dots & 0 \\ 1 & (a_2 - a_1) & \dots & a_2^{n-2}(a_2 - a_1) \\ \vdots & \vdots & & \vdots \\ 1 & (a_n - a_1) & \dots & a_n^{n-2}(a_n - a_1) \end{vmatrix} =$$

$$= \begin{vmatrix} (a_2 - a_1) & \dots & a_2^{n-2}(a_2 - a_1) \\ \vdots & & \vdots \\ (a_n - a_1) & \dots & a_n^{n-2}(a_n - a_1) \end{vmatrix} =$$

$$= (a_2 - a_1) \dots (a_n - a_1) \begin{vmatrix} 1 & a_2 & \dots & a_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \dots & a_n^{n-2} \end{vmatrix} =$$

$$= \prod_{k=2}^{n} (a_k - a_1) V(a_2, \dots, a_n) =$$

$$= \prod_{k=2}^{n} (a_k - a_1) \prod_{2 \leqslant j < i \leqslant n} (a_i - a_j) = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j). \quad \square$$

Следствие 6.9.2. *$V(a_1, \dots, a_n) \neq 0$ тогда и только тогда, когда $a_i \neq a_j$ при $i \neq j$ (т. е. когда все элементы $a_1, a_2, \dots, a_n$ различны).*

Теорема 6.9.3 (интерполяционная формула Лагранжа).

1) *Если $a_1, \dots, a_n$ — различные элементы поля K, $b_1, \dots, b_n$ — любые элементы поля K, то существует и единственный многочлен $f(x) \in K[x]$ такой, что $\deg f(x) \leqslant n - 1$ и $f(a_i) = b_i$ для всех $1 \leqslant i \leqslant n$ (здесь $\deg f(x)$ — степень многочлена $f(x)$).*

2) *Этот многочлен имеет вид*

$$f(x) = \sum_{i=1}^{n} b_i \frac{(x - a_1) \dots \widehat{(x - a_i)} \dots (x - a_n)}{(a_i - a_1) \dots \widehat{(a_i - a_i)} \dots (a_i - a_n)}$$

(здесь $\widehat{(x - a_i)}$, $\widehat{(a_i - a_i)}$ означает, что эти множители не входят в произведения).

3) *Интерполяционный многочлен $f(x) \in K[x]$, $\deg f(x) \leqslant n - 1$, для которого $f(a_i) = b_i$, $i = 1, \dots, n$, можно находить методом Ньютона в виде*

$$f(x) = \\ = \lambda_0 + \lambda_1(x - a_1) + \lambda_2(x - a_1)(x - a_2) + \dots + \lambda_{n-1} \prod_{i=1}^{n-1} (x - a_i),$$

при этом коэффициенты определяются последовательно: при $x = a_1$ *имеем* $b_1 = f(a_1) = \lambda_0$, *т. е.* $\lambda_0 = b_1$; *при* $x = a_2$ *имеем* $b_2 = f(a_2) = b_1 + \lambda_1(a_2 - a_1)$, *т. е.* $\lambda_1 = (b_2 - b_1)/(a_2 - a_1)$; ...; *при* $x = a_{n-1}$ *получаем*

$$b_{n-1} = \lambda_0 + \lambda_1(a_{n-1} - a_1) + \ldots + \lambda_{n-2} \prod_{i=1}^{n-2} (a_{n-1} - a_i)$$

и находим λ_{n-2} (*коэффициент при* λ_{n-2} *отличен от нуля*); *полагая* $x = a_n$, *имеем коэффициент* $\prod_{i=1}^{n-1} (a_n - a_i) \neq 0$ *при* λ_{n-1} *в равенстве*

$$b_n = \lambda_0 + \lambda_1(a_n - a_1) + \ldots + \lambda_{n-1} \prod_{i=1}^{n-1} (a_n - a_i)$$

и находим λ_{n-1}.

Доказательство.

1) Будем искать многочлен

$$f(x) = f_0 + f_1 x + \ldots + f_{n-1} x^{n-1},$$

где $f_0, f_1, \ldots, f_{n-1}$ — неизвестные коэффициенты (элементы поля K), такой, что

$$\begin{aligned} f(a_1) &= f_0 + f_1 a_1 + \ldots + f_{n-1} a_1^{n-1} = b_1, \\ &\vdots \\ f(a_n) &= f_0 + f_1 a_n + \ldots + f_{n-1} a_n^{n-1} = b_n. \end{aligned}$$

Определитель этой системы

$$V(a_1, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & \ldots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_n & \ldots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j) \neq 0,$$

поскольку все элементы $a_1, \ldots, a_n$ различны. Поэтому такой многочлен $f(x)$ существует (и единственный).

2) Очевидно, что приведённый многочлен в форме Лагранжа

$$f(x) = \sum_{i=1}^{n} b_i \frac{(x-a_1)\ldots\widehat{(x-a_i)}\ldots(x-a_n)}{(a_i-a_1)\ldots\widehat{(a_i-a_i)}\ldots(a_i-a_n)}$$

удовлетворяет двум условиям:

$$\deg f(x) \leqslant n-1;$$
$$f(a_i) = b_i, \quad i = 1, 2, \ldots, n.$$

3) Многочлен $f(x)$ в форме Ньютона удовлетворяет двум условиям:

$$\deg f(x) \leqslant n-1;$$
$$f(a_i) = b_i, \quad i = 1, 2, \ldots, n. \quad \square$$

Упражнение 6.9.4. Пусть $0 \leqslant k_1 < k_2 < \ldots < k_n \in \mathbb{Z}$, $0 < a_1 < a_2 < \ldots < a_n \in \mathbb{R}$, $A = (a_{ij})$, где $a_{ij} = a_i^{k_j}$. Тогда $|A| > 0$.

Упражнение 6.9.5. Пусть $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$, где $a_{ij} = \dfrac{1}{a_i + b_j}$, $a_i, b_j \in \mathbb{R}$. Тогда

$$|A| = \frac{\prod\limits_{1 \leqslant i < j \leqslant n} (a_j - a_i)(b_j - b_i)}{\prod\limits_{i,j=1}^{n} (a_i + b_j)}.$$

Глава 7

Линейные преобразования линейных пространств столбцов, задаваемые (прямоугольной) матрицей

Рассмотрим линейные пространства столбцов над полем K (например, над полем $\mathbb{R}$ действительных чисел)

$$U = \hat{K}^n = \left\{ X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \, x_i \in K \right\},$$

$$V = \hat{K}^m = \left\{ Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \middle| \, y_i \in K \right\}.$$

Каждая $(m \times n)$-матрица $F = (f_{ij})$, $f_{ij} \in K$, *задаёт отображение* $f : U \to V$,

$$f(X) = Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

для всех

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = X \in U = \hat{K}^n,$$

где

$$\begin{aligned} y_1 &= f_{11}x_1 + \ldots + f_{1n}x_n, \\ &\vdots \qquad\quad \vdots \qquad\qquad \vdots \\ y_m &= f_{m1}x_1 + \ldots + f_{mn}x_n. \end{aligned}$$

Теорема 7.0.6. *Отображение*

$$f\colon U = \hat{K}^n \to V = \hat{K}^m,$$

задаваемое прямоугольной $(m \times n)$-матрицей $F = (f_{ij})$, обладает следующими свойствами:

1) $f(X + X') = f(X) + f(X')$ *для всех* $X, X' \in U$;

2) $f(cX) = cf(X)$ *для всех* $c \in K$, $X \in U$.

Доказательство. Для

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \quad c \in K$$

имеем

$$X + X' = \begin{pmatrix} x_1 + x'_1 \\ \vdots \\ x_n + x'_n \end{pmatrix}, \quad cX = \begin{pmatrix} cx_1 \\ \vdots \\ cx_n \end{pmatrix}.$$

Применяя отображение f, определяемое прямоугольной матрицей $F = (f_{ij})$, к $X + X'$ и cX, соответственно получаем

$$f(X + X') = f(X) + f(X'), \quad f(cX) = cf(X). \quad \square$$

Замечание 7.0.7. Отображение $f\colon U \to V$ из одного линейного пространства U в другое линейное пространство V, удовлетворяющее свойствам

1) $f(X + X') = f(X) + f(X')$ для всех $X, X' \in U$,

2) $f(cX) = cf(X)$ для всех $c \in K$, $X \in U$,

называется *линейным отображением* (преобразованием). Тем самым мы показали, что отображение, задаваемое прямоугольной $(m \times n)$-матрицей $F = (f_{ij})$, определяет линейное преобразование соответствующих линейных пространств столбцов:

$$f\colon U = \hat{K}^n \to V = \hat{K}^m.$$

Пример 7.0.8. Если $m = 1$, то имеем линейную функцию

$$y = f_1 x_1 + \ldots + f_m x_n$$

из $U = \hat{K}^n$ в $\hat{K}^1 = K$.

Пример 7.0.9. Поворот плоскости вокруг точки $(0, 0)$ на угол α является линейным отображением $f\colon \hat{\mathbb{R}}^2 \to \hat{\mathbb{R}}^2$, задаваемым матрицей поворота

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Теорема 7.0.10 (об однозначной определяемости матрицы, задающей линейное отображение столбцов). *Пусть*

$$f\colon U = \hat{K}^n \to V = \hat{K}^m, \quad g\colon U = \hat{K}^n \to V = \hat{K}^m -$$

два линейных отображения, задаваемых $(m \times n)$-матрицами $F = (f_{ij})$ и $G = (g_{ij})$ соответственно. Тогда $f = g$ в том и только в том случае, когда $F = G$ (т. е. $f_{ij} = g_{ij}$ для всех i, j).

Доказательство.

1) Если $F = G$, то ясно, что $f = g$.

2) Пусть $f = g$. Рассмотрим

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

где 1 стоит в j-й строке, а остальные элементы равны нулю. Тогда

$$\begin{pmatrix} f_{1j} \\ \vdots \\ f_{ij} \\ \vdots \\ f_{mj} \end{pmatrix} = f(e_j) = g(e_j) = \begin{pmatrix} g_{1j} \\ \vdots \\ g_{ij} \\ \vdots \\ g_{mj} \end{pmatrix},$$

поэтому для любого i имеем $f_{ij} = g_{ij}$, т. е. $F = (f_{ij}) = (g_{ij}) = G$. □

Теорема 7.0.11 (о задании любого линейного отображения линейных пространств столбцов матрицей). *Пусть*

$$f : U = \hat{K}^n \to V = \hat{K}^m -$$

линейное отображение линейных пространств столбцов, т. е.

1) $f(X + X') = f(X) + f(X')$ *для всех* $X, X' \in U$,

2) $f(cX) = cf(X)$ *для всех* $c \in K$, $X \in U$.

Тогда найдётся (и единственная) $(m \times n)$*-матрица* $F = (f_{ij})$ *такая, что определяемое с её помощью линейное отображение совпадает с линейным отображением* f.

Доказательство. Пусть

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j, \quad f(e_j) = \begin{pmatrix} f_{1j} \\ \vdots \\ f_{ij} \\ \vdots \\ f_{mj} \end{pmatrix} \in V = \hat{K}^m, \quad f_{ij} \in K.$$

Получили $(m \times n)$-матрицу $F = (f_{ij})$.

Для любого

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U = \hat{K}^n$$

имеем

$$X = x_1 e_1 + \ldots + x_n e_n.$$

Тогда

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = f(X) = x_1 f(e_1) + \ldots + x_n f(e_n) = \\ = x_1 \begin{pmatrix} f_{11} \\ \vdots \\ f_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} f_{1n} \\ \vdots \\ f_{mn} \end{pmatrix},$$

т. е.

$$\begin{aligned} y_1 &= f_{11}x_1 + \ldots + f_{1n}x_n, \\ &\vdots \qquad\quad \vdots \qquad\qquad \vdots \\ y_m &= f_{m1}x_1 + \ldots + f_{mn}x_n. \end{aligned}$$

Итак, линейное отображение f задаётся $(m \times n)$-матрицей $F = (f_{ij})$.

Как мы показали, матрица $F = (f_{ij})$ определена однозначно. □

7.1. Произведение линейных отображений

Теорема 7.1.1. *Если ${}_KU$, ${}_KV$, ${}_KW$ — линейные пространства над полем K,*

$$U \xrightarrow{f} V \xrightarrow{g} W,$$

f и g — линейные отображения линейных пространств, то их произведение

$$h = gf : U \to W$$

является линейным отображением.

Доказательство. Пусть $u, u_1, u_2 \in {}_KU$ и $\lambda \in K$. Тогда

$$\begin{aligned} h(u_1 + u_2) &= (gf)(u_1 + u_2) = g(f(u_1 + u_2)) = \\ &= g(f(u_1) + f(u_2)) = g(f(u_1)) + g(f(u_2)) = h(u_1) + h(u_2); \\ h(\lambda u) &= (gf)(\lambda u) = g(f(\lambda u)) = g(\lambda f(u)) = \lambda(g(f(u))) = \lambda h(u). \end{aligned}$$ □

7.2. Матрица произведения линейных отображений пространств столбцов

Если $U = \hat{K}^n$, $V = \hat{K}^m$, $W = \hat{K}^r$ — пространства столбцов над полем K, линейное отображение $f : \hat{K}^n \to \hat{K}^m$ задаётся $(m \times n)$-мат-

рицей $F = (f_{ij})$, линейное отображение $g\colon \hat{K}^m \to \hat{K}^r$ задаётся $(r \times m)$-матрицей $G = (g_{ij})$, то вычислим однозначно определённую матрицу линейного отображения $h = gf\colon \hat{K}^n \to \hat{K}^r$.

Пусть

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \hat{K}^n,$$

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = f(X) \in \hat{K}^m, \quad Z = \begin{pmatrix} z_1 \\ \vdots \\ z_r \end{pmatrix} = g(Y) \in \hat{K}^r.$$

Тогда для $1 \leqslant k \leqslant r$

$$z_k = \sum_{i=1}^{m} g_{ki} y_i = \sum_{i=1}^{m} g_{ki} \left(\sum_{l=1}^{n} f_{il} x_l \right) = \sum_{i=1}^{m} \sum_{l=1}^{n} g_{ki} f_{il} x_l =$$

$$\overset{(*)}{=} \sum_{l=1}^{n} \sum_{i=1}^{m} g_{ki} f_{il} x_l = \sum_{l=1}^{n} \left(\sum_{i=1}^{m} g_{ki} f_{il} \right) x_l = \sum_{l=1}^{n} h_{kl} x_l,$$

где

$$h_{kl} = \sum_{i=1}^{m} g_{ki} f_{il} = g_{k1} f_{1l} + \ldots + g_{km} f_{ml},$$

т. е. матрицей линейного отображения $h = gf$ является $(r \times n)$-матрица $H = (h_{kl})$.

Замечание $(*)$. Использованное в доказательстве равенство

$$\sum_{i=1}^{m} \left(\sum_{l=1}^{n} \gamma_{il} \right) = \sum_{l=1}^{n} \left(\sum_{i=1}^{m} \gamma_{il} \right)$$

означает разный порядок суммирования элементов прямоугольной $(m \times n)$-матрицы $(\gamma_{il}) \in \mathrm{M}_{m,n}(K)$.

Это вычисление приводит нас к следующему определению *произведения согласованных по размеру* матриц.

Определение 7.2.1. Пусть

$$G = (g_{ij}) \in \mathrm{M}_{r,m}(K), \quad F = (f_{ij}) \in \mathrm{M}_{m,n}(K) -$$

прямоугольные матрицы согласованных размеров (т. е. длина m строки матрицы G совпадает с длиной m столбца матрицы F). Тогда определим *произведение* $H = GF$ как $(r \times n)$-матрицу $H = (h_{kl})$, где

$$h_{kl} = \sum_{i=1}^{m} g_{ki} f_{il} = g_{k1} f_{1l} + \ldots + g_{km} f_{ml}.$$

Таким образом, нами фактически доказана

Теорема 7.2.2. *Для диаграммы*

$$\hat{K}^n \xrightarrow{f} \hat{K}^m \xrightarrow{g} \hat{K}^r$$

с линейными отображениями, задаваемыми матрицами $F = (f_{ij}) \in \in \mathrm{M}_{m,n}(K)$ *и* $G = (g_{ij}) \in \mathrm{M}_{r,m}(K)$ *соответственно, произведение*

$$h = gf \colon \hat{K}^n \to \hat{K}^r$$

является линейным отображением, задаваемым матрицей $H = (h_{ij})$, *являющейся произведением* $H = GF$ *матриц линейных отображений* G *и* F. □

Глава 8

Алгебра матриц

8.1. Линейное пространство $\mathrm{M}_{m,n}(K)$ прямоугольных матриц размера $m \times n$

Через $\mathrm{M}_{m,n}(K)$ обозначим совокупность всех прямоугольных матриц над полем K фиксированного размера $m \times n$ (для краткости обозначения, $\mathrm{M}_n(K) = \mathrm{M}_{n,n}(K)$ — совокупность всех квадратных $(n \times n)$-матриц). Как для пространства строк $K^n = \mathrm{M}_{1,n}(K)$ и для пространства столбцов $\hat{K}^n = \mathrm{M}_{n,1}(K)$, так и для $\mathrm{M}_{m,n}(K)$ определены операции *сложения* матриц

$$C = A + B \quad (c_{ij} = a_{ij} + b_{ij} \text{ для каждого места } (i,j))$$

и *умножения* матрицы на число $c \in K$

$$D = cA \quad (d_{ij} = ca_{ij} \text{ для каждого места } (i,j)).$$

Как и для совокупности строк $K^n = \mathrm{M}_{1,n}(K)$, так и для $\mathrm{M}_{m,n}(K)$ непосредственно проверяется выполнение всех аксиом линейного пространства (в частности, нейтральным элементом в $\mathrm{M}_{m,n}(K)$ будет нулевая матрица 0 с нулями на всех местах, $-A = (-1)A$).

8.2. Произведение матриц

Если

$$A = (a_{ij}) \in \mathrm{M}_{r,m}(K), \quad B = (b_{ij}) \in \mathrm{M}_{m,n}(K)$$

то мы определили их *произведение*

$$AB = U = (u_{ij}) \in \mathrm{M}_{r,n}(K),$$

полагая

$$u_{il} = \sum_{k=1}^{m} a_{ik} b_{kl}$$

(т. е. элемент матрицы AB, стоящий на пересечении i-й строки и j-го столбца получается «умножением» i-й строки (длины m) матрицы A на j-й столбец (длины m) матрицы B). Таким образом, условие возможности перемножить две прямоугольные матрицы A и B заключается в том, что *длина строк левого множителя A совпадает с длиной столбцов правого множителя B.*

Примеры вычисления произведения AB

Пример 8.2.1.

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & m+n \\ 0 & 1 \end{pmatrix}, \quad m, n \in \mathbb{Z}.$$

Пример 8.2.2.

$$\begin{pmatrix} k_1 & \dots & k_n \end{pmatrix} \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix} = (k_1 l_1 + \dots k_n l_n) \in \mathrm{M}_1(K) = K.$$

Пример 8.2.3.

$$\begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -5 & -2 \\ 3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} -1 & -10 & -8 \\ 1 & -5 & -2 \\ 9 & 15 & 22 \end{pmatrix};$$

$$\begin{pmatrix} 1 & -5 & -2 \\ 3 & 0 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & -9 \\ -6 & 13 \end{pmatrix}.$$

Пример 8.2.4. Пусть

$$E_r = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathrm{M}_r(K)$$

(единичная матрица размера $r \times r$), $A \in \mathrm{M}_{r,m}(K)$, тогда $E_r A = A$, $AE_m = A$. В частности, если $E = E_n$, $A \in \mathrm{M}_n(K)$, то $EA = A = AE$.

8.3. Матричные единицы E_{ij}

Обозначим через E_{ij} матрицу, в которой на пересечении i-й строки и j-го столбца стоит 1, а на всех остальных местах стоит 0. Тогда в $\mathrm{M}_n(K)$ имеем

$$E_{ij}E_{kl} = \begin{cases} E_{il}, & \text{если } j = k, \\ 0 \text{ (нулевая матрица)}, & \text{если } j \neq k \end{cases}$$

(или $E_{ij}E_{kl} = \delta_{jk}E_{il}$, где

$$\delta_{jk} = \begin{cases} 1, & \text{если } j = k, \\ 0, & \text{если } j \neq k \end{cases} \quad -$$

символ Кронекера).

Важные следствия умножения матричных единиц

Следствие 8.3.1. *Так как в* $\mathrm{M}_n(K)$ *при* $n \geqslant 2$

$$E_{11}E_{12} = E_{12} \neq 0 = E_{12}E_{11},$$

то:

а) *умножение матриц некоммутативно*;

б) *имеются делители нуля* (*ненулевые элементы, произведение которых равно нулю*).

Задача 8.3.2. Найти в $\mathrm{M}_n(K)$ все делители нуля. Точнее, доказать, что для $A \in \mathrm{M}_n(K)$ следующие условия равносильны:

1) $AX = 0$ для некоторой матрицы $0 \neq X \in \mathrm{M}_n(K)$;

2) $YA = 0$ для некоторой матрицы $0 \neq Y \in \mathrm{M}_n(K)$;

3) $|A| = 0$.

Матрицы элементарных преобразований

Следствие 8.3.3. *Пусть* $c \in K$, $i \neq j$, *и*

$$e_{ij}^c = E + cE_{ij} \in \mathrm{M}_m(K)$$

(в этой матрице в отличие от единичной матрицы на месте (i,j) *вне диагонали стоит* c*). Ясно, что* $|e_{ij}^c| = 1$.

а) *Если* $i \neq j$, $e_{ij}^c \in \mathrm{M}_m(K)$ *и* $A \in \mathrm{M}_{m,n}(K)$, *то матрица* $A' = e_{ij}^c A$ *получается из матрицы* A *элементарным преобразованием строк 1-го типа:* $A'_i = A_i + cA_j$.

б) *Если* $i \neq j$, $e_{ij}^c \in \mathrm{M}_n(K)$ *и* $A \in \mathrm{M}_{m,n}(K)$, *то матрица* $A' = Ae_{ij}^c$ *получается из матрицы* A *элементарным преобразованием столбцов 1-го типа:* $\hat{A}'_j = \hat{A}_j + c\hat{A}_i$.

Следствие 8.3.4. *Пусть* $i \neq j$ *и* t_{ij} *— матрица, полученная из единичной матрицы* $E_m \in \mathrm{M}_m(K)$ *перестановкой* i*-й и* j*-й строк (или, что то же самое, перестановкой* i*-го и* j*-го столбцов). Ясно, что* $|t_{ij}| = -1$.

а) *Если* $t_{ij} \in \mathrm{M}_m(K)$ *и* $A \in \mathrm{M}_{m,n}(K)$, *то матрица* $A' = t_{ij}A$ *получается из матрицы* A *элементарным преобразованием строк 2-го типа:* $A'_i = A_j$, $A'_j = A_i$.

б) *Если* $t_{ij} \in \mathrm{M}_n(K)$ *и* $A \in \mathrm{M}_{m,n}(K)$, *то матрица* $A' = At_{ij}$ *получается из матрицы* A *элементарным преобразованием столбцов 2-го типа:* $\hat{A}'_i = \hat{A}_j$, $\hat{A}'_j = \hat{A}_i$.

Следствие 8.3.5. *Пусть* $\lambda_1, \ldots, \lambda_m \in K$,

$$d(\lambda_1, \ldots, \lambda_m) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_m \end{pmatrix} \in \mathrm{M}_m(K) -$$

диагональная матрица с элементами $\lambda_1, \lambda_2, \ldots, \lambda_m \in K$ *на диагонали. Ясно, что* $|d(\lambda_1, \ldots, \lambda_m)| = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_m$.

а) *Если* $d(\lambda_1, \ldots, \lambda_m) \in \mathrm{M}_m(K)$ *и* $A \in \mathrm{M}_{m,n}(K)$, *то*

$$d(\lambda_1, \ldots, \lambda_m)A = \begin{pmatrix} \lambda_1 A_1 \\ \lambda_2 A_2 \\ \vdots \\ \lambda_m A_m \end{pmatrix} -$$

матрица, получаемая из матрицы A *умножением строк* $A_1, \ldots, A_m$ *соответственно на «числа»* $\lambda_1, \ldots, \lambda_m$.

б) *Если* $d(\lambda_1, \ldots, \lambda_n) \in \mathrm{M}_n(K)$ *и* $A \in \mathrm{M}_{m,n}(K)$, *то*

$$A\, d(\lambda_1, \ldots, \lambda_n) = (\lambda_1 \hat{A}_1, \ldots, \lambda_n \hat{A}_n) -$$

матрица, получаемая из матрицы A *умножением столбцов* $\hat{A}_1, \ldots, \hat{A}_n$ *соответственно на «числа»* $\lambda_1, \ldots, \lambda_n$.

В частности, умножение слева матрицы A *на матрицу* $d(1, \ldots, \lambda_i = c, \ldots, 1)$, $c \neq 0$, *равносильно применению к строкам матрицы* A *элементарного преобразования 3-го типа* $A'_i = cA_i$ (*умножение справа на матрицу такого типа даёт применение к столбцам матрицы* A *элементарного преобразования 3-го типа* $\hat{A}'_i = c\hat{A}_i$).

Замечание 8.3.6. Ясно, что $\lambda E = d(\lambda, \ldots, \lambda)$ и $(\lambda E)A = \lambda A = A(\lambda E)$ для $E = E_n$, $A \in \mathrm{M}_n(K)$, т. е. *скалярная* матрица λE перестановочна с любой другой матрицей из $\mathrm{M}_n(K)$.

Задача 8.3.7. Пусть K — поле, $n \in \mathbb{N}$, $n \geqslant 2$,

$$Z(\mathrm{M}_n(K)) = \{A \in \mathrm{M}_n(K) \mid AB = BA \ \ \forall B \in \mathrm{M}_n(K)\}.$$

Тогда $A \in Z(\mathrm{M}_n(K))$ в том и только в том случае, когда $A = \lambda E_n$, $\lambda \in K$.

Следствие 8.3.8 (матричная запись системы линейных уравнений). *Для системы линейных уравнений*

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \vdots \qquad\qquad \vdots \qquad\quad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

возможна матричная запись

$$AX = B,$$

где $A = (a_{ij})$ *—* (m, n)*-матрица коэффициентов,*

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} -$$

столбец неизвестных,

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathrm{M}_{m,1}(K) -$$

столбец свободных членов.

Таким образом, строка $(k_1, \ldots, k_n)$ *является решением системы линейных уравнений, если столбец*

$$\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} \in \mathrm{M}_{n,1}(K)$$

является решением матричного уравнения

$$A \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Замечание 8.3.9 (Штрассен, 1969). Умножение двух (2×2)-матриц можно осуществить с использованием 7 умножений и 18 сложений (вместо 8 умножений и 4 сложений в обычном определении произведения матриц):

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} (a-d)(e-h)+(b-d)(g+h)+ & -(a-b)h+a(h+f) \\ +d(e+g)+(a-b)h & \\ (-d+c)e+d(e+g) & (a-d)(e-h)-(a-c)(e+f)+ \\ & +a(h+f)-(c-d)e \end{pmatrix}.$$

Это соображение развивает идею алгоритма А. А. Карацубы (1962 г.) быстрого умножения многочленов. Дальнейший прогресс в теории быстрого умножения чисел, многочленов, матриц связан, в частности, с использованием быстрого преобразования Фурье.

8.4. Ассоциативность произведения матриц

Теорема 8.4.1 (об ассоциативности произведения матриц). *Пусть*

$$A = (a_{ij}) \in \mathrm{M}_{r,m}(K),$$
$$B = (b_{ij}) \in \mathrm{M}_{m,n}(K),$$
$$C = (c_{ij}) \in \mathrm{M}_{n,p}(K).$$

Тогда

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

Первое доказательство. Пусть

$$U = AB = (u_{ij}) \in \mathrm{M}_{r,n}(K),$$
$$V = BC = (v_{ij}) \in \mathrm{M}_{m,p}(K),$$
$$S = (AB)C = UC = (s_{ij}) \in \mathrm{M}_{r,p}(K),$$
$$T = A(BC) = AV = (t_{ij}) \in \mathrm{M}_{r,p}(K).$$

Так как

$$u_{il} = \sum_{k=1}^{m} a_{ik}b_{kl}, \quad v_{kj} = \sum_{l=1}^{n} b_{kl}c_{lj},$$

то

$$s_{ij} = \sum_{l=1}^{n} u_{il}c_{lj} = \sum_{l=1}^{n}\sum_{k=1}^{m} a_{ik}b_{kl}c_{lj},$$
$$t_{ij} = \sum_{k=1}^{m} a_{ik}v_{kj} = \sum_{k=1}^{m}\sum_{l=1}^{n} a_{ik}b_{kl}c_{lj}$$

для всех (i, j), и, следовательно, $S = T$.

Второе доказательство. Пусть в диаграмме

$$\hat{K}^p \xrightarrow[C]{\mathcal{C}} \hat{K}^n \xrightarrow[B]{\mathcal{B}} \hat{K}^m \xrightarrow[A]{\mathcal{A}} \hat{K}^r$$

линейные преобразования $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ определены соответственно матрицами A, B, C. Тогда (в силу ассоциативности произведения отображений) $(\mathcal{AB})\mathcal{C} = \mathcal{A}(\mathcal{BC})$. Вычисляя матрицу этого линейного преобразования (по теореме о матрице произведения линейных преобразований), получаем, что $(AB)C = A(BC)$. □

Следствие 8.4.2. *Квадратные $(n \times n)$-матрицы $\mathrm{M}_n(K)$ относительно операции умножения являются моноидом (т. е. операция умножения определена на $\mathrm{M}_n(K)$, ассоциативна и обладает нейтральным элементом $E = E_n$).*

Теорема 8.4.3 (о дистрибутивности для матриц). *Пусть*

$$A = (a_{ij}), B = (b_{ij}) \in \mathrm{M}_{m,n}(K);$$
$$C = (c_{ij}) \in \mathrm{M}_{r,m}(K); \quad D = (d_{ij}) \in \mathrm{M}_{n,p}(K).$$

Тогда

$$C(A+B) = CA + CB \quad \text{в} \quad \mathrm{M}_{r,n}(K),$$
$$(A+B)D = AD + BD \quad \text{в} \quad \mathrm{M}_{m,p}(K).$$

Доказательство. Действительно, для любого места (i,j) имеем

$$\sum_{k=1}^{m} c_{ik}(a_{kj} + b_{kj}) = \sum_{k=1}^{m} c_{ik}a_{kj} + \sum_{k=1}^{m} c_{ik}b_{kj},$$
$$\sum_{l=1}^{n} (a_{il} + b_{il})d_{lj} = \sum_{l=1}^{n} a_{il}d_{lj} + \sum_{l=1}^{n} b_{il}d_{lj},$$

что доказывает наши утверждения. □

Следствие 8.4.4. *Для любых квадратных матриц $A, B, C \in \in \mathrm{M}_n(K)$ имеем*

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

8.5. Итоговая теорема об алгебре матриц

Теорема 8.5.1.

I. *Совокупность $\mathrm{M}_{m,n}(K)$ прямоугольных матриц размера $m \times n$ над K (в частности, квадратные матрицы $\mathrm{M}_n(K)$) относительно операции сложения образуют* абелеву (коммутативную) группу, *т. е.*

I.1) *операция сложения* ассоциативна;

I.1′) *операция сложения* коммутативна (*т. е.* $A+B=B+A$ *для всех* $A, B \in \mathrm{M}_{m,n}(K)$);

I.2) *существует* нейтральный элемент 0 (*нулевая матрица*), $0+A=A+0=A$ *для всех* $A \in \mathrm{M}_{m,n}(K)$;

I.3) *для каждой матрицы* $A \in \mathrm{M}_{m,n}(K)$ *существует* противоположный элемент $-A\,(=(-1)A)=(-a_{ij})$, $A+(-A)=0$.

II. *Операции умножения матрицы* A *на элемент* $c \in K$, $A \mapsto cA$, *в* $\mathrm{M}_{m,n}(K)$ *удовлетворяют условиям*:

II.1) $1 \cdot A = A$;

II.2) $(c_1c_2)A = c_1(c_2A)$.

III. *Операции сложения и умножения на элементы* $c \in K$ *в* $\mathrm{M}_{m,n}(K)$ *удовлетворяют условиям*

III.1) $c(A+B)=cA+cB$;

III.2) $(c_1+c_2)A=c_1A+c_2A$.

Таким образом, I, II, III *означают, что* $\mathrm{M}_{m,n}(K)$ — линейное пространство *над полем* K.

IV. *С операциями сложения* $A+B$ *и умножения матриц* AB *совокупность квадратных матриц* $\mathrm{M}_n(K)$ *является* кольцом, *т. е.*

IV.1) *по сложению* $\mathrm{M}_n(K)$ — абелева группа;

IV.2) *с умножением матриц* $\mathrm{M}_n(K)$ — моноид, *т. е.*

2а) *умножение матриц ассоциативно,*

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

для любых $A, B, C \in \mathrm{M}_n(K)$;

2б) *единичная матрица* E *является нейтральным элементом для операции умножения,*

$$AE = A = EA$$

для всех $A \in \mathrm{M}_n(K)$;

IV.3) *операции сложения и умножения матриц удовлетворяют законам дистрибутивности*

3а) $(A+B)C = AC + BC$;

3б) $C(A+B) = CA + CB$.

V. *С операциями сложения $A+B$ и умножения AB матриц и операциями умножения cA матрицы A на элемент $c \in K$ квадратные матрицы $\mathrm{M}_n(K)$ являются* алгеброй, *т. е.*

V.1) кольцом (*относительно сложения и умножения матриц*);

V.2) *линейным пространством* (*относительно сложения матриц и умножений матрицы на элемент K*)

и дополнительно

V.3) $(cA)B = c(AB) = A(cB)$ *для* $c \in K$, $A, B \in \mathrm{M}_n(K)$.

Доказательство свойства V.3. Для любого места (i,j) имеем

$$\sum_{k=1}^{n}(ca_{ik})b_{kj} = c\sum_{k=1}^{n}a_{ik}b_{kj} = \sum_{k=1}^{n}a_{ik}(cb_{kj}). \quad \square$$

Теорема 8.5.2 (о транспонировании произведения матриц). *Пусть*

$$A \in \mathrm{M}_{m,n}(K), \quad B \in \mathrm{M}_{n,r}(K),$$

тогда

$$(AB)^* = B^*A^*.$$

Доказательство. Ясно, что $G = AB \in \mathrm{M}_{m,r}(K)$ и $H = G^* = (AB)^* \in \mathrm{M}_{r,m}(K)$. Так как $D = B^* \in \mathrm{M}_{r,n}(K)$ и $C = A^* \in \mathrm{M}_{n,m}(K)$, то произведение $U = B^*A^*$ существует и лежит в $\mathrm{M}_{r,m}(K)$, как и $(AB)^* \in \mathrm{M}_{r,m}(K)$.

Для любого места (i,j) имеем для $U = B^*A^* = DC$, где $B^* = D$, $A^* = C$,

$$u_{ij} = \sum_{k=1}^{n} d_{ik}c_{kj} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki} = g_{ji} = h_{ij}.$$

Итак, $B^*A^* = U = H = (AB)^*$. $\square$

Теорема 8.5.3 (об определителе произведения матриц). *Для любых квадратных матриц $A, B \in \mathrm{M}_n(K)$ имеем*

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

Доказательство. Пусть $C = AB$. Рассмотрим определитель размера $2n \times 2n$:

$$\left|\begin{array}{c|c} A & 0 \\ \hline \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} & B \end{array}\right| = \left|\begin{array}{c|c} A^* & \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} \\ \hline 0 & B^* \end{array}\right| = |A^*|\,|B^*| = |A|\,|B|.$$

С другой стороны, прибавляя к каждому столбцу, проходящему через матрицу B, соответствующую линейную комбинацию столбцов, проходящих через матрицу A, т. е.

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hline b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} + b_{1j} \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \\ \hline -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + b_{nj} \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \\ \hline 0 \\ \vdots \\ -1 \end{pmatrix} = \begin{pmatrix} c_{11} \\ \vdots \\ c_{n1} \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix},$$

получаем, что

$$\left|\begin{array}{c|c} A & 0 \\ \hline \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} & B \end{array}\right| = \left|\begin{array}{c|c} A & C = AB \\ \hline \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} & 0 \end{array}\right| =$$

$$= (-1)^n \left|\begin{array}{c|c} \begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix} & 0 \\ \hline A & C \end{array}\right| = (-1)^{2n}|C| = |C|. \quad \square$$

Следствие 8.5.4.

1) *Если* $A, B \in \mathrm{M}_n(K)$, *то* $|AB| = |A|\,|B| = |BA|$ (*т. е. хотя матрицы* AB *и* BA *могут быть различны, их определители равны*).

2) $|A^k| = |A|^k$ *для* $k \in \mathbb{N}$.

Упражнение 8.5.5. Покажите, что любые две матрицы в $\mathrm{M}_2(\mathbb{Z})$, коммутирующие с

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

коммутируют между собой.

Упражнение 8.5.6. Если

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z}),$$

то

$$A^m = \begin{pmatrix} f_{m-1} & f_m \\ f_m & f_{m+1} \end{pmatrix}$$

для $m \in \mathbb{N}$, где

$$f_0 = 0, \quad f_1 = 1, \quad f_2 = 1, \quad f_3 = 2,$$
$$f_{m+1} = f_m + f_{m-1}$$

(числа Фибоначчи). Если

$$B = \begin{pmatrix} -\dfrac{\lambda_2}{5} & \dfrac{1}{5} \\ -\sqrt{5}\lambda_1 & \sqrt{5} \end{pmatrix},$$

где

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2},$$

то

$$A = B^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} B,$$

и поэтому

$$A^m = B^{-1} \begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix} B,$$

откуда

$$f_m = \frac{\lambda_1^m - \lambda_2^m}{\sqrt{5}} = \frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^m - \left(\frac{1-\sqrt{5}}{2}\right)^m\right\},$$

$$f_m \sim \frac{1}{\sqrt{5}}\lambda_1^m,$$

поскольку

$$\lim_{m\to\infty}\left(\frac{1-\sqrt{5}}{2}\right)^m = 0.$$

Упражнение 8.5.7. Пусть $A \in \mathrm{M}_n(K)$. Покажите, что

$$\{B \in \mathrm{M}_n(K) \mid AB = BA\} -$$

подалгебра в алгебре матриц $\mathrm{M}_n(K)$.

Упражнение 8.5.8. Если $D = d(\lambda_1, \ldots, \lambda_n) \in \mathrm{M}_n(K)$, $\lambda_i \neq \lambda_j$ при $i \neq j$, $A \in \mathrm{M}_n(K)$ и $DA = AD$, то A — также диагональная матрица.

Упражнение 8.5.9 (внешнее произведение векторов). Если $A \in \mathrm{M}_{m,1}(K)$, $B \in \mathrm{M}_{1,r}(K)$, то $C = AB \in \mathrm{M}_{m,r}(K)$:

$$\begin{pmatrix}1\\2\\3\end{pmatrix}\begin{pmatrix}4 & 5\end{pmatrix} = \begin{pmatrix}4 & 5\\8 & 10\\12 & 15\end{pmatrix}.$$

Упражнение 8.5.10 (скалярное произведение векторов). Если $A \in \mathrm{M}_{1,n}(K)$, $B \in \mathrm{M}_{n,1}(K)$, то $C = AB \in \mathrm{M}_1(K) = K$:

$$\begin{pmatrix}1 & 2 & 3\end{pmatrix}\begin{pmatrix}4\\5\\6\end{pmatrix} = (1\cdot 4 + 2\cdot 5 + 3\cdot 6) = (32).$$

Упражнение 8.5.11. Пусть $H = (h_{ij}) \in \mathrm{M}_n(\mathbb{R})$, $h_{ij} \in \{1, -1\}$, $H^*H = nE\,(= HH^*)$ (такая матрица называется *матрицей Адамара*). Например,

$$\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}, \quad \begin{pmatrix}1 & 1 & 1 & 1\\1 & -1 & 1 & -1\\1 & 1 & -1 & -1\\1 & -1 & -1 & 1\end{pmatrix}.$$

Докажите, что для n, отличных от 1, 2 и $4k$, где $k \in \mathbb{N}$, не существует матриц Адамара.

Упражнение 8.5.12. Пусть $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{C})$ — *матрица Маркова* (это означает, что для каждого j, $1 \leqslant j \leqslant n$, $\sum\limits_{i=1}^{n} a_{ij} = 1$, т. е. сумма элементов по каждому столбцу равна 1). Докажите, что если $A, B \in \mathrm{M}_n(\mathbb{C})$ — матрицы Маркова, то

1) AB и A^k, $k \in \mathbb{N}$, — матрицы Маркова;

2) если $|a_{ij}| \leqslant 1$ и $|b_{ij}| \leqslant 1$ для всех i, j, то $|c_{ij}| \leqslant 1$ для всех i, j для матрицы $(c_{ij}) = C = AB$.

8.6. Многочлены от матриц, теорема Гамильтона—Кэли

Пусть K — поле,

$$f(t) = a_0 + a_1 t + \ldots + a_n t^n \in K[t] -$$

многочлен с коэффициентами из поля K, $A \in \mathrm{M}_n(K)$. Тогда определим

$$f(A) = a_0 E + a_1 A + \ldots + a_n A^n \in \mathrm{M}_n(K),$$

где

$$E = E_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathrm{M}_n(K) -$$

единичная $(n \times n)$-матрица, т. е.

$$f(A) = \sum_{i=0}^{n} a_i A^i,$$

здесь $A^0 = E$.

Пример 8.6.1. Пусть $f(t) = t^2 + 2t + 1 = (t+1)^2, g(t) = t+1 \in \mathbb{R}[t]$,

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}).$$

Тогда

$$f(A) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 + 2\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$
$$\left(= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}^2 = \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)^2 = (g(A))^2 \right).$$

Упражнение 8.6.2. Пусть

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(K)$$

и

$$f(\lambda) = |A - \lambda E| = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - bc =$$
$$= \lambda^2 - (a+d)\lambda + (ad - bc)$$
$$(= \lambda^2 - \operatorname{tr} A\lambda + |A|) -$$

характеристический многочлен матрицы A (здесь $\operatorname{tr} A = a + d$). Тогда

$$f(A) = A^2 - (a+d)A + (ad-bc)E =$$
$$= \begin{pmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{pmatrix} - \begin{pmatrix} (a+d)a & (a+d)b \\ (a+d)c & (a+d)d \end{pmatrix} +$$
$$+ \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

(т. е. в этом частном случае мы видим, что справедлива теорема Гамильтона—Кэли о том, что матрица A является корнем своего характеристического многочлена $f(\lambda) = |A - \lambda E|$ для (2×2)-матриц).

Теорема 8.6.3. *Пусть K — поле, $A \in \mathrm{M}_n(K)$,*

$$\Delta_A \colon K[t] \to \mathrm{M}_n(K) -$$

отображение, для которого $\Delta_A(f(t)) = f(A)$ для $f(t) \in K[t]$. Тогда

1) $\Delta = \Delta_A$ — *гомоморфизм* K*-алгебр, т. е.*

$$\Delta(f+g) = (f+g)(A) = f(A) + g(A) = \Delta(f) + \Delta(g),$$
$$\Delta(fg) = (fg)(A) = f(A)g(A) = \Delta(f)\Delta(g),$$
$$\Delta(\lambda f) = (\lambda f)(A) = \lambda f(A) = \lambda\Delta(f)$$

для всех $f, g \in K[t]$, $\lambda \in K$;

2) $\operatorname{Ker}\Delta_A = \{f(t) \in K[t] \mid f(A) = 0\}$ — *ненулевой идеал кольца* $K[t]$.

Доказательство.

1) Пусть

$$f(t) = a_0 + a_1 t + \ldots + a_n t^n, \quad g(t) = b_0 + b_1 t + \ldots + b_m t^m,$$

где $a_i, b_j \in K$, и пусть $\lambda \in K$. Тогда

а) если $n \geqslant m$, то

$$(f+g)(A) = \sum_{i=0}^{n}(a_i + b_i)A^i = \sum_{i=0}^{n} a_i A^i + \sum_{i=0}^{m} b_i A^i = f(A) + g(A)$$

(здесь $b_n = \ldots = b_{m+1} = 0$);

б) если

$$(fg)(t) = c_0 + c_1 t + \ldots + c_{m+n} t^{m+n},$$

где

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

то

$$(fg)(A) = \sum_{k=0}^{m+n} c_k A^k;$$

с другой стороны,

$$f(A)g(A) = \left(\sum_{i=0}^{n} a_i A^i\right)\left(\sum_{j=0}^{m} b_j A^j\right) =$$
$$= \sum_{k=0}^{m+n}\left(\sum_{i=0}^{k} a_i b_{k-i}\right)A^k = \sum_{k=0}^{m+n} c_k A^k,$$

т. е. $(fg)(A) = f(A)g(A)$;

в)

$$(\lambda f)(A) = \sum_{i=0}^{n} (\lambda a_i) A^i = \lambda \left(\sum_{i=0}^{n} a_i A^i \right) = \lambda f(A).$$

2) Если $f(t), g(t) \in \operatorname{Ker} \Delta$, $h(t) \in K[t]$, $\lambda \in K$, то $f(A) = 0$, $g(A) = 0$, и поэтому

$$\begin{aligned} (f+g)(A) &= f(A) + g(A) = 0 + 0 = 0, \\ (fh)(A) &= f(A)h(A) = 0 \cdot h(A) = 0, \\ (\lambda f)(A) &= \lambda f(A) = \lambda \cdot 0 = 0. \end{aligned}$$

Итак, $\operatorname{Ker} \Delta \lhd K[t]$ (т. е. $\operatorname{Ker} \Delta$ — идеал K-алгебры $K[t]$).

Так как система матриц

$$E, A, A^2, \ldots, A^{n^2+1}$$

линейно зависима в $\mathrm{M}_n(K)$ (поскольку $\dim_K \mathrm{M}_n(K) = n^2$), то найдутся (не все нулевые) элементы $a_0, a_1, \ldots, a_{n^2+1} \in K$, для которых

$$a_0 E + a_1 A + \ldots + a_{n^2+1} A^{n^2+1} = 0,$$

т. е.

$$0 \neq f(t) = a_0 + a_1 t + \ldots + a_{n^2+1} t^{n^2+1} \in \operatorname{Ker} \Delta.$$

Итак, $\operatorname{Ker} \Delta \neq 0$. □

Замечание 8.6.4. Более сильное утверждение о том, что

$$|A - tE| \in \operatorname{Ker} \Delta,$$

является содержанием следующей теоремы.

Теорема 8.6.5 (теорема Гамильтона—Кэли). *Пусть K — поле (или даже коммутативное ассоциативное кольцо с 1), $A \in \mathrm{M}_n(K)$, $p(t) = |A - tE| \in K[t]$ — характеристический многочлен квадратной матрицы A, $\deg p(t) = n$. Тогда*

$$p(A) = 0 \in \mathrm{M}_n(K).$$

Доказательство. Для матрицы

$$D = A - tE = (d_{ij}) \in \mathrm{M}_n(K[t]),$$

$d_{ij} \in K[t]$, рассмотрим присоединённую матрицу

$$B = (b_{ij}) \in \mathrm{M}_n(K[t]),$$

$b_{ij} = D_{ji} \in K[t]$ — алгебраическое дополнение элемента d_{ji}. Тогда $\deg(b_{ij}(t)) \leqslant n-1$, и поэтому

$$B = B(t) = B_0 + tB_1 + \ldots + t^{n-1}B_{n-1},$$

где $B_i \in \mathrm{M}_n(K)$. Так как

$$p(t) = |A - tE| = (-1)^n t^n + c_{n-1}t^{n-1} + \ldots + c_1 t + c_0,$$

где $c_i \in K$, $i = 0, 1, \ldots, n-1$, $D \cdot B = |D| \cdot E$, то

$$(A - tE)B(t) = p(t)E. \tag{8.1}$$

Приравнивая матричные коэффициенты при степенях t^k, $0 \leqslant k \leqslant n$, в левой и правой частях этого равенства, получаем:

$$\begin{aligned}
&t^n: && -B_{n-1} = (-1)^n E,\\
&t^{n-1}: && A \cdot B_{n-1} - B_{n-2} = c_{n-1}E,\\
&t^{n-2}: && A \cdot B_{n-2} - B_{n-3} = c_{n-2}E,\\
&\ldots && \ldots\\
&t: && A \cdot B_1 - B_0 = c_1 E,\\
&t^0: && A \cdot B_0 = c_0 E.
\end{aligned} \tag{8.2}$$

Умножая слева равенства (8.2) на $A^n, A^{n-1}, \ldots, A, E$ соответственно, получаем

$$\begin{aligned}
&-A^n \cdot B_{n-1} = (-1)^n A^n,\\
&A^n \cdot B_{n-1} - A^{n-1} \cdot B_{n-2} = c_{n-1}A^{n-1},\\
&\ldots\\
&A^2 \cdot B_1 - A \cdot B_0 = c_1 A,\\
&A \cdot B_0 = c_0 E.
\end{aligned} \tag{8.3}$$

Складывая равенства (8.3), получаем

$$\mathrm{M}_n(K) \ni 0 = p(A). \quad \square$$

Замечание 8.6.6. Отметим, что равенства (8.2) показывают, что матрицы $B_0, B_1, \ldots, B_{n-1}$ являются многочленами от матрицы A, в частности, $B_iA = AB_i$, $i = 0, 1, \ldots, n-1$. Поэтому можно было подставить в (8.1) вместо переменной t матрицу A, и тогда

$$\mathrm{M}_n(K) \ni 0 = (A - AE)(B_0 + AB_1 + \ldots + A^{n-1}B_{n-1}) =$$
$$\overset{(8.1)}{=} p(A) \cdot E = p(A). \quad \square$$

Замечание 8.6.7. Очевидное равенство $|A - AE| = 0 \in K$ не является доказательством теоремы Гамильтона—Кэли.

Упражнение 8.6.8. Аннулирующий многочлен минимальной степени $\varphi_A(t)$ жордановой клетки r-го порядка

$$A = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ 0 & 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda \end{pmatrix}$$

равен

$$\varphi_A(t) = (\lambda - t)^r = |A - tE|.$$

Упражнение 8.6.9. Если

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

то

$$\varphi_A(t) = (1-t)^2, \quad |A - tE| = (1-t)^3.$$

8.7. Обратная матрица

Определение 8.7.1. Пусть $A \in \mathrm{M}_n(K)$ — квадратная матрица. Будем говорить, что матрица $B \in \mathrm{M}_n(K)$ является *обратной* к A, если $AB = E = BA$.

Замечание 8.7.2 (для любой ассоциативной операции). Если обратная матрица B к матрице A существует, то она однозначно определена. Действительно, пусть $AB = E = BA$ и $AC = E = CA$, тогда $C = EC = (BA)C = B(AC) = BE = B$ (это повтор того, что мы уже отмечали ранее: единственность обратного элемента, если он существует, для любого элемента моноида). В этом случае однозначно определённую обратную матрицу B мы будем обозначать через A^{-1}: $AA^{-1} = E = A^{-1}A$.

Теорема 8.7.3 (об обратной матрице). *Пусть $A \in \mathrm{M}_n(K)$ — квадратная $(n \times n)$-матрица. Тогда:*

1) *обратная матрица $B = (b_{ij}) = A^{-1}$ существует тогда и только тогда, когда $|A| \neq 0$;*
2) *в этом случае $b_{ij} = \dfrac{A_{ji}}{|A|}$ (формула для элемента обратной матрицы);*
3) $|A^{-1}| = \dfrac{1}{|A|}$.

Доказательство.

а) Если $AB = E$, то $1 = |E| = |AB| = |A|\,|B|$, поэтому $|A| \neq 0$ и, более того, $|A^{-1}| = |B| = \dfrac{1}{|A|}$.

б) Если $|A| \neq 0$, то рассмотрим $B = (b_{ij})$, где $b_{ij} = \dfrac{A_{ji}}{|A|}$. Ясно, что $AB = E = BA$ (принимая во внимание разложение определителя по строкам и столбцам, а также «фальшивое» разложение), т. е. $B = A^{-1}$. □

Следствие 8.7.4. *Если $A, B \in \mathrm{M}_n(K)$, то из $AB = E$ следует, что $BA = E$ (матрица, имеющая правую обратную, обратима (двусторонне)).*

Доказательство. Если $AB = E$, то $|A|\,|B| = |AB| = |E| = 1$, поэтому $|A| \neq 0$, но тогда существует двусторонняя обратная матрица A^{-1}. Таким образом, $A^{-1} = A^{-1}E = A^{-1}(AB) = (A^{-1}A)B = E \cdot B = B$, следовательно, $BA = A^{-1}A = E$. □

Следствие 8.7.5. *Для $A, B \in \mathrm{M}_n(K)$ имеем $|AB| = |A|\,|B|$, поэтому $|AB| \neq 0$ тогда и только тогда, когда $|A| \neq 0$ и $|B| \neq 0$, т. е.*

обратная матрица $(AB)^{-1}$ *существует тогда и только тогда, когда существуют* A^{-1} *и* B^{-1}. *Более того, в этом случае* $(AB)^{-1} = B^{-1}A^{-1}$.

Доказательство. $(AB)(B^{-1}A^{-1}) = E = (B^{-1}A^{-1})(AB)$. □

Следствие 8.7.6. *Если существуют обратные матрицы* $A_1^{-1}, \ldots, A_r^{-1}$ *для* $A_1, \ldots, A_r \in \mathrm{M}_n(K)$, *то*

$$(A_1 A_2 \cdot \ldots \cdot A_r)^{-1} = A_r^{-1} \cdot \ldots \cdot A_2^{-1} A_1^{-1}.$$

Следствие 8.7.7. *Если существует обратная матрица* A^{-1} *для* $A \in \mathrm{M}_n(K)$, *то* $(A^{-1})^{-1} = A$.

Доказательство. $A^{-1}A = E = AA^{-1}$ (с точки зрения матрицы A^{-1}: $A = (A^{-1})^{-1}$). □

Упражнение 8.7.8. Пусть

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(K), \quad |A| = ad - bc \neq 0.$$

Тогда:

$$B = (b_{ij} = A_{ji}) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix};$$

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

Упражнение 8.7.9. Пусть

$$A = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 0 & 1 & \ldots & 1 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \end{pmatrix}.$$

Тогда

$$A^{-1} = \begin{pmatrix} 1 & -1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ 0 & & & 1 \end{pmatrix}.$$

Упражнение 8.7.10. Найти

$$\begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 0 \end{pmatrix}^{-1} \in \mathrm{M}_n(\mathbb{Q})$$

(матрица размера $n \times n$, на главной диагонали которой стоят нули, а все остальные элементы равны **1**).

Упражнение 8.7.11. Пусть

$$A = \begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ n & 1 & 2 & \dots & n-2 & n-1 \\ n-1 & n & 1 & \dots & n-3 & n-2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 2 & 3 & 4 & \dots & n & 1 \end{pmatrix} \in \mathrm{M}_n(\mathbb{Q}).$$

Тогда

$$A^{-1} = \frac{1}{ns} \begin{pmatrix} 1-s & 1+s & 1 & \dots & 1 & 1 \\ 1 & 1-s & 1+s & \dots & 1 & 1 \\ 1 & 1 & 1-s & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1+s & 1 & 1 & \dots & 1 & 1-s \end{pmatrix},$$

где $s = \dfrac{n(n+1)}{2}$.

Теорема 8.7.12 (о линейных группах).

а) *Множество обратимых матриц*

$$\mathrm{GL}_n(K) = \{A \in \mathrm{M}_n(K) \mid |A| \neq 0\}$$

с операцией умножения является группой (линейная группа).

б) *Множество матриц с единичным определителем*

$$\mathrm{SL}_n(K) = \{A \in \mathrm{M}_n(K) \mid |A| = 1\}$$

с операцией умножения является группой (специальная линейная группа).

Доказательство.

а) Все проверки для $\mathrm{GL}_n(K)$ уже были проведены.

б) Если $A, B \in \mathrm{SL}_n(K)$, то $|A| = 1$, $|B| = 1$, поэтому $|AB| = |A|\,|B| = 1 \cdot 1 = 1$, следовательно, $AB \in \mathrm{SL}_n(K)$. Ясно, что $|E| = 1$, т. е. $E \in \mathrm{SL}_n(K)$. Если $A \in \mathrm{SL}_n(K)$, то $|A| = 1 \neq 0$, т. е. существует A^{-1}, при этом $|A^{-1}| = \dfrac{1}{|A|} = 1$, поэтому $A^{-1} \in \mathrm{SL}_n(K)$. □

Лемма 8.7.13. *Если $A \in \mathrm{GL}_n(K)$ (т. е. $A \in \mathrm{M}_n(K)$ и $|A| \neq 0$), то $|A^*| = |A| \neq 0$ (т. е. $A^* \in \mathrm{GL}_n(K)$) и, более того, $(A^*)^{-1} = (A^{-1})^*$.*

Доказательство.

$$(A^{-1})^* A^* = (AA^{-1})^* = E^* = E;$$
$$A^*(A^{-1})^* = (A^{-1}A)^* = E^* = E$$

(с точки зрения матрицы A^*: $(A^*)^{-1} = (A^{-1})^*$). □

Определение 8.7.14. Квадратная матрица $A \in \mathrm{M}_n(K)$ называется *ортогональной матрицей*, если $A^{-1} = A^*$.

Теорема 8.7.15. *Совокупность ортогональных матриц $\mathrm{O}_n(K) = \{A \in \mathrm{M}_n(K) \mid A^{-1} = A^*\}$ относительно умножения матриц является группой.*

Доказательство.

а) Если $A, B \in \mathrm{O}_n(K)$, то $A^{-1} = A^*$ и $B^{-1} = B^*$. Тогда $(AB)^{-1} = B^{-1}A^{-1} = B^*A^* = (AB)^*$, поэтому $AB \in \mathrm{O}_n(K)$.

б) $E^{-1} = E = E^*$, т. е. $E \in \mathrm{O}_n(K)$.

в) Если $A \in \mathrm{O}_n(K)$, то для $B = A^{-1}$ имеем $B^{-1} = (A^{-1})^{-1} = (A^*)^{-1} = (A^{-1})^* = B^*$, следовательно, $B = A^{-1} \in \mathrm{O}_n(K)$. □

Задача 8.7.16. Пусть $A \in \mathrm{M}_n(K)$ и существует такое число k, что A^k — нулевая матрица. Покажите, что матрицы $E - A$, $E + A$ обратимы (здесь E — единичная матрица в $\mathrm{M}_n(K)$).

Задача 8.7.17. Для $A, B \in \mathrm{M}_n(K)$ равносильны условия:

1) матрица $E - AB$ обратима;

2) матрица $E - BA$ обратима

(этот факт полезен при построении теории определителей над произвольным кольцом R: в алгебраической K-теории — функтор $K_1(R)$).

Более того, можно доказать, что если $A \in \mathrm{M}_{m,n}(K)$, $B \in \in \mathrm{M}_{n,m}(K)$, то $E_m - AB$ — обратимая матрица тогда и только тогда, когда $E_n - BA$ — обратимая матрица.

Задача 8.7.18. Найти число элементов в группах $\mathrm{GL}_2(\mathbb{Z}_2)$, $\mathrm{SL}_2(\mathbb{Z}_2)$, $\mathrm{GL}_n(K)$, где K — конечное поле из q элементов.

Упражнение 8.7.19. Рассмотрим отображение

$$f\colon \mathrm{S}_n \to \mathrm{GL}_n(K),$$

где

$$f(\sigma) = \sum_{j=1}^{n} E_{\sigma(j)j}$$

(т. е. в j-м столбце единственная единица стоит в $\sigma(j)$-й строке, остальные элементы нулевые). Тогда

$$|f(\sigma)| = \varepsilon(\sigma) = \begin{cases} 1, & \sigma \in \mathrm{A}_n, \\ -1 & \sigma \in \mathrm{S}_n \setminus \mathrm{A}_n, \end{cases}$$

поэтому $f(\sigma) \in \mathrm{GL}_n(K)$. Покажите, что f — инъективный гомоморфизм (т. е. группа $\mathrm{GL}_n(K)$ содержит подгруппу, изоморфную группе S_n).

Действительно, для $\sigma, \tau \in \mathrm{S}_n$ имеем

$$f(\sigma)f(\tau) = \left(\sum_{j=1}^{n} E_{\sigma(j)j}\right)\left(\sum_{j=1}^{n} E_{\tau(j)j}\right) =$$

$$= \sum_{j=1}^{n} E_{\sigma(i)i}E_{i=\tau(j)j} = \sum_{j=1}^{n} E_{\sigma(\tau(j))j} = \sum_{j=1}^{n} E_{(\sigma\tau)(j)j} = f(\sigma\tau),$$

т. е. f — гомоморфизм. Если $f(\sigma) = E$, то

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Итак, f — инъективный гомоморфизм. □

Контрольные вопросы 8.7.20.

1) $i \neq j$, $(e_{ij}^c)^{-1} = (E + cE_{ij})^{-1} = E - cE_{ij}$;

2) $i \neq j$, $t_{ij}^{-1} = t_{ij}$;

3) $\lambda_1 \neq 0, \ldots, \lambda_n \neq 0$, $d(\lambda_1, \ldots, \lambda_n)^{-1} = d(\lambda_1^{-1}, \ldots, \lambda_n^{-1})$.

8.8. Нахождение обратной матрицы A^{-1}

Пусть дана квадратная матрица $A \in \mathrm{M}_n(K)$ такая, что $|A| \neq 0$.

Первый способ. $A^{-1} = B = (b_{ij})$, $b_{ij} = \dfrac{A_{ji}}{|A|}$ (к сожалению, требуется вычислить n^2 определителей A_{ji} размера $(n-1) \times (n-1)$).

Второй способ. Найдём матрицу $X \in \mathrm{M}_n(K)$ такую, что $AX = E$ (тогда, по следствию 8.7.4, $XA = E$, $X = A^{-1}$). Это равносильно нахождению таких столбцов $\hat{X}_1, \ldots, \hat{X}_n$, что

$$A\hat{X}_1 = \hat{E}_1, \ldots, \ A\hat{X}_n = \hat{E}_n,$$

т. е. решению n систем линейных уравнений с матрицей A для коэффициентов и столбцами свободных членов $\hat{E}_1, \ldots, \hat{E}_n$ (столбцы единичной матрицы). Так как $|A| \neq 0$, то элементарными преобразованиями строк 1-го, 2-го и 3-го типов мы можем матрицу A привести к единичной матрице E. Применяя эти преобразования одновременно к n нашим системам, получаем

$$(A \mid E) \mapsto \ldots \mapsto (E \mid B).$$

Но тогда столбцы матрицы B — решения наших n систем, $AB = E$ (как мы уже отметили, в этом случае $BA = E$, $B = A^{-1}$).

Замечания 8.8.1.

1) Можно предложить другое обоснование этого алгоритма. Найдутся элементарные матрицы T_i 1-го, 2-го или 3-го типа такие, что $T_r \cdot \ldots \cdot T_2 T_1 A = E$, т. е. $TA = E$ для $T = T_r \cdot \ldots \cdot T_1$ и, следовательно, $T = A^{-1}$. Но тогда $B = TE = T = A^{-1}$.

   Отсюда следует также, что группа $G = \mathrm{GL}_n(K)$ порождается элементарными матрицами 1-го, 2-го и 3-го типа.

2) Этот алгоритм можно применять и для выяснения, существует ли обратная матрица, так как если определитель $|A|$ равен 0, то мы не сможем привести элементарными преобразованиями матрицу A к E (ступенчатый вид матрицы A будет треугольной матрицей с хотя бы одним нулём на диагонали). Это означает, что можно не вычислять определитель матрицы A перед применением алгоритма.

Пример 8.8.2.

$$A = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}, \quad |A| = 1 \neq 0,$$

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

т. е.

$$A^{-1} = \begin{pmatrix} 1 & -m \\ 0 & 1 \end{pmatrix}.$$

Пример 8.8.3. Найти обратную матрицу для матрицы

$$\begin{pmatrix} 5 & 2 & 0 \\ 2 & 1 & 1 \\ 3 & 3 & 8 \end{pmatrix},$$

если она существует.

Решение.

$$\left(\begin{array}{ccc|ccc} 5 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 3 & 8 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 5 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 7 & 0 & -1 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & 7 & 0 & -1 & 1 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 5 & 2 & 0 & 1 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 2 & 7 & 0 & -1 & 1 \\ 0 & -3 & -13 & 0 & 3 & -2 \\ 0 & -8 & -35 & 1 & 5 & -5 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & 7 & 0 & -1 & 1 \\ 0 & -3 & -13 & 0 & 3 & -2 \\ 0 & 1 & 4 & 1 & -4 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & 7 & 0 & -1 & 1 \\ 0 & 1 & 4 & 1 & -4 & 1 \\ 0 & -3 & -13 & 0 & 3 & -2 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & 7 & 0 & -1 & 1 \\ 0 & 1 & 4 & 1 & -4 & 1 \\ 0 & 0 & -1 & 3 & -9 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 2 & 7 & 0 & -1 & 1 \\ 0 & 1 & 4 & 1 & -4 & 1 \\ 0 & 0 & 1 & -3 & 9 & -1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 21 & -64 & 8 \\ 0 & 1 & 0 & 13 & -40 & 5 \\ 0 & 0 & 1 & -3 & 9 & -1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & 16 & -2 \\ 0 & 1 & 0 & 13 & -40 & 5 \\ 0 & 0 & 1 & -3 & 9 & -1 \end{array}\right).$$

Итак,

$$\begin{pmatrix} 5 & 2 & 0 \\ 2 & 1 & 1 \\ 3 & 3 & 8 \end{pmatrix}^{-1} = \begin{pmatrix} -5 & 16 & -2 \\ 13 & -40 & 5 \\ -3 & 9 & -1 \end{pmatrix}.$$

Замечания о матричных уравнениях $AX = B$ (случай $YA = B$ сводится к этому, $A^*Y^* = B^*$)

Случай 1. $A \in \mathrm{M}_n(K)$, $|A| \neq 0$. Тогда существует обратная матрица A^{-1}, и поэтому существует единственное решение $X = A^{-1}B$ уравнения $AX = B$ (для уравнения $YA = B$ существует единственное решение $Y = BA^{-1}$). При этом можно отдельно не вычислять матрицу A^{-1}, а применять наш алгоритм, приписывая к матрице A матрицу B, $(A \mid B)$, и приводя элементарными преобразованиями строк к $(E \mid A^{-1}B)$.

Пример 8.8.4. Пусть

$$A = \begin{pmatrix} 1 & 3 & 2 \\ -2 & -1 & 1 \\ 1 & 2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 8 & -16 & 9 \\ -6 & 7 & -3 \\ 7 & -13 & 7 \end{pmatrix}.$$

Требуется найти матрицу $A^{-1}B$

Решение.

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 8 & -16 & 9 \\ -2 & -1 & 1 & -6 & 7 & -3 \\ 1 & 2 & 2 & 7 & -13 & 7 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & -3 & 2 \\ -2 & -1 & 1 & -6 & 7 & -3 \\ 1 & 2 & 2 & 7 & -13 & 7 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & -3 & 2 \\ 0 & 3 & 5 & 8 & -19 & 11 \\ 1 & 2 & 2 & 7 & -13 & 7 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & -3 & 2 \\ 0 & 0 & 5 & 5 & -10 & 5 \\ 1 & 2 & 2 & 7 & -13 & 7 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 1 & -2 & 1 \\ 1 & 2 & 2 & 7 & -13 & 7 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 1 & -2 & 1 \\ 1 & 0 & 2 & 5 & -7 & 3 \end{array}\right) \to$$

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

Следовательно,

$$A^{-1}B = \begin{pmatrix} 3 & -3 & 1 \\ 1 & -3 & 2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Общий случай матричного уравнения $AX = B$, $A \in \mathrm{M}_{m,n}(K)$, $X \in \mathrm{M}_{n,r}(K)$, $B \in \mathrm{M}_{m,r}(K)$, равносилен рассмотрению r систем линейных уравнений с матрицей A и столбцами $\hat{B}_1, \ldots, \hat{B}_r$ в качестве столбцов свободных членов. Приведение матрицы A к ступенчатому виду $\bar{A}$,

$$(A \mid B) \mapsto (\bar{A} \mid \bar{B}),$$

сводит задачу к анализу r ступенчатых систем с одной матрицей $\bar{A}$ коэффициентов и столбцами свободных членов $\hat{\bar{B}}_1, \ldots, \hat{\bar{B}}_r$.

Замечание 8.8.5. Вычисление матрицы $Y = BA^{-1}$ можно провести, используя элементарные преобразования столбцов:

$$\left(\frac{A}{B}\right) \to \left(\frac{E}{BA^{-1}}\right).$$

8.9. Замечания об обратимом (биективном) линейном отображении

Замечание 8.9.1. Пусть U, V — линейные пространства, $f\colon U \to V$ — линейное отображение (т. е. $f(u_1 + u_2) = f(u_1) + f(u_2)$

и $f(ru) = rf(u)$ для всех $u, u_1, u_2 \in U$ и $r \in K$). Если отображение f биективно, то его обратное отображение f^{-1} также является линейным отображением.

Доказательство. Для всех $v_1, v_2 \in V$

$$f(f^{-1}(v_1 + v_2)) = v_1 + v_2 = \\ = f(f^{-1}(v_1)) + f(f^{-1}(v_2)) = f(f^{-1}(v_1) + f^{-1}(v_2)).$$

Так как f инъективно, то

$$f^{-1}(v_1 + v_2) = f^{-1}(v_1) + f^{-1}(v_2).$$

Аналогично, для $v \in V$ и $r \in K$ из

$$f(f^{-1}(rv)) = rv = rf(f^{-1}(v)) = f(rf^{-1}(v))$$

следует, что

$$f^{-1}(rv) = rf^{-1}(v).$$

Итак, f^{-1} — линейное отображение. □

Замечание 8.9.2. Если $f\colon \hat{K}^n \to \hat{K}^m$ — линейное отображение с матрицей $F = (f_{ij}) \in \mathrm{M}_{m,n}(K)$, то f — биективное отображение тогда и только тогда, когда

а) $m = n$,

б) $|F| \neq 0$.

При этом матрица линейного отображения $g = f^{-1}$ равна $G = F^{-1}$.

Доказательство.
1) Если $m = n$ и $|F| \neq 0$, то для системы

$$\begin{cases} f_{11}x_1 + \ldots + f_{1n}x_n = y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_{n1}x_1 + \ldots + f_{nn}x_n = y_n \end{cases}$$

по правилу Крамера знаем, что решение существует и единственно, при этом

$$x_i = \frac{D_i}{|F|} = \frac{F_{1i}}{|F|}y_1 + \ldots + \frac{F_{ni}}{|F|}y_n = g_{i1}y_1 + \ldots + g_{in}y_n,$$

где $G = (g_{ij}) = F^{-1}$. Итак, $g = f^{-1}$ существует и является линейным отображением с матрицей $G = F^{-1}$.

2) Для линейного отображения $f\colon \hat{K}^n \to \hat{K}^m$ с матрицей $F = (f_{ij}) \in \mathrm{M}_{m,n}(K)$, где

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \quad \begin{cases} f_{11}x_1 + \ldots + f_{1n}x_n = y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_{m1}x_1 + \ldots + f_{mn}x_n = y_m, \end{cases}$$

из нашего исследования систем линейных уравнений (метод Гаусса) имеем:

а) если $m < n$, то отображение f не является инъективным (даже для нулевого столбца свободных членов есть отличный от нуля прообраз (ненулевое решение));

б) если $m > n$, то отображение f не является сюръективным (так как если $m > n \geqslant r$, то в ступенчатой форме $\bar{F}$ нашей системы для столбца правых частей, дающего «экзотическое» уравнение $0x_1 + \ldots + 0x_n = \bar{y}_m \neq 0$, уже нет прообраза (решения)).

Итак, если f биективно, то $m = n$, т. е. $f\colon U \to U$, где $U = \hat{K}^n$. Если $g = f^{-1}$, то g — линейное отображение. Пусть $G = (g_{ij})$ — его матрица. Так как $fg = 1_U = gf$, то $FG = E = GF$, и поэтому $|F| \neq 0$ и $G = F^{-1}$. □

Упражнение 8.9.3 (ещё одна очень хорошая функция от матриц).

1) Пусть $A = (a_{ij}) \in \mathrm{M}_n(K)$. Положим

$$\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

(след матрицы A). Тогда:

а) tr — линейная функция,

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B), \quad \mathrm{tr}(\lambda A) = \lambda\,\mathrm{tr}(A)$$

для всех $A, B \in \mathrm{M}_n(K)$ и $\lambda \in K$;

б) $\mathrm{tr}(E) = n$;

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

2) Функция $\mathrm{tr}\colon \mathrm{M}_n(K) \to K$ однозначно определяется свойствами а), б) и в).

3) Если $\mathrm{char}\, K = 0$ (например, $K = \mathbb{R}$), то в алгебре матриц $\mathrm{M}_n(K)$ единичная матрица E не представима в виде $AB - BA$ для $A, B \in \mathrm{M}_n(K)$.

8.10. Матричное построение поля комплексных чисел

Поле комплексных чисел $\mathbb{C}$ можно найти как изоморфное подполе в кольце (2×2)-матриц $\mathrm{M}_2(\mathbb{R})$ над полем действительных чисел $\mathbb{R}$.

Рассмотрим совокупность $\mathbb{C}'$ всех (2×2)-матриц вида

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}),$$

где $a, b \in \mathbb{R}$. Так как

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} = \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

то подмножество $\mathbb{C}'$ в $\mathrm{M}_2(\mathbb{R})$ замкнуто относительно операций сложения и умножения, о которых мы уже знаем, что они ассоциативны, умножение в $\mathbb{C}'$ коммутативно, сложение и умножение связаны законом дистрибутивности.

Так как

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}', \quad -\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} -a & -b \\ -(-b) & aa \end{pmatrix} \in \mathbb{C}',$$

то $(\mathbb{C}', +)$ — абелева группа.

Итак, $(\mathbb{C}', +, \cdot)$ — коммутативное кольцо.

Если

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

то

$$\begin{vmatrix} a & b \\ -b & a \end{vmatrix} = a^2 + b^2 \neq 0,$$

и поэтому существует обратная матрица

$$A^{-1} = \begin{pmatrix} \dfrac{a}{a^2+b^2} & \dfrac{-b}{a^2+b^2} \\ -\left(\dfrac{-b}{a^2+b^2}\right) & \dfrac{a}{a^2+b^2} \end{pmatrix} \in \mathbb{C}',$$

таким образом, $\mathbb{C}'$ — поле (подполе в кольце матриц $\mathrm{M}_2(\mathbb{R})$).

Отождествляя действительное число $a \in \mathbb{R}$ со скалярной матрицей

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \in \mathbb{C}',$$

получаем (изоморфное) вложение поля $\mathbb{R}$ в $\mathbb{C}'$ ($\mathbb{R} \subseteq \mathbb{C}'$),

$$a \mapsto \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

Обозначив

$$i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{C}',$$

получаем

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

Если

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in \mathbb{C}',$$

то

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = a + bi.$$

Замечание 8.10.1. Фактически, нами установлено, что отображение f из $\mathbb{C}$ в $\mathbb{C}'$,

$$f((a,b)) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

является изоморфизмом построенных полей $\mathbb{C}$ и $\mathbb{C}'$, т. е. биекцией, для которой $f(z_1+z_2) = f(z_1)+f(z_2)$, $f(z_1z_2) = f(z_1)f(z_2)$ для всех $z_1, z_2 \in \mathbb{C}$.

Глава 9

Линейные пространства

9.1. Вывод свойств линейного пространства из аксиом

Пусть K — поле (например, $K = \mathbb{R}$ — поле действительных чисел). Многочисленные конкретные примеры линейных пространств, с которыми мы уже столкнулись (линейные пространства строк K^n, столбцов $\hat{K}^n$, пространства прямоугольных и квадратных матриц $\mathrm{M}_{m,n}(K)$ и $\mathrm{M}_n(K)$, пространство многочленов $K[x]$, пространство непрерывных вещественных функций $C[0,1]$ на отрезке $[0,1]$ и т. д.), оправдывают введение и рассмотрение понятия *абстрактного линейного пространства* ${}_K V$ над полем K как множества V с операцией сложения ($V \times V \to V$, $(a,b) \mapsto a+b$) и операциями умножения на элементы $c \in K$ ($V \to V$, $v \mapsto cv$), удовлетворяющими следующим условиям:

I.1) ассоциативность сложения (т. е. $(u+v)+w = u+(v+w)$ для всех $u, v, w \in V$);

I.2) коммутативность сложения (т. е. $u+v = v+u$ для всех $u, v \in V$);

I.3) существование нейтрального элемента 0 для операции сложения (т. е. $v+0 = v$ для всех $v \in V$);

I.4) существование противоположного элемента $-v$ для всякого $v \in V$ (т. е. $v+(-v) = 0$);

II.1) $1 \cdot v = v$ для всех $v \in V$;

II.2) $(rs)v = r(sv)$ для всех $r, s \in K$, $v \in V$;

III.1) $r(v_1 + v_2) = rv_1 + rv_2$ для всех $r \in K$, $v_1, v_2 \in V$;

III.2) $(r + s)v = rv + sv$ для всех $r, s \in K$, $v \in V$.

Приведём вывод ряда следствий из этих аксиом линейного пространства (хотя, конечно, в каждом конкретном случае они достаточно очевидны).

1) Уравнение $u+x = v$ для $u, v \in {}_KV$ имеет, причём единственное, решение $x = (-u) + v$.

Действительно, прибавляя $-u$ к левой и правой части, получаем, что $x = (-u) + v$. С другой стороны, $u + (-u) + v = v$.

2) Если $x + x = x$ для $x \in {}_KV$, то $x = 0$.

Действительно, прибавляя к левой и правой части противоположный элемент $-x$, получаем, что $x = (-x) + x + x = (-x) + x = 0$.

3) $0v = 0$ для любого $v \in {}_KV$.

Действительно, если $x = 0v$ (здесь $0 \in K$), то $x + x = 0v + 0v = (0 + 0)v = 0v = x$, и поэтому $x = 0 \in {}_KV$.

4) $r0 = 0$ для $r \in K$, $0 \in V$.

Действительно, если $x = r0$, то $x+x = r0+r0 = r(0+0) = r0 = x$, и поэтому $x = 0$.

5) $(-1)v = -v$ для всех $v \in V$.

Действительно, $(-1)v + v = (-1 + 1)v = 0v = 0$, т. е. $(-1)v = -v$.

6) $rv = 0$ для $r \in K$, $v \in V$ тогда и только тогда, когда либо $r = 0$, либо $v = 0$.

Действительно, если $r \neq 0$, то в поле K существует элемент $r^{-1} \in K$, и поэтому $v = 1v = r^{-1}rv = r^{-1}0 = 0$.

7) $r(u - v) = ru - rv$ для всех $r \in K$, $u, v \in V$.

Действительно, $r(u - v) + rv = r(u - v + v) = ru$, т. е. $r(u - v) = ru - rv$.

8) $-(-v) = v$ для всех $v \in V$.

Действительно, $v + (-v) = 0$, и поэтому $-(-v) = v$.

9.2. Линейная зависимость в линейных пространствах

Пусть ${}_KV$ — линейное пространство над полем K. Если $v_1, \ldots, v_r \in V$, $k_1, \ldots, k_r \in K$, то элемент

$$k_1v_1 + \ldots + k_rv_r \in V$$

называется *линейной комбинацией* элементов $v_1, \ldots, v_r$ с коэффициентами $k_1, \ldots, k_r \in K$.

Систему элементов $v_1, \ldots, v_r \in {}_KV$ назовём *линейно зависимой*, если найдутся элементы $k_1, \ldots, k_r \in K$ такие, что

а) не все k_i равны нулю (т. е. хотя бы один элемент k_i отличен от нуля);

б) $k_1v_1 + k_2v_2 + \ldots + k_rv_r = 0$.

Для краткости в этой ситуации мы будем говорить, что «*нетривиальная*» линейная комбинация элементов $v_1, \ldots, v_r$ равна нулю (конечно, *тривиальная* линейная комбинация всегда равна нулю, $0v_1 + \ldots + 0v_r = 0$).

Система элементов $v_1, \ldots, v_r \in {}_KV$ называется *линейно независимой*, если она не является линейно зависимой, это означает, что из равенства

$$k_1v_1 + \ldots + k_rv_r = 0, \quad k_1, \ldots, k_r \in K,$$

следует, что

$$k_1 = k_2 = \ldots = k_r = 0.$$

Теорема 9.2.1. *Система элементов $v_1, \ldots, v_r \in {}_KV$ линейно зависима тогда и только тогда, когда для некоторого i, $1 \leqslant i \leqslant r$,*

$$v_i = \sum_{j \neq i} l_jv_j, \quad l_j \in K$$

(т. е. элемент v_i является линейной комбинацией остальных элементов системы $v_1, \ldots, v_r$).

Доказательство.

1) Пусть система $v_1, \ldots, v_r$ линейно зависима, т. е.

$$k_1 v_1 + \ldots + k_r v_r = 0, \quad k_i \neq 0.$$

Тогда

$$v_i = \sum_{j \neq i} \frac{(-k_j)}{k_i} v_j.$$

2) Если

$$v_i = \sum_{j \neq i} l_j v_j,$$

то

$$\sum_{j \neq i} l_j v_j + (-1) v_i = v_i + (-1) v_i = 0,$$

т. е. система $v_1, \ldots, v_r$ линейно зависима, поскольку $-1 \neq 0$. □

Пример 9.2.2. Если в системе элементов $v_1, \ldots, v_r \in {}_K V$ есть нулевой элемент, скажем, $v_i = 0$, то система $v_1 \ldots, v_r$ линейно зависима.

Действительно, $0v_1 + \ldots + 1v_i + \ldots + 0v_r = 0$, или, другим способом, $v_i = 0 = \sum\limits_{j \neq i} 0 v_j$.

Пример 9.2.3. Если $v_i = v_j$ для $i \neq j$, то система $v_1, \ldots, v_r \in {}_K V$ линейно зависима.

Действительно, $0v_1 + \ldots + 1v_i + \ldots + (-1)v_j + \ldots + 0v_r = 0$, или, иначе, $v_i = v_j + \sum\limits_{\substack{k \neq i \\ k \neq j}} 0 v_k$.

Пример 9.2.4. Система строк $\varepsilon_1, \ldots, \varepsilon_n \in {}_K K^n$, где

$$\begin{aligned} \varepsilon_1 &= (1, 0, \ldots, 0), \\ \varepsilon_2 &= (0, 1, \ldots, 0), \\ &\ldots \\ \varepsilon_n &= (0, 0, \ldots, 1), \end{aligned}$$

линейно независима. Кроме того, любая строка $\alpha = (k_1, \ldots, k_n) \in {}_K K^n$ является линейной комбинацией элементов $\varepsilon_1, \ldots, \varepsilon_n$, а именно, $\alpha = (k_1, \ldots, k_n) = k_1 \varepsilon_1 + \ldots + k_n \varepsilon_n$.

Действительно,

$$k_1\varepsilon_1 + \ldots + k_n\varepsilon_n = (k_1, \ldots, k_n),$$

и поэтому если

$$k_1\varepsilon_1 + \ldots + k_n\varepsilon_n = (0, \ldots, 0),$$

то

$$k_1 = k_2 = \ldots = k_n = 0,$$

следовательно, система строк $\{\varepsilon_1, \ldots, \varepsilon_n\}$ линейно независима.

Пример 9.2.5. Пусть $v_1, v_2, v_3 \in {}_{\mathbb{R}}V$ — линейно независимая система в линейном пространстве ${}_{\mathbb{R}}V$. Тогда

$$u_1 = v_1 + v_2, \quad u_2 = v_1 + v_3, \quad u_3 = v_2 + v_3 -$$

также линейно независимая система.

Действительно, если

$$k_1u_1 + k_2u_2 + k_3u_3 = 0,$$

то

$$0 = k_1(v_1 + v_2) + k_2(v_1 + v_3) + k_3(v_2 + v_3) = \\ = (k_1 + k_2)v_1 + (k_1 + k_3)v_2 + (k_2 + k_3)v_3,$$

поэтому

$$\begin{cases} k_1 + k_2 = 0, \\ k_1 + k_3 = 0, \\ k_2 + k_3 = 0. \end{cases}$$

Следовательно, $k_1 = 0$, $k_2 = 0$, $k_3 = 0$, и система элементов u_1, u_2, u_3 линейно независима.

Упражнения 9.2.6.

1) Подсистема линейно независимой системы линейно независима.

2) Если подсистема линейно зависима, то линейно зависима и вся система.

Замечание 9.2.7. Для системы строк в K^n

$$\alpha_1 = (a_{11}, \ldots, a_{1n}),$$
$$\ldots$$
$$\alpha_r = (a_{r1}, \ldots, a_{rn})$$

вопрос о её линейной зависимости равносилен существованию ненулевого решения $(k_1, \ldots, k_r)$ следующей однородной системы линейных уравнений:

$$\begin{cases} a_{11}x_1 + \ldots + a_{r1}x_r = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}x_1 + \ldots + a_{rn}x_r = 0 \end{cases}$$

с транспонированной матрицей A^*, где

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{r1} & \ldots & a_{rn} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}.$$

Таким образом, метод Гаусса даёт нам в этом случае алгоритмическое решение задачи о линейной зависимости строк.

Теорема 9.2.8. *Пусть $A = (a_{ij}) \in \mathrm{M}_n(K)$ — квадратная матрица. Тогда следующие условия равносильны:*

1) $|A| = 0$;

2) *система строк $A_1, \ldots, A_n$ матрицы A линейно зависима (в пространстве строк K^n);*

3) *система столбцов $\hat{A}_1, \ldots, \hat{A}_n$ матрицы A линейно зависима (в пространстве столбцов $\hat{K}^n$).*

Доказательство.

1) Если строки матрицы A линейно зависимы, скажем, i-я строка A_i является линейной комбинацией остальных, $A_i = \sum\limits_{j \neq i} l_j A_j$, то, как мы показали, $|A| = 0$, т. е. 2) $\Longrightarrow$ 1).

2) Пусть $|A| = 0$. Тогда

$$k_1 A_1 + \ldots + k_n A_n = 0$$

в том и только в том случае, если $(k_1, \ldots, k_n)$ является решением однородной системы линейных уравнений с матрицей A^*. Так как $|A^*| = |A| = 0$, то существует ненулевое решение $(k_1, \ldots, k_n)$, т. е. система строк $A_1, \ldots, A_n$ матрицы A линейно зависима. Итак, 1) $\Longrightarrow$ 2).

3) Так как $|A^*| = |A|$, то 1) $\Longleftrightarrow$ 3). □

Задача 9.2.9. Пусть $A = (a_{ij}) \in \mathrm{M}_n(K)$, $B = (b_{ij}) \in \mathrm{M}_n(K)$, где $b_{ij} = A_{ji}$. Покажите, что если $|A| = 0$, то $|B| = 0$.

Теорема 9.2.10. *Любая система из m строк в K^n при $m > n$ линейно зависима.*

Доказательство. Если

$$\alpha_1 = (a_{11}, \ldots, a_{1n}),$$
$$\ldots$$
$$\alpha_m = (a_{m1}, \ldots, a_{mn}),$$

то равенство $k_1\alpha_1 + \ldots + k_m\alpha_m = 0$ равносильно тому, что $(k_1, \ldots, k_m)$ является решением следующей однородной системы линейных уравнений:

$$\begin{cases} a_{11}x_1 + \ldots + a_{m1}x_m = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}x_1 + \ldots + a_{mn}x_m = 0. \end{cases}$$

Так как число n уравнений меньше числа m переменных, то однородная система обладает ненулевым решением, т. е. система $\alpha_1, \ldots, \alpha_m$ линейно зависима. □

Следствие 9.2.11. *Если система $\alpha_1, \ldots, \alpha_r \in K^n$ линейно независима, то $r \leqslant n$.*

Лемма 9.2.12. *Если система элементов $\alpha_1, \ldots, \alpha_r \in {}_KV$ линейного пространства ${}_KV$ над полем K линейно независима, $\beta \in {}_KV$ и система $\alpha_1, \ldots, \alpha_r, \beta$ линейно зависима, то β является линейной комбинацией элементов $\alpha_1, \ldots, \alpha_r$.*

Доказательство. Пусть

$$k_1\alpha_1 + \ldots + k_r\alpha_r + k_{r+1}\beta = 0, \quad k_1, \ldots, k_{r+1} \in K,$$

где не все k_i, $1 \leqslant i \leqslant r+1$, равны нулю. Если бы $k_{r+1} = 0$, то нетривиальная линейная комбинация $k_1\alpha_1 + \ldots + k_r\alpha_r = 0$, равная нулю, означала бы, что система $\alpha_1, \ldots, \alpha_r$ линейно зависима, что противоречит предположению.

Итак, $k_{r+1} \neq 0$, и поэтому

$$\beta = \frac{-k_1}{k_{r+1}}\alpha_1 + \ldots + \frac{-k_r}{k_{r+1}}\alpha_r. \quad \square$$

Лемма 9.2.13 (единственность представления элемента линейного пространства ${}_KV$ в виде линейной комбинации линейно независимой системы элементов). *Пусть $\{\alpha_1, \ldots, \alpha_r\}$ — линейно независимая система элементов линейного пространства ${}_KV$ и*

$$\beta = k_1\alpha_1 + \ldots + k_r\alpha_r = k_1'\alpha_1 + \ldots + k_r'\alpha_r, \quad k_i, k_i' \in K.$$

Тогда $k_1 = k_1', \ldots, k_r = k_r'$.

Доказательство. Действительно,

$$(k_1 - k_1')\alpha_1 + \ldots + (k_r - k_r')\alpha_r = 0,$$

и поэтому $k_1 - k_1' = 0, \ldots, k_r - k_r' = 0$. $\square$

9.3. Максимальные линейно независимые подсистемы систем элементов линейных пространств, базис линейного пространства

Пусть $S \subseteq {}_KV$. Наиболее важные для нас случаи:

а) S — конечное подмножество элементов в ${}_KV$;

б) $S = {}_KV$.

Подсистема $v_1, \ldots, v_r \in S \subseteq {}_KV$ называется *максимальной линейно независимой подсистемой* в S, если:

1) $v_1, \ldots, v_r$ — линейно независимая система;

2) $v_1, \ldots, v_r, v$ — линейно зависимая система для всякого $v \in S$,

или, что эквивалентно,

2') любой элемент $v \in S$ является линейной комбинацией элементов $v_1, \ldots, v_r$.

Максимальная линейно независимая подсистема $v_1, \ldots, v_r$ в $S = {}_KV$ (если в ${}_KV$ существует такая *конечная* система) называется *базисом* линейного пространства ${}_KV$. Линейное пространство ${}_KV$ с конечным базисом $v_1, \ldots, v_r$ называется *конечномерным линейным пространством* (при этом будет показано, что любой другой базис линейного пространства содержит то же самое число элементов).

Пример 9.3.1. Как мы уже видели, система строк

$$\begin{aligned} \varepsilon_1 &= (1, 0, \ldots, 0), \\ \varepsilon_2 &= (0, 1, \ldots, 0), \\ &\ldots \\ \varepsilon_n &= (0, 0, \ldots, 1) \end{aligned}$$

является базисом линейного пространства строк K^n.

Лемма 9.3.2. *Любую линейно независимую подсистему $v_1, \ldots, v_r$ в $S \subseteq K^n$ можно дополнить до максимальной линейно независимой подсистемы в $S \subseteq K^n$.*

Доказательство. Если $v_1, \ldots, v_r$ — максимальная линейно независимая подсистема в $S \subseteq K^n$, то все доказано. Если нет, то найдётся элемент $v \in S$ такой, что $v_1, v_2, \ldots, v_r, v = v_{r+1}$ — линейно независимая подсистема в S. После конечного числа шагов процесс остановится, так как любые системы из $n+1$ элементов в линейном пространстве K^n оказываются линейно зависимыми. □

Следствие 9.3.3. *Любой ненулевой элемент $0 \neq v \in S \subseteq K^n$ дополняем до максимальной линейно независимой подсистемы в S.*

Следствие 9.3.4. *В $S = \mathbb{R}^n$ (или $S = K^n$ для бесконечного поля K) бесконечно много различных базисов. Если поле K конечно, $|K| = q$ (например, $K = \mathbb{Z}_2$), то число элементов в K^n равно q^n, и поэтому число базисов в K^n конечно. Найдите их число.*

Замечание 9.3.5. Пусть строки $a_1, \ldots, a_s \in K^n$ линейно независимы, $s < n$. Тогда существуют такие строки $a_{s+1}, \ldots, a_n \in K^n$,

что $\{a_1, \ldots, a_n\}$ — базис линейного пространства K^n. Практическое нахождение строк $a_{s+1}, \ldots, a_n$ можно осуществить следующим образом. Запишем строки $a_1, \ldots, a_s$ по столбцам и приведём полученную матрицу к ступенчатому виду: $\varphi(a_1^*, \ldots, a_s^*) = A_{\text{ступ}}$, где $(a_1^*, \ldots, a_s^*), A_{\text{ступ}} \in \mathrm{M}_{n,s}(K)$, φ — последовательность элементарных преобразований строк. Так как строки $a_1, \ldots, a_s$ линейно независимы, то в $A_{\text{ступ}}$ имеется ровно s ненулевых строк (первые s строк). Пусть $\hat{b}_{s+1}, \ldots, \hat{b}_n \in \hat{K}^n$ — столбцы, на i-м месте которых стоит 1, а остальные элементы равны 0, $i = s+1, \ldots, n$. Припишем эти столбцы справа к матрице $A_{\text{ступ}}$. Пусть $B \in \mathrm{M}_n(K)$ — полученная матрица. Применяя к матрице B последовательность элементарных преобразований строк, обратную к φ, приходим к матрице $\tilde{B}$. При этом $(\tilde{B})^*$ — матрица, в которой первые s строк — это $a_1, \ldots, a_s$, а последующие строки дополняют их до базиса линейного пространства K^n.

9.4. Замечание о линейной выражаемости конечных систем элементов в линейном пространстве

Пусть ${}_K V$ — линейное пространство, $S_1 \subseteq {}_K V$, $S_2 \subseteq {}_K V$. Будем говорить, что система S_2 элементов $u_1, \ldots, u_s$ *линейно выражается* через систему S_1 элементов $v_1, \ldots, v_r$, если каждый элемент $u_i \in S_2$, $1 \leqslant i \leqslant s$, является линейной комбинацией элементов $v_1, \ldots, v_r$ системы S_1,

$$u_i = \sum_{j=1}^{r} m_{ij} v_j, \quad m_{ij} \in K.$$

Если к тому же система S_3 элементов $w_1, \ldots, w_t$ линейно выражается через систему S_2,

$$w_k = \sum_{i=1}^{s} l_{ki} u_i, \quad l_{ki} \in K, \quad 1 \leqslant k \leqslant t,$$

то

$$w_k = \sum_{i=1}^{s} l_{ki} u_i = \sum_{i=1}^{s} \sum_{j=1}^{r} (l_{ki} m_{ij}) v_j = \sum_{j=1}^{r} \left(\sum_{i=1}^{s} l_{ki} m_{ij} \right) v_j,$$

т. е. система S_3 линейно выражается через систему S_1.

Системы S_1 и S_2 называются *эквивалентными*, если они линейно выражаются друг через друга (обозначение: $S_1 \sim S_2$).

Следствие 9.4.1. *Отношение «быть эквивалентными системами», $S_1 \sim S_2$, является отношением эквивалентности.*

Следствие 9.4.2. *Если элемент $v \in {}_KV$ является линейной комбинацией элементов $v_1, \ldots, v_r$ системы S_1, $S_1 \sim S_2$, где S_2 — система элементов $u_1, \ldots, u_s$, то элемент v является линейной комбинацией элементов $u_1, \ldots, u_s$ системы S_2.*

Следствие 9.4.3. *Любая (конечная) система элементов $S \subseteq {}_KV$ эквивалентна своей максимальной линейно независимой подсистеме.*

Следствие 9.4.4. *Любые две (конечные) максимально независимые подсистемы любой системы $S \subseteq {}_KV$ эквивалентны.*

Замечание 9.4.5. Если $A, B \in \mathrm{M}_{m,n}(K)$ и матрица B получена из матрицы A конечным числом элементарных преобразований 1-го, 2-го и 3-го типов, то каждая строка матрицы B является линейной комбинацией строк матрицы A (поскольку от матрицы B мы можем вернуться к матрице A с помощью элементарных преобразований строк 1-го, 2-го и 3-го типов, то каждая строка матрицы A является линейной комбинацией строк матрицы B). Таким образом, в линейном пространстве строк K^n системы строк $A_1, \ldots, A_m$ матрицы A и $B_1, \ldots, B_m$ матрицы B линейно выражаются друг через друга.

Теорема 9.4.6 (основная теорема о линейной зависимости). *Пусть в линейном пространстве ${}_KV$ линейно независимая система элементов $v_1, \ldots, v_r$ линейно выражается через другую систему элементов $u_1, \ldots, u_s$. Тогда $r \leqslant s$.*

Доказательство. Допустим противное: пусть $r > s$. В силу нашего предположения

$$v_1 = a_{11}u_1 + \ldots + a_{1s}u_s,$$

$$\ldots$$

$$v_r = a_{r1}u_1 + \ldots + a_{rs}u_s, \quad a_{ij} \in K.$$

Так как $r > s$, то r строк

$$(a_{11}, \ldots, a_{1s}),$$
$$\ldots$$
$$(a_{r1}, \ldots, a_{rs})$$

в линейном пространстве строк K^s линейно зависимы: найдётся их линейная комбинация с коэффициентами $k_1, \ldots, k_r$, где $k_i \neq 0$ для некоторого i, равная нулевой строке $(0, \ldots, 0) \in K^s$. Но тогда и линейная комбинация элементов $v_1, \ldots, v_r$ с этими же коэффициентами $k_1, \ldots, k_r$, равна нулю, $k_1 v_1 + \ldots + k_r v_r = 0$. Таким образом, система элементов $v_1, \ldots, v_r$ линейно зависима, что приводит нас к противоречию. □

Следствие 9.4.7. *Две эквивалентные конечные линейно независимые системы в линейном пространстве* ${}_K V$ *содержат равное число элементов.*

Следствие 9.4.8. *Для системы* $S \subseteq {}_K V$, *где* ${}_K V$ *— конечномерное линейное пространство, любые две (конечные) максимальные линейно независимые подсистемы содержат одинаковое число элементов* $\mathrm{r}(S)$, *называемое* рангом *системы* S.

Следствие 9.4.9. *Если* $S = {}_K V$ *и* ${}_K V$ *— конечномерное линейное пространство, то любые два базиса в* ${}_K V$ *состоят из одного и того же числа элементов* n, *это число* n *называется размерностью линейного пространства* ${}_K V$, *обозначение*: $\dim {}_K V = n$.

Как мы видели ранее, одним из базисов в линейном пространстве строк ${}_K K^n$ является система строк

$$\varepsilon_1 = (1, 0, \ldots, 0),$$
$$\ldots$$
$$\varepsilon_n = (0, 0, \ldots, 1),$$

и поэтому $\dim {}_K K^n = n$.

Следствие 9.4.10. *Если в конечномерном линейном пространстве* ${}_K V$ *одна система элементов* S_1 *линейно выражается через другую систему* S_2, *то* $\mathrm{r}(S_1) \leqslant \mathrm{r}(S_2)$.

Следствие 9.4.11. *Если в линейном пространстве ${}_KV$ система M из m элементов имеет ранг r, то любая её подсистема S из s элементов ($s \leqslant m$) имеет ранг не меньше чем $r + s - m$.*

Доказательство. Действительно, если R — максимальная линейно независимая подсистема в M, $|R| = r$, то $R \setminus (R \cap S) \subset M \setminus S$, и поэтому $|R \setminus (R \cap S)| \leqslant m - s$. Следовательно, $|R \cap S| \geqslant r - (m - s) = r + s - m$. □

Следствие 9.4.12. *Для системы строк $v_1, \ldots, v_r \in K^n$ следующие условия эквивалентны:*

1) *система строк $v_1, \ldots, v_r$ является базисом линейного пространства строк K^n (т. е. максимальной линейно независимой подсистемой строк в K^n; и тогда $r = n$);*

2) *каждая строка $v \in K^n$ единственным образом представляется в виде линейной комбинации*

$$v = \lambda_1 v_1 + \ldots + \lambda_r v_r, \quad \lambda_1, \ldots, \lambda_r \in K$$

(и тогда $r = n$);

3) *$r = n$ и система строк $v_1, \ldots, v_n$ линейно независима;*

4) *$r = n$ и каждая строка $v \in K^n$ представима в виде линейной комбинации*

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n, \quad \lambda_1, \ldots, \lambda_n \in K.$$

Доказательство. Мы уже показали, что 1) $\Longrightarrow$ 2). Покажем, что 2) $\Longrightarrow$ 1). Если $v_1, \ldots, v_r$ — линейно зависимая система строк, $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$ с некоторым $\lambda_i \neq 0$, то нулевая строка имеет два различных представления

$$0 = 0 \cdot v_1 + \ldots + 0 \cdot v_r = \lambda_1 v_1 + \ldots + \lambda_r v_r, \quad \lambda_i \neq 0.$$

При этом $r = n$, так как любые базисы в K^n содержат n элементов.

Ясно, что 1) $\Longrightarrow$ 3). Покажем, что 3) $\Longrightarrow$ 1). Для любой строки $v \in K^n$ система строк $v_1, \ldots, v_n, v$ линейно зависима ($n + 1 > n$). Так

как $v_1, \ldots, v_n$ — линейно независимая система, то $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ для некоторых $\lambda_1, \ldots, \lambda_n \in K$.

Ясно, что 1) $\Longrightarrow$ 4). Покажем, что 4) $\Longrightarrow$ 1). Допустим, что $v_1, \ldots, v_n$ — линейно зависимая система. Тогда её максимально линейно независимая подсистема $v_{i_1}, \ldots, v_{i_r}$, $r < n$, является максимальной линейно независимой подсистемой в K^n, что противоречит $r = n$. □

9.5. Единственность главного ступенчатого вида матрицы

Теорема 9.5.1. *Пусть $A, B, C \in \mathrm{M}_{m,n}(K)$, B и C — ступенчатые матрицы, полученные из ненулевой матрицы A конечным числом элементарных преобразований строк 1-го, 2-го и 3-го типов. Тогда:*

1) *системы строк $\{B_1, \ldots, B_m\}$ матрицы B и $\{C_1, \ldots, C_m\}$ матрицы C в линейном пространстве строк K^n линейно выражаются друг через друга (другими словами, линейные оболочки строк матриц A, B и C в K^n совпадают: $\langle A_1, \ldots, A_m\rangle = \langle B_1, \ldots, B_m\rangle = \langle C_1, \ldots, C_m\rangle$, см. с. 107);*

2) *числа r_1 и r_2 ненулевых строк в ступенчатых матрицах B и C соответственно совпадают (при этом $r = r_1 = r_2 = \dim_K\langle A_1, \ldots, A_m\rangle$; другие интерпретации числа $r = r(A)$ будут даны в теореме 9.16.1 о ранге матрицы);*

3) *лидеры строк ступенчатых матриц B и C располагаются в одних и тех же столбцах;*

4) *если B и C — главные ступенчатые виды ненулевой матрицы $A \in \mathrm{M}_{m,n}(K)$, то $B = C$.*

Доказательство.

1) В силу замечания 9.4.5, в линейном пространстве строк K^n системы строк $\{A_1, \ldots, A_m\}$ матрицы A и $\{B_1, \ldots, B_m\}$ матрицы B линейно выражаются друг через друга. Аналогично, системы строк $\{A_1, \ldots, A_m\}$ матрицы A и $\{C_1, \ldots, C_m\}$ матрицы C также линейно выражаются друг через друга. Принимая во внимание транзитивность линейной выражаемости систем строк (см. следствие 9.4.2), по-

лучаем, что системы строк $\{B_1, \ldots, B_m\}$ матрицы B и $\{C_1, \ldots, C_m\}$ матрицы C линейно выражаются друг через друга. Следовательно,

$$\langle A_1, \ldots, A_m \rangle = \langle B_1, \ldots, B_m \rangle = \langle C_1, \ldots, C_m \rangle.$$

2) Так как ненулевые строки ступенчатой матрицы образуют максимально независимую подсистему строк, то из 1) следует, что $r_1 = r_2$ (см. следствие 9.4.10), при этом

$$r = r_1 = r_2 = \dim\langle B_1, \ldots, B_m \rangle = \\ = \dim\langle C_1, \ldots, C_m \rangle = \dim\langle A_1, \ldots, A_m \rangle.$$

3) Пусть лидеры r ненулевых строк $B_1, B_2, \ldots, B_r$ ступенчатой матрицы B расположены в столбцах с номерами $k_1, k_2, \ldots, k_r$, $k_1 < k_2 < \ldots < k_r$, а лидеры r ненулевых строк $C_1, C_2, \ldots, C_r$ ступенчатой матрицы C расположены в столбцах с номерами $l_1, l_2, \ldots, l_r$, $l_1 < l_2 < \ldots < l_r$. Так как системы строк $\{B_1, B_2, \ldots, B_r\}$, $\{C_1, C_2, \ldots, C_r\}$ линейно выражаются друг через друга, то, в силу леммы 3.5.5 и следствия 3.5.6, $k_1 = l_1$ ($k_1 \geqslant \min\{l_i\} = l_1$; $l_1 \geqslant \min\{k_i\} = k_1$).

Если

$$B_2 = \sum_{j=1}^{r} \lambda_{2j} C_j, \quad C_2 = \sum_{j=1}^{r} \mu_{2j} B_j,$$

то $\lambda_{21} = 0 = \mu_{21}$. Применяя наше рассуждение для систем $\{B_2, \ldots, B_r\}$ и $\{C_2, \ldots, C_r\}$, которые линейно выражаются друг через друга, получаем, что $k_2 = l_2$.

Продолжая этот процесс, убеждаемся в том, что $k_3 = l_3, \ldots, k_r = l_r$.

4) В 2) и 3) доказано, что число ненулевых строк r и номера столбцов $l_1, \ldots, l_r$, $1 \leqslant l_1 < l_2 < \ldots < l_r \leqslant n$, в которых находятся главные неизвестные главных ступенчатых видов B и C, определены однозначно. Таким образом, разбиения на главные и свободные неизвестные, определяемые ступенчатыми видами B и C, совпадают. Поскольку главные неизвестные однозначно выражаются через свободные (в эквивалентных однородных системах линейных уравнений с главными ступенчатыми матрицами B и C), при этом главный ступенчатый вид определяется этим выражением однозначно (см. замечание 3.6.9), то $B = C$. □

Замечание 9.5.2 (матричное доказательство п. 4 теоремы о единственности главного ступенчатого вида). Для $A \in \mathrm{M}_{m,n}(K)$ существуют такие обратимые матрицы $F, G \in \mathrm{M}_m(K)$ (произведения матриц, соответствующих элементарным преобразованиям строк), что

$$A = F \cdot B = G \cdot C.$$

Следовательно,

$$B = D \cdot C, \quad \text{где} \quad D = F^{-1}G.$$

Используя определение главного ступенчатого вида и переставляя столбцы матриц B и C, имеем:

$$B \cdot Q = \left(\begin{array}{c|c} E_r & * \\ \hline 0 & 0 \end{array}\right) = D \cdot \left(\begin{array}{c|c} E_r & *' \\ \hline 0 & 0 \end{array}\right) = D \cdot C \cdot Q, \qquad (9.1)$$

где $Q \in \mathrm{M}_n(K)$ (матрица Q — обратимая матрица, соответствующая последовательности элементарных преобразований столбцов; мы уже доказали в п. 2 и 3, что числа r и столбцы $j_1, \ldots, j_r$, в которых стоят лидеры строк, одинаковы для ступенчатых матриц B и C, соответственно; нулевые блоки могут отсутствовать (если $k = r = m$)). Следовательно, матрица D имеет следующий блочный вид:

$$D = \left(\begin{array}{c|c} E_r & \\ \hline 0 & \tilde{*} \end{array}\right),$$

где матрица $\tilde{*} \in \mathrm{M}_{m,m-r}(K)$ (если $r < m$) состоит из произвольных элементов поля K. Поэтому, умножая D на

$$\left(\begin{array}{c|c} E_r & *' \\ \hline 0 & 0 \end{array}\right)$$

и приравнивая к

$$\left(\begin{array}{c|c} E_r & * \\ \hline 0 & 0 \end{array}\right),$$

получаем, что $* = *' \in \mathrm{M}_{m-r,n-r}(K)$. Умножая (9.1) справа на Q^{-1}, получаем $B = C$.

9.6. Изоморфизм линейных пространств

Пусть ${}_KU$, ${}_KV$ — линейные пространства над полем K. Биективное отображение

$$f\colon {}_KU \to {}_KV,$$

для которого

$$f(u_1 + u_2) = f(u_1) + f(u_2),$$
$$f(ku) = kf(u)$$

для всех $u_1, u_2, u \in {}_KU$, $k \in K$, называется *изоморфизмом* линейных пространств ${}_KU$ и ${}_KV$ (в этом случае будем говорить, что линейные пространства ${}_KU$ и ${}_KV$ *изоморфны*, обозначение: ${}_KU \cong {}_KV$).

Упражнение 9.6.1. Отношение ${}_KU \cong {}_KV$ является отношением эквивалентности.

Лемма 9.6.2. *Если* $f\colon {}_KU \to {}_KV$ *— изоморфизм линейных пространств,* $\dim {}_KU = n$, $\{e_1, \ldots, e_n\}$ *— базис в* ${}_KU$, *то* $\{f(e_1), \ldots, f(e_n)\}$ *— базис в* ${}_KV$, *и поэтому* $\dim {}_KV = n = \dim {}_KU$.

Доказательство.

1) Если $v \in {}_KV$, то $f(u) = v$ для некоторого $u \in {}_KU$. Пусть $u = k_1e_1 + \ldots + k_ne_n$, где $k_1, \ldots, k_n \in K$. Тогда

$$v = f(u) = k_1f(e_1) + \ldots + k_nf(e_n).$$

2) Пусть $k_1f(e_1) + \ldots + k_nf(e_n) = 0$ для $k_1, \ldots, k_n \in K$. Тогда

$$0 = k_1f(e_1) + \ldots + k_nf(e_n) = f(k_1e_1 + \ldots + k_ne_n),$$

и поэтому

$$k_1e_1 + \ldots + k_ne_n = 0,$$

следовательно, $k_1 = k_2 = \ldots = k_n = 0$.

Итак, в силу 1) и 2), $\{f(e_1), \ldots, f(e_n)\}$ — базис линейного пространства ${}_KV$. □

Лемма 9.6.3. *Если* $\dim {}_K V = n$ *и* $\{e_1, \ldots, e_n\}$ *— базис линейного пространства* ${}_K V$, *то, сопоставляя каждому элементу* $v = k_1 e_1 + \ldots + k_n e_n \in {}_K V$ *однозначно определённую строчку его координат* $(k_1, \ldots, k_n)$ *в базисе* $\{e_1, \ldots, e_n\}$, *получаем изоморфизм линейных пространств* ${}_K V \cong K^n$, *таким образом, каждое* n*-мерное линейное пространство* ${}_K V$ *над полем* K *изоморфно линейному пространству строк* K^n.

Доказательство. Соответствие

$$\Delta\colon {}_K V \ni v = k_1 e_1 + \ldots + k_n e_n \mapsto (k_1, \ldots, k_n) \in K^n$$

является биекцией, для которой

$$\begin{aligned}
\Delta(v + v') &= \Delta((k_1 e_1 + \ldots + k_n e_n) + (k'_1 e_1 + \ldots + k'_n e_n)) = \\
&= \Delta((k_1 + k'_1) e_1 + \ldots + (k_n + k'_n) e_n) = \\
&= (k_1 + k'_1, \ldots, k_n + k'_n) = (k_1, \ldots, k_n) + (k'_1, \ldots, k'_n) = \\
&= \Delta(v) + \Delta(v'); \\
\Delta(kv) &= \Delta(k(k_1 e_1 + \ldots + k_n e_n)) = \Delta((kk_1) e_1 + \ldots + (kk_n) e_n) = \\
&= (kk_1, \ldots, kk_n) = k(k_1, \ldots, k_n) = k\Delta(v). \quad \square
\end{aligned}$$

Теорема 9.6.4. *Конечномерные линейные пространства* ${}_K U$ *и* ${}_K V$ *изоморфны тогда и только тогда, когда* $\dim {}_K U = \dim {}_K V = n$, *и в этом случае* ${}_K U \cong K^n \cong {}_K V$.

Доказательство теоремы следует из лемм 9.6.2 и 9.6.3. $\square$

Упражнение 9.6.5. Покажите, что следующие линейные пространства являются бесконечномерными линейными пространствами (это означает, что в них нет базиса из конечного числа элементов):

1) ${}_{\mathbb{R}} C[0,1]$ — линейное пространство вещественных непрерывных функций на отрезке $[0,1]$;

2) ${}_K K[x]$ — линейное пространство многочленов от переменной x с коэффициентами из поля K;

3) ${}_K K^{\mathbb{N}}$ — линейное пространство всех счётных последовательностей $(k_1, k_2, \ldots, k_n, \ldots)$ элементов из поля K.

Упражнение 9.6.6. Докажите, что

а) $\dim {}_K\mathrm{M}_{m,n}(K) = mn$;

б) $\dim {}_{\mathbb{R}}\{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = A\} = \dfrac{n(n+1)}{2}$;

в) $\dim {}_{\mathbb{R}}\{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = -A\} = \dfrac{n(n-1)}{2}$.

9.7. Замена базиса линейного пространства

Пусть V — конечномерное линейное пространство над полем K, $\dim V = n < \infty$, $\{v_1, \ldots, v_n\}$ — базис в V, $\{v'_1, \ldots, v'_n\}$ — другой базис в V,

$$v'_j = c_{1j}v_1 + c_{2j}v_2 + \ldots + c_{nj}v_n, \quad j = 1, \ldots, n, \quad c_{ij} \in K$$

(запись по столбцу!). $C = (c_{ij}) \in \mathrm{M}_n(K)$ — *матрица перехода* от первого базиса ко второму.

Замечание 9.7.1. Так как умножение в поле K коммутативно, то левое линейное пространство ${}_KV$ можно рассматривать и как правое линейное пространство V_K, полагая $v\lambda = \lambda v$ для всех $\lambda \in K$, $v \in V$. Тогда определение матрицы перехода может быть записано в матричном виде как

$$(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n)C.$$

Ограничиваясь левыми линейными пространствами, мы можем использовать эквивалентную форму записи:

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

или, кратко, $\mathcal{E}' = C^*\mathcal{E}$, где

$$\mathcal{E} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \mathcal{E}' = \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} \in \mathrm{M}_{n,1}(V).$$

Если $V = K^n$, то $v_1, \ldots, v_n, v'_1, \ldots, v'_n \in K^n$, $\mathcal{E}, \mathcal{E}' \in \mathrm{M}_n(K)$ и $\mathcal{E}' = C^*\mathcal{E}$ означает равенство квадратных $(n \times n)$-матриц.

9.8. Обратимость матрицы перехода

1) Если $|C| = 0$, то $|C^*| = 0$ и строки матрицы C^* линейно зависимы. Поэтому из

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \text{т. е.} \quad \mathcal{E}' = C^*\mathcal{E},$$

следует, что $v'_1, \ldots, v'_n$ — линейно зависимая система в V, что приводит к противоречию с тем, что $v'_1, \ldots, v'_n$ — базис. Итак, мы показали, что $|C| \neq 0$ и существует обратная матрица C^{-1} (тогда $(C^*)^{-1} = (C^{-1})^*$).

2) Другое доказательство обратимости матрицы C даёт интерпретация матрицы $B = C^{-1}$ как матрицы перехода от второго базиса к первому.

Действительно, элементы $v_1, \ldots, v_n$ также выражаются как линейные комбинации элементов базиса $\{v'_1, \ldots, v'_n\}$:

$$v_i = b_{1i}v'_1 + \ldots + b_{ni}v'_n, \quad i = 1, \ldots, n, \quad b_{ij} \in K,$$

$B = (b_{ij}) \in \mathrm{M}_n(K)$. Тогда $\mathcal{E} = B^*\mathcal{E}'$. Так как $\mathcal{E}' = C^*\mathcal{E}$, то

$$\mathcal{E} = B^*(C^*\mathcal{E}) = (B^*C^*)\mathcal{E} = (CB)^*\mathcal{E}.$$

Так как $\{v_1, \ldots, v_n\}$ — базис в V, то $(CB)^* = E$, следовательно, $CB = E$, и поэтому $B = C^{-1}$. □

3) Для любой обратимой матрицы $C \in \mathrm{M}_n(K)$, $|C| \neq 0$, и любого базиса $\{v_1, \ldots, v_n\}$ конечномерного линейного пространства ${}_KV$, $\dim {}_KV = n$, элементы $v'_1, \ldots, v'_n \in {}_KV$, где

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

образуют базис линейного пространства ${}_KV$.

Действительно, в этом случае

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (C^*)^{-1} \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix}, \quad (C^*)^{-1} = (C^{-1})^*,$$

т. е. n линейно независимых элементов $v_1, \ldots, v_n$ линейно выражаются через $v'_1, \ldots, v'_n$. По основной лемме о линейной зависимости элементы $v'_1, \ldots, v'_n$ линейно независимы. Так как $\dim {}_K V = n$, то $\{v'_1, \ldots, v'_n\}$ — базис линейного пространства ${}_K V$. □

9.9. Замена координат элемента линейного пространства при замене базиса

Пусть $\{v_1, \ldots, v_n\}$, $\{v'_1, \ldots, v'_n\}$ — два базиса линейного пространства ${}_K V$, $\dim {}_K V = n$, $C \in \mathrm{M}_n(K)$, $|C| \neq 0$, — матрица перехода от первого базиса ко второму,

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

$x = x_1 v_1 + \ldots + x_n v_n = x'_1 v'_1 + \ldots + x'_n v'_n \in {}_K V$. Так как

$$x = x_1 v_1 + \ldots + x_n v_n = (x_1, \ldots, x_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} =$$

$$= (x_1, \ldots, x_n)(C^{-1})^* \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = (x'_1, \ldots, x'_n) \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix},$$

то

$$(x'_1, \ldots, x'_n) = (x_1, \ldots, x_n)(C^{-1})^*,$$

или

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = C^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

что эквивалентно

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}. \quad \square$$

Пример 9.9.1. Пусть $V = \mathbb{R}^3$, $v_1 = (2, 1, -3)$, $v_2 = (3, 2, -5)$, $v_3 = (1, -1, 1)$. Необходимо выяснить, образуют ли элементы v_1, v_2, v_3 базис в $\mathbb{R}^3$, и если да, то найти координаты строки $x = (6, 2, -7)$ в базисе $\{v_1, v_2, v_3\}$.

Решение.

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = C^* \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix},$$

где $\{e_1, e_2, e_3\}$ — стандартный базис в $\mathbb{R}^3$,

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

Строки v_1, v_2, v_3 образуют базис в $\mathbb{R}^3$ тогда и только тогда, когда матрица C обратима. Если матрица C обратима, то столбец координат строки x в базисе $\{v_1, v_2, v_3\}$ равен

$$C^{-1} \begin{pmatrix} 6 \\ 2 \\ -7 \end{pmatrix}.$$

Для вычисления этого столбца применим алгоритм вычисления матрицы $A^{-1}B$ (см. с. 182), в процессе работы которого проверяется, обратима ли матрица $A = C$:

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 1 & 2 & -1 & 2 \\ -3 & -5 & 1 & -7 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 3 & 1 & 6 \\ -3 & -5 & 1 & -7 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -1 & 3 & 2 \\ -3 & -5 & 1 & -7 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 1 & -2 & -1 \end{array}\right) \to$$

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

$$\to \left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Таким образом, матрица C обратима, $(1,1,1)$ — координаты строки x в базисе $\{v_1, v_2, v_3\}$, $x = v_1 + v_2 + v_3$.

Этот же результат можно было получить, используя формулу $(6,2,-7)(C^*)^{-1} = (1,1,1)$,

$$\left(\begin{array}{ccc} 2 & 1 & -3 \\ 3 & 2 & -5 \\ 1 & -1 & 1 \\ \hline 6 & 2 & -7 \end{array}\right) \to \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 1 & 1 \end{array}\right)$$

(здесь применяем элементарные преобразования столбцов).

9.10. Линейные подпространства линейных пространств

Пусть K — поле, ${}_KV$ — линейное пространство над полем K. Непустое подмножество $\varnothing \neq U \subseteq {}_KV$ называется *линейным подпространством* линейного пространства ${}_KV$, если:

1) $u_1 + u_2 \in U$ для всех $u_1, u_2 \in U$;

2) $ku \in U$ для всех $k \in K$, $u \in U$.

Ясно, что ${}_KU$ — линейное пространство относительно тех же операций сложения элементов и умножения на элементы из поля K, что и в линейном пространстве ${}_KV$.

Если U — линейное подпространство в конечномерном линейном пространстве ${}_KV$, $n = \dim {}_KV < \infty$, то $\dim {}_KU \leqslant \dim {}_KV$. Действительно, если элементы $u_1, \ldots, u_s \in {}_KU$ линейно независимы в ${}_KU$, то эти элементы линейно независимы и в линейном пространстве ${}_KV$, $s \leqslant n$, поэтому $\dim {}_KU \leqslant \dim {}_KV$.

Если ${}_KU$ — линейное подпространство линейного пространства ${}_KV$, ${}_KU \subseteq {}_KV$ и $\dim {}_KU = \dim {}_KV = n$, то ${}_KU = {}_KV$. Действительно, если $\{u_1, \ldots, u_n\}$ — базис линейного пространства ${}_KU \subseteq {}_KV$, то эти n элементов $u_1, \ldots, u_n$ линейно независимы в ${}_KV$ и $\dim {}_KV = n$, поэтому $\{u_1, \ldots, u_n\}$ — базис линейного пространства ${}_KV$. Итак, каждый элемент $v \in V$ имеет вид $v = k_1u_1 + \ldots + k_nu_n \in {}_KU$, $k_i \in K$, т. е. ${}_KV = {}_KU$.

9.11. Пересечение линейных подпространств

Лемма 9.11.1. *Пересечение*

$$U = \bigcap_{i \in I} U_i$$

любого семейства линейных подпространств $\{U_i \subset {}_K V \mid i \in I\}$ *линейного пространства* ${}_K V$ *является линейным подпространством.*

Доказательство. Если $u, u_1, u_2 \in U = \bigcap\limits_{i \in I} U_i$, $k \in K$, то $u, u_1, u_2 \in U_i$ для любого $i \in I$, поэтому $u_1 + u_2, ku \in U_i$ для любого $i \in I$, т. е. $u_1 + u_2, ku \in U = \bigcap\limits_{i \in I} U_i$. □

Следствие 9.11.2. *Если* U_1 *и* U_2 *— линейные подпространства линейного пространства* ${}_K V$, *то* $U_1 \cap U_2$ *— линейное подпространство в* ${}_K V$ *(наибольшее подпространство среди подпространств, лежащих одновременно в* U_1 *и в* U_2*).*

9.12. Сумма линейных подпространств

Если U_1 и U_2 — линейные подпространства линейного пространства ${}_K V$, то *сумма линейных подпространств*

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\ u_2 \in U_2\}$$

также является линейным подпространством. Действительно, если $u_1 + u_2, u_1' + u_2' \in U_1 + U_2$, $u_1, u_1' \in U_1$, $u_2, u_2' \in U_2$, $k \in K$, то

$$(u_1 + u_2) + (u_1' + u_2') = (u_1 + u_1') + (u_2 + u_2') \in U_1 + U_2;$$
$$k(u_1 + u_2) = ku_1 + ku_2 \in U_1 + U_2. \quad \square$$

Замечание 9.12.1. $U_1 + U_2$ — наименьшее линейное подпространство среди линейных подпространств, содержащих одновременно U_1 и U_2. Более того,

$$U_1 + U_2 = \bigcap_{\substack{U \subseteq {}_K V \\ U_1 \subseteq U,\ U_2 \subseteq U}} U.$$

Замечание 9.12.2. Если U, U_1, U_2, U_3 — линейные подпространства в ${}_KV$, то

$$U \cap U = U, \quad U + U = U,$$
$$U_1 \cap U_2 = U_2 \cap U_1, \quad U_1 + U_2 = U_2 + U_1,$$
$$U_1 \cap (U_2 \cap U_3) = (U_1 \cap U_2) \cap U_3,$$
$$U_1 + (U_2 + U_3) = (U_1 + U_2) + U_3,$$
$$U_1 \cap (U_1 + U_2) = U_1, \quad U_1 + (U_1 \cap U_2) = U_1.$$

9.13. Линейная оболочка элементов линейного пространства

Пусть ${}_KV$ — линейное пространство, $v_1, \ldots, v_m \in {}_KV$. Рассмотрим

$$\langle v_1, \ldots, v_m \rangle = \{k_1 v_1 + \ldots + k_m v_m \mid k_1, \ldots, k_m \in K\} -$$

совокупность всех линейных комбинаций $k_1 v_1 + \ldots + k_m v_m$ элементов $v_1, \ldots, v_m$ с коэффициентами $k_1, \ldots, k_m \in K$, называемую *линейной оболочкой* элементов $v_1, \ldots, v_m$. Линейная оболочка $\langle v_1, \ldots, v_m \rangle$ является наименьшим линейным подпространством, содержащим элементы $v_1, \ldots, v_m$. Действительно,

$$(k_1 v_1 + \ldots + k_m v_m) + (l_1 v_1 + \ldots + l_m v_m) = \\ = (k_1 + l_1) v_1 + \ldots + (k_m + l_m) v_m;$$

$$k(k_1 v_1 + \ldots + k_m v_m) = (kk_1) v_1 + \ldots + (kk_m) v_m;$$

если U — линейное подпространство в ${}_KV$, $v_1, \ldots, v_m \in U$, то $k_1 v_1 + \ldots + k_m v_m \in U$, следовательно, $\langle v_1, \ldots, v_m \rangle \subseteq U$. Более того,

$$\langle v_1, \ldots, v_m \rangle = \bigcap_{\substack{U \subseteq {}_KV \\ v_1, \ldots, v_m \in U}} U.$$

Замечание 9.13.1. Если $0 \neq v \in {}_KV$, то $\langle v \rangle = Kv = \{kv \mid k \in K\}$, $\dim \langle v \rangle = 1$; если $v = 0$, $\langle v \rangle = Kv = \{0\}$.

Замечание 9.13.2. $\langle v_1, \ldots, v_m \rangle = Kv_1 + \ldots + Kv_m$.

Замечание 9.13.3. $\dim_K\langle v_1,\ldots,v_m\rangle = \mathrm{r}\{v_1,\ldots,v_m\}$; любая максимальная линейно независимая подсистема в $\{v_1,\ldots,v_m\}$ является базисом линейного подпространства $\langle v_1,\ldots,v_m\rangle$.

Основная лемма о линейной зависимости может быть сформулирована в следующей эквивалентной форме.

Теорема 9.13.4 (о замене). *Пусть $v_1,\ldots,v_s \in {}_KV$ — линейно независимая система, $u_1,\ldots,u_r \in \langle v_1,\ldots,v_s\rangle$, $\{u_1,\ldots,u_r\}$ — линейно независимая система элементов. Тогда $r \leqslant s$ и*

$$\langle v_1,\ldots,v_s\rangle = \langle u_1,\ldots,u_r,v_{i_{r+1}},\ldots,v_{i_s}\rangle,$$

где

$$1 \leqslant i_{r+1} < \ldots < i_s \leqslant s.$$

Доказательство. Так как $s = \dim_K\langle v_1,\ldots,v_s\rangle$, то $r \leqslant s$. Если $r = s$, то $\langle v_1,\ldots,v_s\rangle = \langle u_1,\ldots,u_r\rangle$. Если $r < s$, то найдётся $v_{i_{r+1}} \notin \langle u_1,\ldots,u_r\rangle$ (индекс i_{r+1} — минимальный с этим свойством). Продолжая этот процесс, построим базис $\{u_1,\ldots,u_r,v_{i_{r+1}},\ldots,v_{i_s}\}$ в $\langle v_1,\ldots,v_s\rangle$. □

Следствие 9.13.5. *Пусть U, W — линейные подпространства в ${}_KV$ и $U \subseteq W$, $\dim_K U = l$, $\dim_K W = m$. Тогда $l \leqslant m$ и любой базис подпространства U можно дополнить $m-l$ элементами до базиса подпространства W. В частности, если $U \subseteq W$ и $l = m$, то $U = W$.*

Теорема 9.13.6 (формула размерности). *Пусть U, W — линейные подпространства в ${}_KV$, $\dim_K V = n < \infty$. Тогда*

$$\dim_K U + \dim_K W = \dim_K(U \cap W) + \dim_K(U + W),$$

или, что эквивалентно,

$$\dim_K(U + W) = \dim_K U + \dim_K W - \dim_K(U \cap W).$$

Доказательство. Пусть $\dim_K(U \cap W) = d$, $\dim_K U = s$, $\dim_K W = t$. Ясно, что $0 \leqslant d \leqslant s$, $d \leqslant t$. При $d = 0$ утверждение очевидно (объединение базисов в U и W даёт базис в $U + W$). Выберем базис $v_1,\ldots,v_d$ линейного пространства $U \cap W$ и дополним

его до базиса $v_1, \ldots, v_d, u_1, \ldots, u_{s-d}$ линейного пространства U и до базиса $v_1, \ldots, v_d, w_1, \ldots, w_{t-d}$ линейного пространства W. Ясно, что

$$U + W = \langle v_1, \ldots, v_d, u_1, \ldots, u_{s-d}, w_1, \ldots, w_{t-d} \rangle.$$

Если

$$\lambda_1 v_1 + \ldots + \lambda_d v_d + \mu_1 u_1 + \ldots + \mu_{s-d} u_{s-d} + \gamma_1 w_1 + \ldots + \gamma_{t-d} w_{t-d} = 0,$$

то

$$\sum_{i=1}^{d} \lambda_i v_i + \sum_{j=1}^{s-d} \mu_j u_j = -\sum_{k=1}^{t-d} \gamma_k w_k \in U \cap W,$$

поэтому $\mu_1 = \ldots = \mu_{s-d} = 0$, $\gamma_1 = \ldots = \gamma_{t-d} = 0$. Следовательно, $\lambda_1 = \ldots = \lambda_d = 0$. Таким образом,

$$\{v_1, \ldots, v_d, u_1, \ldots, u_{s-d}, w_1, \ldots, w_{t-d}\} -$$

базис линейного подпространства $U + W$, откуда

$$s + t = d + (s - d) + d + (t - d) = d + (d + (s - d) + (t - d)),$$

поэтому

$$\dim {}_K U + \dim {}_K W = \dim {}_K U \cap W + \dim {}_K (U + W). \quad \square$$

Теорема 9.13.7 (о существовании прямого дополнения подпространства). *Пусть* $\dim {}_K V = n < \infty$, U *— линейное подпространство в* ${}_K V$. *Тогда существует линейное подпространство* W *в* ${}_K V$ *такое, что*

$$U + W = V, \quad U \cap W = \{0\},$$

(*называемое прямым дополнением подпространства* U *в* ${}_K V$; *в этом случае также говорят, что линейное пространство* ${}_K V$ *является прямой суммой линейных подпространств* U *и* W, *обозначение:* ${}_K V = U \oplus W$).

Доказательство. Если $\dim {}_K U = r$ и $\{u_1, \ldots, u_r\}$ — базис в ${}_K U$, то дополним его до базиса линейного пространства ${}_K V$: $u_1, \ldots, u_r, v_1, \ldots, v_{n-r}$. Пусть $W = \langle v_1, \ldots, v_{n-r} \rangle$. Тогда ${}_K V = U + W$, $U \cap W = \{0\}$. $\square$

Замечание 9.13.8. Конечно, прямое дополнение определено неоднозначно, однако все прямые дополнения линейного пространства изоморфны (а именно, все они имеют размерность $\dim {}_KV - \dim {}_KU$).

Замечание 9.13.9. Если ${}_KV = U \oplus W$, то представление элемента $v \in V$ в виде $v = u + w$, $u \in U$, $w \in W$, определено однозначно (действительно, если $v = u + w = u' + w'$, $u' \in U$, $w' \in W$, то $u - u' = w' - w \in U \cap W = \{0\}$, следовательно, $u = u'$, $w = w'$), и поэтому линейное пространство ${}_KV = U \oplus W$ изоморфно *внешней прямой* сумме $\{(u, w) \mid u \in U,\ w \in W\}$ линейных пространств ${}_KU$ и ${}_KW$ с естественными операциями сложения пар и их умножения на $c \in K$.

Пример 9.13.10 (прямого разложения). Пусть

$$V = \mathrm{M}_n(\mathbb{R}), \quad U = \{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = A\},$$
$$W = \{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = -A\}.$$

Тогда

$${}_{\mathbb{R}}V = U \oplus W.$$

Действительно, $A = \dfrac{A + A^*}{2} + \dfrac{A - A^*}{2}$. Если $A = A^* = -A$, то $A = 0 \in \mathrm{M}_n(\mathbb{R})$. □

9.14. Решётка подпространств линейного пространства

Рассмотрим частично упорядоченное множество всех линейных подпространств U линейного пространства ${}_KV$:

$$\mathcal{L}({}_KV) = \{U \mid U \subseteq {}_KV\},$$

где $U_1 \leqslant U_2$ означает $U_1 \subseteq U_2$. Для любых двух элементов $U_1, U_2 \in \mathcal{L}({}_KV)$ существует точная верхняя грань $U_1 \vee U_2 = U_1 + U_2$ и точная нижняя грань $U_1 \wedge U_2 = U_1 \cap U_2$, таким образом, частично упорядоченное множество $\mathcal{L}({}_KV)$ является *решёткой* (*решёткой линейных подпространств* линейного пространства ${}_KV$), при этом $\mathcal{L}({}_KV)$ — решётка с дополнениями (т. е. для всякого $U \in \mathcal{L}({}_KV)$ существует такой элемент $W \in \mathcal{L}({}_KV)$, что $U \vee W = V$, $U \wedge W = \{0\}$).

Теорема 9.14.1. *В решётке $\mathcal{L}({}_KV)$ выполнено следующее модулярное тождество Дедекинда: если $X, Y, Z \in \mathcal{L}({}_KV)$, $X \subseteq Z$, то*

$$X + (Y \cap Z) = (X + Y) \cap Z.$$

Доказательство.

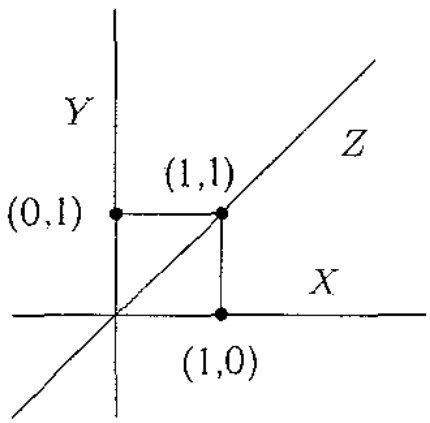
1) Пусть $x + a \in X + (Y \cap Z)$, где $x \in X$, $a \in Y \cap Z$, тогда $a \in Y$, и поэтому $x + a \in X + Y$; $x \in X \subseteq Z$, $a \in Y \cap Z \subseteq Z$, и следовательно, $x + a \in Z$; итак, $x + a \in (X + Y) \cap Z$.

2) Пусть $z \in (X + Y) \cap Z$, $z = x + y$, где $x \in X$, $y \in Y$. Тогда $y = z - x \in Y \cap Z$, поскольку $X \subseteq Z$, и поэтому $z = x + y \in X + (Y \cap Z)$. □

Замечание 9.14.2. Если $\dim {}_KV \geqslant 2$, то в $\mathcal{L}({}_KV)$ не выполняется тождество дистрибутивности

$$(X + Y) \cap Z = (X \cap Z) + (Y \cap Z).$$

Действительно, в ${}_KV = K^2$ имеем для



$$X + Y = {}_KV = K^2, \quad X \cap Z = \{0\}, \quad Y \cap Z = \{0\},$$
$$(X + Y) \cap Z = Z \neq \{0\} = (X \cap Z) + (Y \cap Z). \quad \square$$

Замечание 9.14.3. Итак, мы убедились в том, что решётка $\mathcal{L}({}_KV)$ всех линейных подпространств линейного пространства ${}_KV$ является модулярной (дедекиндовой) решёткой с дополнениями.

9.15. Проективная размерность подпространств и проективная геометрия $\mathrm{PG}({}_K V)$

Если $\dim {}_K V = n$, $U \in \mathcal{L}({}_K V)$ — линейное подпространство в ${}_K V$, то определим *проективную размерность*

$$\mathrm{p.dim}\, U = \dim {}_K U - 1.$$

Таким образом, нулевое подпространство в ${}_K V$ имеет проективную размерность, равную -1; одномерные линейные подпространства имеют нулевую проективную размерность (их называют *точками* проективной геометрии); двумерные линейные подпространства имеют проективную размерность, равную 1 (их называют *прямыми* проективной геометрии); и т. д., $\mathrm{p.dim}\, V = n-1$. Обозначая через G_i совокупность всех $(i+1)$-мерных линейных подпространств в ${}_K V$, получаем $(n-1)$-*мерную проективную геометрию*

$$\mathrm{PG}({}_K V) = \{G_0, G_1, \ldots, G_{n-1}\},$$

где G_0 — множество точек, G_1 — множество прямых, G_2 — плоскостей, G_i — множество i-мерных плоскостей, с отношением инцидентности $U \prec W$ для $U \in G_i$, $W \in G_j$, где $0 \leqslant i \leqslant j \leqslant n-1$, означающим, что $U \subseteq W$.

9.16. Теорема о ранге матрицы

Пусть $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ — прямоугольная $(m \times n)$-матрица с элементами a_{ij} из поля K. Определитель $M_{i_1,\ldots,i_k;j_1,\ldots,j_k}$ квадратной $(k \times k)$-матрицы, состоящей из элементов на пересечении k строк с номерами $i_1, \ldots, i_k$ и k столбцов с номерами $j_1, \ldots, j_k$, называется *минором k-го порядка* матрицы A. Наивысший порядок *ненулевого* минора матрицы A обозначим через $\mathrm{r}(A)$.

Теорема 9.16.1 (о ранге матрицы). *Следующие четыре числовые характеристики матрицы $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ совпадают:*

1) $\mathrm{r}(A_1, \ldots, A_m)$ (*ранг системы строк, в* K^n);

2) $\mathrm{r}(\hat{A}_1, \ldots, \hat{A}_n)$ (*ранг системы столбцов, в* $\hat{K}^m$);

3) $\mathrm{r}(A)$ (*наивысший порядок ненулевого минора*);

4) *число ненулевых строк* r *в ступенчатом виде* $\bar{A}$ *матрицы* A.

(*Это совпадающее число называется рангом матрицы* A *и будет обозначаться через* $\mathrm{r}(A)$).

Доказательство разобьём на четыре леммы.

Лемма 9.16.2. *Пусть матрица* $\tilde{A}$ *получена из матрицы* A *элементарным преобразованием строк (столбцов) 1-го или 2-го типа, тогда* $\mathrm{r}(A) = \mathrm{r}(\tilde{A})$. *Если* $\bar{A}$ *— ступенчатая форма, к которой приводится матрица* A, *то* $\mathrm{r}(A) = \mathrm{r}(\bar{A})$.

Доказательство проведём для преобразований строк (для столбцов всё аналогично).

Случай 1. $A'_i = A_i + cA_j$, $c \in K$, $i \neq j$. Для $k > \mathrm{r}(A)$ рассмотрим минор $\tilde{M} = \tilde{M}_{i_1,\ldots,i_k;j_1,\ldots,j_k}$ в $\tilde{A}$.

а) Если $i \notin \{i_1, \ldots, i_k\}$, то $\tilde{M} = M_{i_1,\ldots,i_k;j_1,\ldots,j_k} = 0$.

б) Если $i, j \in \{i_1, \ldots, i_k\}$, то $\tilde{M} = M_{i_1,\ldots,i_k;j_1,\ldots,j_k} = 0$.

в) Если $i \in \{i_1, \ldots, i_k\} \not\ni j$, то разложим определитель $\tilde{M}$ по i-й строке $A'_i = A_i + cA_j$ в сумму двух определителей: $\tilde{M} = M + c\tilde{\Delta} = 0$, так как $M = M_{i_1,\ldots,i_k;j_1,\ldots,j_k} = 0$, поскольку $k > \mathrm{r}(A)$, определитель $\tilde{\Delta}$ в качестве i-й строчки имеет часть строки A_j, но $j \notin \{i_1, \ldots, i_k\}$, и поэтому $\tilde{\Delta}$ отличается от минора матрицы порядка k перестановкой двух строк, и поэтому $\tilde{\Delta} = 0$. Итак, $\mathrm{r}(\tilde{A}) \leqslant \mathrm{r}(A)$. Поскольку от A к $\tilde{A}$ можно вернуться элементарным преобразованием строк, то $\mathrm{r}(A) \leqslant \mathrm{r}(\tilde{A})$.

Случай 2. $A_i \leftrightarrow A_j$ разбирается аналогично ($i, j \in \{i_1, \ldots, i_k\}$; $i, j \notin \{i_1, \ldots, i_k\}$; $i \in \{i_1, \ldots, i_k\} \not\ni j$). □

Лемма 9.16.3 (о сохранении линейных соотношений между столбцами при элементарных преобразованиях строк). *Пусть от матрицы* A *к матрице* A' *мы перешли элементарными преобразованиями строк, тогда столбцы матриц* A *и* A' *имеют одни и те же линейные соотношения, а именно,* $k_1\hat{A}_1 + \ldots + k_n\hat{A}_n = 0$ *тогда и только тогда, когда* $k_1\hat{A}'_1 + \ldots + k_n\hat{A}'_n = 0$.

Доказательство. Ясно, что элементарные преобразования 1-го и 2-го типа для строк сохраняют линейное соотношение для столбцов и эти преобразования обратимы. □

Следствие 9.16.4. *Система столбцов $\hat{A}_{j_1}, \ldots, \hat{A}_{j_r}$ матрицы A линейно зависима (соответственно, линейно независима или является максимальной линейно независимой подсистемой в $\hat{A}_1, \ldots, \hat{A}_n \in \hat{K}^m$) тогда и только тогда, когда соответствующая система столбцов (с теми же номерами) $\hat{A}'_{j_1}, \ldots, \hat{A}'_{j_r}$ матрицы A' линейно зависима (соответственно линейно независима или является максимальной линейно независимой подсистемой в $\hat{A}'_1, \ldots, \hat{A}'_n \in \hat{K}^m$).*

Следствие 9.16.5. $\mathrm{r}\{\hat{A}_1, \ldots, \hat{A}_n\} = \mathrm{r}\{\hat{A}'_1, \ldots, \hat{A}'_n\}$.

Лемма 9.16.6. *Если $\bar{A}$ — ступенчатая матрица, то наивысший порядок ненулевого минора $\mathrm{r}(\bar{A})$ совпадает с числом r ненулевых строк.*

Доказательство.

1) Минор r-го порядка на пересечении r ненулевых строк и столбцов, проходящих через уголки ступенек, является определителем треугольной матрицы с ненулевыми элементами на главной диагонали, и поэтому отличен от нуля.

2) Все миноры, порядок которых больше r, нулевые, так как имеют нулевую строку. □

Лемма 9.16.7. *В ступенчатой матрице $\bar{A}$ ранг системы столбцов совпадает с числом r ненулевых строк (а именно, столбцы, проходящие через уголки ступенек, образуют максимальную линейно независимую подсистему столбцов).*

Доказательство.

1) Указанные столбцы линейно независимы, так как проходят через $(r \times r)$-матрицу с ненулевым определителем.

2) Любой столбец ступенчатой матрицы является линейной комбинацией указанных. □

Следствие 9.16.8 (алгоритм нахождения максимальной линейно независимой подсистемы в системе столбцов прямоугольной матрицы). *От матрицы A перейдём к ступенчатой матрице $\bar{A}$ с помощью элементарных преобразований строк 1-го и 2-го типов, запомним номера столбцов $j_1, \ldots, j_r$, проходящих через уголки ступенек в $\bar{A}$, в матрице A возьмём столбцы с этими номерами $\hat{A}_{j_1}, \ldots, \hat{A}_{j_r}$.*

Пример 9.16.9. Найти какую-либо максимальную линейно независимую подсистему строк в системе $a_1, a_2, a_3, a_4 \in \mathbb{R}^4$,

$$a_1 = (-1, 4, -3, -2), \quad a_2 = (3, -7, 5, 3),$$
$$a_3 = (3, -2, 1, 0), \quad a_4 = (-4, 1, 0, 1),$$

а остальные строки выразить как линейные комбинации строк этой подсистемы.

Решение. Записываем строки a_1, a_2, a_3, a_4 как столбцы и приводим полученную матрицу к главному ступенчатому виду с помощью элементарных преобразований строк:

$$\begin{pmatrix} -1 & 3 & 3 & -4 \\ 4 & -7 & -2 & 1 \\ -3 & 5 & 1 & 0 \\ -2 & 3 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & 3 & -4 \\ 0 & 5 & 10 & -15 \\ 0 & -4 & -8 & 12 \\ 0 & -3 & -6 & 9 \end{pmatrix} \to$$
$$\to \begin{pmatrix} -1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 & -5 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Записываем номера столбцов в ступенчатом виде, проходящие через уголки ступенек: 1, 2. Поэтому $\{a_1, a_2\}$ — максимальная линейно независимая подсистема, $a_3 = 3a_1 + 2a_2$, $a_4 = -5a_1 - 3a_2$; ранг системы строк a_1, a_2, a_3, a_4 равен 2.

Завершение доказательства теоремы о ранге:

$$\mathrm{r}(\hat{A}_1, \ldots, \hat{A}_n) \overset{\text{лемма 9.16.3}}{=} \mathrm{r}(\hat{\bar{A}}_1, \ldots, \hat{\bar{A}}_n) \text{ (ранг столбцов}$$
$$\text{ступенчатой матрицы } \bar{A}) \overset{\text{лемма 9.16.7}}{=} r \overset{\text{лемма 9.16.6}}{=}$$
$$= \mathrm{r}(\bar{A}) \overset{\text{лемма 9.16.2}}{=} \mathrm{r}(A) = \mathrm{r}(A^*) \overset{\text{лемма 9.16.3}}{=} \mathrm{r}(A_1, \ldots, A_m). \quad \square$$

Теорема 9.16.10. *Пусть* $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$, $B = (b_{ij}) \in \in \mathrm{M}_{n,r}(K)$. *Тогда*

$$\mathrm{r}(AB) \leqslant \mathrm{r}(A), \quad \mathrm{r}(AB) \leqslant \mathrm{r}(B).$$

Доказательство. Пусть $C = (c_{ij}) = AB$. Тогда

$$c_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj},$$
$$C_i = a_{i1}B_1 + \ldots + a_{in}B_n,$$
$$\hat{C}_j = \hat{A}_1 b_{1j} + \ldots + \hat{A}_n b_{nj},$$

т. е. строки матрицы C линейно выражаются через строки матрицы B, столбцы матрицы C линейно выражаются через столбцы матрицы A. Поэтому $\mathrm{r}(C) \leqslant \mathrm{r}(B)$ и $\mathrm{r}(C) \leqslant \mathrm{r}(A)$. □

Следствие 9.16.11. *При умножении на квадратную матрицу A с $|A| \neq 0$ ранг не меняется.*

Доказательство. Так как $|A| \neq 0$, то существует обратная матрица A^{-1}. Поэтому

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

и следовательно,

$$\mathrm{r}(B) \leqslant \mathrm{r}(BA), \quad \mathrm{r}(B) \leqslant \mathrm{r}(AB).$$

Ранее мы доказали, что

$$\mathrm{r}(B) \geqslant \mathrm{r}(BA), \quad \mathrm{r}(B) \geqslant \mathrm{r}(AB).$$

Поэтому

$$\mathrm{r}(B) = \mathrm{r}(BA), \quad \mathrm{r}(B) = \mathrm{r}(AB). \quad \square$$

Задачи 9.16.12.

1) В условиях теоремы:

$$\mathrm{r}(A) + \mathrm{r}(B) - n \leqslant \mathrm{r}(AB).$$

2) Если $A, B, C \in \mathrm{M}_n(K)$ и $ABC = 0$, то

$$\mathrm{r}(A) + \mathrm{r}(B) + \mathrm{r}(C) \leqslant 2n.$$

3) Пусть $A \in \mathrm{M}_{m,n}(K)$, $B \in \mathrm{M}_{n,m}(K)$ и $m > n$. Покажите, что $\det(AB) = 0$.

Доказательство. Так как $AB \in \mathrm{M}_m(K)$, то

$$\mathrm{r}(AB) \leqslant \mathrm{r}(B) \leqslant n < m. \quad \square$$

4) Если $A^2 = A \in \mathrm{M}_n(K)$, то

$$\mathrm{r}(A) + \mathrm{r}(E - A) = n.$$

5) Если $A, B \in \mathrm{M}_n(K)$ и $A^2 = A$, $AB = 0 = BA$, то

$$\mathrm{r}(A + B) = \mathrm{r}(A) + \mathrm{r}(B).$$

6) Если $A, B \in \mathrm{M}_n(K)$, $AB = BA$, $\mathrm{r}(A^2) = \mathrm{r}(A)$ и $\mathrm{r}(B^2) = \mathrm{r}(B)$, то

$$\mathrm{r}((AB)^2) = \mathrm{r}(AB).$$

7) Если $A_1, \ldots, A_k \in \mathrm{M}_n(K)$, $k \geqslant 2$, то

$$\mathrm{r}(A_1 \ldots A_k) \geqslant \mathrm{r}(A_1) + \ldots + \mathrm{r}(A_k) - n(k-1).$$

Теорема 9.16.13 (о факториальном ранге). *Пусть $m, n \in \mathbb{N}$, $A \in \mathrm{M}_{m,n}(K)$. Ранг матрицы $\mathrm{r}(A)$ равен наименьшему числу k такому, что*

$$A = B \cdot C, \quad \text{где } B \in \mathrm{M}_{m,k}(K), \quad C \in \mathrm{M}_{k,n}(K)$$

(это число k называется факториальным рангом матрицы A).

Доказательство. Допустим, что $A = B \cdot C$, где $B \in \mathrm{M}_{m,n}(K)$, $C \in \mathrm{M}_{k,n}(K)$. Тогда система столбцов матрицы A линейно выражается через систему столбцов матрицы B (их k штук). Поэтому $\mathrm{r}(A) \leqslant k$.

Пусть $k = \mathrm{r}(A)$. Выберем строки $A_{i_1}, \ldots, A_{i_k}$, образующие максимальную линейно независимую подсистему строк $A_1, \ldots, A_m$ матрицы A,

$$A_i = \beta_{i1} A_{i_1} + \ldots + \beta_{ik} A_{i_k}, \quad \beta_{ij} \in K, \quad 1 \leqslant i \leqslant m.$$

Рассмотрим матрицы $B \in \mathrm{M}_{m,k}(K)$, $B = (\beta_{ij})$, и $C \in \mathrm{M}_{k,n}(K)$, для которой j-я строка $C_j = A_{i_j}$, $j = 1, \ldots, k$. Тогда $A = B \cdot C$. $\square$

Теорема 9.16.14 (теорема Кронекера—Капелли: критерий совместности и определённости системы линейных уравнений в терминах рангов матриц). *Пусть* $(a_{ij} \mid b_i)$ *— система* m *линейных уравнений с* n *неизвестными,* $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ *— матрица коэффициентов,*

$$A' = \left(\begin{array}{c|c} & b_1 \\ A & \vdots \\ & b_m \end{array}\right) -$$

расширенная матрица системы линейных уравнений.

а) *Система линейных уравнений совместна тогда и только тогда, когда ранг матрицы коэффициентов* A *равен рангу расширенной матрицы* $A' = (A, \hat{b})$, $\mathrm{r}(A) = \mathrm{r}(A')$.

б) *Система линейных уравнений определённая тогда и только тогда, когда* $\mathrm{r}(A) = \mathrm{r}(A') = n$.

Доказательство.

1) Используя определение ранга матрицы с помощью столбцов, видим, что всегда $\mathrm{r}(A) \leqslant \mathrm{r}(A')$.

2) Если $(k_1, \ldots, k_n)$ — решение, то

$$k_1\hat{A}_1 + \ldots + k_n\hat{A}_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

т. е. столбцы матрицы A' линейно выражаются через столбцы матрицы A, следовательно, $\mathrm{r}(A') \leqslant \mathrm{r}(A)$, и поэтому $\mathrm{r}(A') = \mathrm{r}(A)$.

3) Пусть $\mathrm{r}(A') = \mathrm{r}(A) = r$. Тогда максимальная линейно независимая система столбцов матрицы A содержит r столбцов, и поэтому она является и максимальной линейно независимой системой столбцов матрицы A'. Таким образом, столбец

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

линейно выражается через эту систему столбцов матрицы A, а поэтому и через *все* столбцы матрицы A,

$$k_1\hat{A}_1 + \ldots + k_n\hat{A}_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Итак, существует решение $(k_1, \ldots, k_n)$ системы линейных уравнений. □

Второе доказательство. Элементарными преобразованиями приведём систему линейных уравнений к ступенчатому виду (ранги матриц не меняются при этом). Совпадение рангов означает отсутствие «экзотических» уравнений в ступенчатом виде, т. е. совместность системы линейных уравнений. □

4) *Доказательство критерия определённости (в терминах рангов).* Если система определена, т. е. $\mathrm{r}(A) = \mathrm{r}(A')$, то она определена тогда и только тогда, когда в ступенчатом виде нет свободных неизвестных, т. е. $\mathrm{r}(A) = \mathrm{r}(A') = n$. □

9.17. Размерность пространства решений однородной системы линейных уравнений

Как мы отметили ранее, совокупность решений $X_{\text{одн}}$ однородной системы линейных уравнений с матрицей $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ является линейным пространством, подпространством в K^n.

Теорема 9.17.1. *Если $r = \mathrm{r}(A) < n$, то $\dim X_{\text{одн}} = n - r$ (т. е. размерность пространства решений равна числу свободных неизвестных). (Если $\mathrm{r}(A) = n$, то система линейных уравнений имеет лишь нулевое решение.)*

Доказательство. Для удобства записи переупорядочим неизвестные, если это необходимо, так, чтобы

$$\underbrace{x_1, \ldots, x_r}_{r \text{ главных неизвестных}} \quad \text{и} \quad \underbrace{x_{r+1}, \ldots, x_n}_{n-r \text{ свободных неизвестных}}.$$

Пусть $E = E_{n-r} \in \mathrm{M}_{n-r}(K)$ — единичная матрица размера $(n-r) \times (n-r)$. Возьмём её строки в качестве наборов значений для свободных неизвестных и дополним их (единственно возможным способом) до решений нашей системы линейных уравнений

$$\alpha_1 = (c_{11}, \ldots, c_{1r}, 1, 0, \ldots, 0),$$
$$\vdots$$
$$\alpha_{n-r} = (c_{(n-r)1}, \ldots, c_{(n-r)r}, 0, 0, \ldots, 1).$$

Эта система $n-r$ строк-решений линейно независима (поскольку строки единичной матрицы, конечно, линейно независимы). Если

$$\beta = (\beta_1, \ldots, \beta_{n-r}, \beta_{n-r+1}, \ldots, \beta_n) \in X_{\text{одн}} -$$

произвольное решение, то

$$\gamma = \beta - \beta_{n-r+1}\alpha_1 - \ldots - \beta_n\alpha_{n-r} = (\gamma_1, \ldots, \gamma_{n-r}, 0, \ldots, 0) \in X_{\text{одн}}.$$

Однако, конечно,

$$(0, \ldots, 0, 0, \ldots, 0) \in X_{\text{одн}},$$

при этом γ и нулевое решение имеют одинаковый набор значений для свободных неизвестных. Так как значения главных неизвестных однозначно определяются по свободным, то $\gamma = 0$, следовательно,

$$\beta = \beta_{n-r+1}\alpha_1 + \ldots + \beta_n\alpha_{n-r}.$$

Итак, мы построили базис $\{\alpha_1, \ldots, \alpha_{n-r}\}$ линейного пространства решений $X_{\text{одн}}$, поэтому $\dim X_{\text{одн}} = n - r$. □

Замечание 9.17.2. Если вместо строк единичной матрицы E_{n-r} для свободных неизвестных брать строки всевозможных матриц $C \in \mathrm{GL}_{n-r}(K)$ (т. е. $C \in \mathrm{M}_n(K)$, $|C| \neq 0$), то этот алгоритм позволяет построить все базисы в $X_{\text{одн}}$.

Замечание 9.17.3. Любой базис линейного пространства решений $X_{\text{одн}}$ однородной системы линейных уравнений называется в ряде алгебраических текстов «фундаментальной системой решений однородной системы линейных уравнений».

9.18. Задание любого подпространства в ${}_KV = K^n$ как пространства решений однородной системы линейных уравнений

Пусть K — поле, $u_1, \ldots, u_m \in {}_KV = K^n$, $U = \langle u_1, \ldots, u_m \rangle$ — подпространство в K^n, являющееся линейной оболочкой строк $u_1, \ldots, u_m$, т. е. множеством всех линейных комбинаций строк $u_1, \ldots, u_m$. Мы найдём такую матрицу $A \in \mathrm{M}_{s,n}(K)$, что множество решений однородной системы линейных уравнений

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

совпадает с U.

Если U — нулевое подпространство, то в качестве A мы можем взять любую матрицу $n \times n$ с ненулевым определителем (например, $A = E$). Если $U = K^n$ (это эквивалентно тому, что $\dim U = n$), то в качестве A мы можем взять нулевую матрицу из $\mathrm{M}_{s,n}$, $s \geqslant 1$. Если же $1 \leqslant \dim U = \mathrm{r}(u_1, \ldots, u_m) < n$, то пусть $u_i = (u_{i1}, u_{i2}, \ldots, u_{in})$, $1 \leqslant i \leqslant m$, $u_{ij} \in K$.

Рассмотрим матрицу $B \in \mathrm{M}_{m,n}(K)$, $B = (b_{ij})$, $b_{ij} = u_{ij}$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, и однородную систему линейных уравнений

$$B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \left.\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\right\} m. \tag{9.2}$$

Ясно, что $r = \mathrm{r}(B) = \dim U$, поэтому $1 \leqslant r < n$. Размерность s пространства решений $X_{\text{одн}}$ этой системы равна $n - r$, и так как $1 \leqslant r < n$, то $1 \leqslant s < n$.

Пусть строки $v_1, \ldots, v_s \in K^n$ образуют фундаментальную систему решений системы (9.2), $v_i = (v_{i1}, \ldots, v_{in})$, $1 \leqslant i \leqslant s$, $v_{ij} \in K$. Пусть $A \in \mathrm{M}_{s,n}(K)$, $A = (a_{ij})$, $a_{ij} = v_{ij}$, $1 \leqslant i \leqslant s$, $1 \leqslant j \leqslant n$. Покажем, что A — искомая матрица.

Действительно, по построению матрицы A любая строка из U (как линейная комбинация строк $u_1, \ldots, u_m$) является решением од-

нородной системы уравнений

$$A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{9.3}$$

т. е. $U \subseteq X_{\text{одн}}$. С другой стороны,

$$\dim X_{\text{одн}} = n - \mathrm{r}(A) = n - s = n - (n - r) = r = \dim U.$$

Следовательно, $U = X_{\text{одн}}$. □

В заключение отметим, что матрица A определена неоднозначно. Например, другая матрица A' может быть получена с помощью другой фундаментальной системы решений системы (9.2).

Полученное задание линейных подпространств оказывается полезным при решении ряда практических задач. Например, пусть $u_1, \ldots, u_m \in \mathbb{R}^n$ — линейно независимые строки, $m < n$. Требуется найти такие строки $u_{m+1}, \ldots, u_n$, что $\{u_1, \ldots, u_n\}$ — базис линейного пространства $\mathbb{R}^n$. Как и выше, пусть $v_1, \ldots, v_s$ — какая-нибудь фундаментальная система решений системы (9.2) (в нашем случае $\mathrm{r}(B) = m$, $s = n - m$). Положим $u_{m+1} = v_1, \ldots, u_n = v_{n-m}$. Покажем, что $\{u_1, \ldots, u_n\}$ — базис в $\mathbb{R}^n$. Достаточно показать, что строки $u_1, \ldots, u_n$ линейно независимы над $\mathbb{R}$. Пусть $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ и $\alpha_1 u_1 + \ldots + \alpha_n u_n = 0 \in \mathbb{R}^n$. Тогда для строки

$$z = \alpha_1 u_1 + \ldots + \alpha_m u_m = -\alpha_{m+1} u_{m+1} - \ldots - \alpha_n u_n$$

имеем $z \in U \cap V$, где $V = \langle u_{m+1}, \ldots, u_n \rangle$. Если $z = (z_1, \ldots, z_n)$, $z_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, то по построению подпространств U и V (см. (9.2), (9.3)) имеем

$$(z_1, \ldots, z_n)\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = 0,$$

$z_1^2 + \ldots + z_n^2 = 0$, следовательно, $z_1 = \ldots = z_n = 0$, и $z = 0 \in \mathbb{R}^n$. Значит,

$$\alpha_1 u_1 + \ldots + \alpha_m u_m = 0 (\in \mathbb{R}^n) = \alpha_{m+1} u_{m+1} + \ldots + \alpha_n u_n.$$

Но $u_1, \ldots, u_m$ — линейно независимые строки, поэтому $\alpha_1 = \ldots = \alpha_m = 0$. Строки $u_{m+1}, \ldots, u_n$ также линейно независимы, следовательно, $\alpha_{m+1} = \ldots = \alpha_n = 0$. Итак, $\alpha_1 = \ldots = \alpha_n = 0$ и строки $u_1, \ldots, u_n$ линейно независимы.

Таким образом, мы рассмотрели два способа задания линейных подпространств в ${}_KV = K^n$:

1) как множество решений $X_{\text{одн}}$ однородной системы линейных уравнений;

2) как линейную оболочку $\langle u_1, \ldots, u_m \rangle$ строк $u_1, \ldots, u_m \in {}_KV = K^n$.

При этом мы научились переходить от первого задания ко второму (фундаментальная система решений) и от второго задания к первому. Первый способ задания удобен для задания пересечения $U \cap W$ подпространств (надо к первой однородной системе уравнений приписать вторую). Второй способ задания удобен для задания суммы подпространств:

$$\langle u_1, \ldots, u_m \rangle + \langle w_1, \ldots, w_t \rangle = \langle u_1, \ldots, u_m, w_1, \ldots, w_t \rangle.$$

В следующем примере мы увидим комбинацию этих приёмов.

Пример 9.18.1. Пусть $V_1 = \langle u_1, u_2, u_3 \rangle \subseteq \mathbb{R}^4$ (линейная оболочка строк $u_1 = (1,1,0,0)$, $u_2 = (0,1,1,0)$, $u_3 = (0,0,1,1)$), $V_2 = \langle v_1, v_2, v_3 \rangle \subseteq \mathbb{R}^4$ (линейная оболочка строк $v_1 = (1,0,1,0)$, $v_2 = (0,2,1,1)$, $v_3 = (1,2,1,2)$). Необходимо найти базисы линейных пространств $V_1 + V_2$ и $V_1 \cap V_2$, при этом строки u_1, u_2, u_3, v_1, v_2, v_3 выразить через базис пространства $V_1 + V_2$.

Решение. Запишем строки u_1, u_2, u_3, v_1, v_2, v_3 по столбцам и приведём полученную матрицу к ступенчатому виду с помощью элементарных преобразований строк:

$$\begin{matrix} u_1 & u_2 & u_3 & v_1 & v_2 & v_3 \end{matrix}$$
$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 & 2 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \to \begin{array}{c} \begin{matrix} u_1' & u_2' & u_3' & v_1' & v_2' & v_3' \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \end{array}.$$

Поскольку $V_1 + V_2 = \langle u_1, u_2, u_3, v_1, v_2, v_3 \rangle$ и элементарные преобразования строк матрицы не меняют линейных соотношений между столбцами, то $\{u_1, u_2, u_3, v_1\}$ — базис в $V_1 + V_2$ (и так как $\dim(V_1 + V_2) = 4$, то $V_1 + V_2 = \mathbb{R}^4$). Из ступенчатого вида мы вычисляем v_2' и v_3' через u_1', u_2', u_3', v_1':

$$v_2' + v_1' = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = u_1' + u_2' + u_3'.$$

Поэтому $v_2' = u_1' + u_2' + u_3' - v_1'$ и, следовательно, $v_2 = u_1 + u_2 + u_3 - v_1$. Для v_3' мы видим, что $v_3' + v_1' = (2, 0, 2, 0)^* = 2u_1' + 2u_3'$, поэтому $v_3 = 2u_1 + 2u_3 - v_1$. Проведённые вычисления равносильны завершению приведения матрицы к главному ступенчатому виду:

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

Рассмотрим теперь $V_1 \cap V_2$. Для этого найдём однородные системы линейных уравнений, чьи множества решений совпадают с V_1 и V_2 соответственно.

Для V_1:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

система уже имеет ступенчатый вид, x_1, x_2, x_3 — главные неизвестные, x_4 — свободная. Фундаментальная система решений состоит из одной строки $(-1, 1, -1, 1)$. Итак, подпространство V_1 совпадает

с пространством решений однородной системы линейных уравнений

$$(-1,1,-1,1)\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=0. \tag{9.4}$$

Для V_2:

$$\begin{pmatrix}1&0&1&0\\0&2&1&1\\1&2&1&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\to\begin{pmatrix}1&0&1&0\\0&2&1&1\\0&2&0&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\to$$

$$\to\begin{pmatrix}1&0&1&0\\0&2&1&1\\0&1&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\to\begin{pmatrix}1&0&1&0\\0&1&0&1\\0&2&1&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\to$$

$$\to\begin{pmatrix}1&0&1&0\\0&1&0&1\\0&0&1&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

и мы приходим к ступенчатому виду, при этом x_1, x_2, x_3 — главные неизвестные, а x_4 — свободная. Фундаментальная система решений состоит из одной строки $(-1,-1,1,1)$. Значит, однородная система линейных уравнений

$$(-1,-1,1,\mathbf{1})\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=0 \tag{9.5}$$

задаёт подпространство V_2.

Ясно, что система

$$\begin{pmatrix}-1&1&-1&1\\-1&-1&1&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

задаёт подпространство $V_1 \cap V_2$. Решим эту систему:

$$\begin{pmatrix} -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

x_1, x_2 — главные неизвестные, x_3, x_4 — свободные неизвестные. Фундаментальная система решений состоит из двух строк

$$u = (0, 1, 1, 0),$$
$$v = (1, 0, 0, 1).$$

Следовательно, $\{u, v\}$ — базис линейного подпространства $V_1 \cap V_2$.

9.19. Собственные числа и собственные векторы матрицы

Пусть K — поле, $A \in \mathrm{M}_n(K)$, $0 \neq \hat{X} \in \hat{K}^n = \mathrm{M}_{n,1}(K)$, $\lambda \in K$. Если $A \cdot \hat{X} = \lambda \cdot \hat{X}$, то λ называется *собственным числом* матрицы A, а $\hat{X}$ — *собственным вектором* матрицы A, отвечающим собственному числу λ.

Условие $A \cdot \hat{X} = \lambda \cdot \hat{X}$ эквивалентно условию

$$(A - \lambda E)\hat{X} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \hat{K}^n,$$

где $E \in \mathrm{M}_n(K)$ — единичная матрица. При фиксированном λ это условие превращается в однородную систему линейных уравнений

относительно неизвестных $x_1, \ldots, x_n$,

$$\hat{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Матрица $A - \lambda E$ этой системы — квадратная матрица размера n. Поэтому наличие ненулевого решения этой системы равносильно тому, что $|A - \lambda E| = 0$. Пусть t — переменная,

$$p(t) = |A - tE| = p_n t^n + p_{n-1} t^{n-1} + \ldots + p_0 \in K[t] —$$

многочлен степени n от переменной t (называемый характеристическим многочленом матрицы A), при этом:

$$p_n = (-1)^n,$$

$$p_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{ii} = (-1)^{n-1} \operatorname{tr} A, \quad p_0 = |A|.$$

Мы показали, что собственные числа и только они являются корнями характеристического многочлена из поля K.

Если $\lambda \in K$ и $p(\lambda) = 0$, то все собственные векторы матрицы A относительно собственного числа λ — это все ненулевые решения системы

$$(A - \lambda E)\hat{X} = (0) \in \hat{K}^n.$$

Отметим, что множество всех собственных векторов матрицы A относительно собственного числа λ не образует линейного подпространства в $\hat{K}^n$, так как все эти векторы ненулевые. Но если к этому множеству добавить нулевой вектор, то получится линейное подпространство всех решений системы

$$(A - \lambda E)\hat{X} = (0).$$

Таким образом, если $p(\lambda) = |A - \lambda E| = 0$, $r = \mathrm{r}(A - \lambda E)$, то $0 \leqslant r < n$, то размерность пространства решений этой системы равна $s = n - r$, поэтому $1 \leqslant s \leqslant n$. Если $\{X_1, \ldots, X_s\}$ — какая-либо фундаментальная система решений системы $(A - \lambda E)\hat{X} = (0)$, то все собственные векторы матрицы A, отвечающие собственному числу λ, — это все нетривиальные линейные комбинации элементов $\hat{X}_1, \ldots, \hat{X}_s$ с коэффициентами из поля K.

Пример 9.19.1.

$$A = \begin{pmatrix} 10 & 3 \\ -5 & 2 \end{pmatrix}, \quad K = \mathbb{R},$$

$$|A - \lambda E| = \begin{vmatrix} 10-\lambda & 3 \\ -5 & 2-\lambda \end{vmatrix} = \lambda^2 - 12\lambda + 35.$$

Корни: $\lambda_1 = 7$, $\lambda_2 = 5$, $\lambda_1, \lambda_2 \in \mathbb{R}$ (собственные числа матрицы A).

Собственные векторы для $\lambda_1 = 7$:

$$A - 7E = \begin{pmatrix} 3 & 3 \\ -5 & -5 \end{pmatrix}, \quad \begin{pmatrix} 3 & 3 \\ -5 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения:

$$\left\{ \begin{pmatrix} s \\ -s \end{pmatrix} \middle| \, s \in \mathbb{R}, \ s \neq 0 \right\}.$$

Собственные векторы для $\lambda_2 = 5$:

$$A - 5E = \begin{pmatrix} 5 & 3 \\ -5 & -3 \end{pmatrix}, \quad \begin{pmatrix} 5 & 3 \\ -5 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения:

$$\left\{ \begin{pmatrix} -3t \\ 5t \end{pmatrix} \middle| \, t \in \mathbb{R}, \ t \neq 0 \right\}.$$

Пример 9.19.2.

$$K = \mathbb{C}, \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

$$p(\lambda) = |A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3.$$

Имеется лишь одно собственное число: $\lambda = 0$. Собственные векторы относительно $\lambda = 0$ задаются системой линейных уравнений

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Система уже имеет ступенчатый вид, x_2, x_3 — главные неизвестные, x_1 — свободная переменная, множество собственных векторов относительно $\lambda = 0$:

$$\left\{ \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} \middle| \, s \in \mathbb{C}, \ s \neq 0 \right\}.$$

Пример 9.19.3. Если

$$A = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} -$$

диагональная матрица, то $\alpha_1, \ldots, \alpha_n$ — все корни характеристического многочлена матрицы A (и следовательно, собственные числа).

Пример 9.19.4.

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

$$p(\lambda) = |A - \lambda E| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

а) $K = \mathbb{R}$: нет действительных корней многочлена $p(\lambda) = \lambda^2 + 1$, поэтому для матрицы A нет действительных собственных чисел (и собственных векторов).

б) $K = \mathbb{C}$: многочлен $p(\lambda)$ имеет корни $\lambda_1 = i \in \mathbb{C}$, $\lambda_2 = -i \in \mathbb{C}$ (собственные числа матрицы A).

Собственные векторы для $\lambda = i$:

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения:

$$\left\{ c \cdot \begin{pmatrix} -i \\ 1 \end{pmatrix} \middle| \, c \in \mathbb{C}, \ c \neq 0 \right\}.$$

Собственные векторы для $\lambda = -i$:

$$\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения

$$\left\{ c \cdot \begin{pmatrix} i \\ 1 \end{pmatrix} \middle| c \in \mathbb{C},\ c \neq 0 \right\}.$$

Пример 9.19.5.

$$K = \mathbb{R}, \quad A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix},$$

$$p(\lambda) = |A - \lambda E| = -\lambda^3 + 12\lambda + 16.$$

Корни многочлена $p(\lambda)$: $\lambda_1 = -2$, $\lambda_2 = -2$, $\lambda_3 = 4$ (собственные числа).

Собственные векторы для $\lambda = -2$:

$$(A + 2E) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ненулевые решения

$$\left\{ \begin{pmatrix} s - t \\ s \\ t \end{pmatrix} \middle| s, t \in \mathbb{R},\ s^2 + t^2 \neq 0 \right\}.$$

Собственные векторы для $\lambda = 4$:

$$(A - 4E) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ненулевые решения

$$\left\{ \begin{pmatrix} s \\ s \\ 2s \end{pmatrix} \middle| s \in \mathbb{R},\ s \neq 0 \right\}.$$

Задача 9.19.6 (уравнение Сильвестера). Пусть $A \in \mathrm{M}_n(\mathbb{C})$, $B \in \mathrm{M}_m(\mathbb{C})$, $C \in \mathrm{M}_{n,m}(\mathbb{C})$ и матрицы A и B не имеют общих собственных чисел. Тогда матричное уравнение Сильвестера $AX - XB = C$ имеет единственное решение $X \in \mathrm{M}_{n,m}(\mathbb{C})$.

Задача 9.19.7. Пусть $A, B \in \mathrm{M}_n(\mathbb{C})$, $AB = BA$. Покажите, что для матриц A и B существует общий собственный вектор.

Трудная задача 9.19.8. Пусть $A, B \in \mathrm{M}_n(\mathbb{C})$ и $\mathrm{r}(AB - BA) = 1$. Тогда для матриц A и B существует общий собственный вектор.

Теорема 9.19.9. *Пусть* $A \in \mathrm{M}_n(K)$, $0 \neq \hat{X}_1, \ldots, \hat{X}_l \in \hat{K}_n$, $\lambda_1, \ldots, \lambda_l \in K$, $\lambda_i \neq \lambda_j$ *при* $1 \leqslant i \neq j \leqslant l$, $A \cdot \hat{X}_i = \lambda_i \cdot \hat{X}_i$, $i = 1, \ldots, l$. *Тогда столбцы* $\hat{X}_1, \ldots, \hat{X}_l$ *линейно независимы, т. е. собственные векторы, отвечающие различным собственным значениям, линейно независимы.*

Доказательство. Доказательство проведём индукцией по l. Основание индукции: $l = 1$, $A \cdot \hat{X}_1 = \lambda_1 \cdot \hat{X}_1$, $0 \neq \hat{X}_1$, $\{\hat{X}_1\}$ — линейно независимая система векторов.

Пусть теперь $l \geqslant 2$ и наше утверждение доказано для всех l', $1 \leqslant l' < l$. Допустим, что

$$\alpha_1 \hat{X}_1 + \ldots + \alpha_l \hat{X}_l = 0 \in \hat{K}^n, \quad \alpha_i \in K, \quad i = 1, \ldots, l. \tag{9.6}$$

Умножая слева на матрицу A обе части равенства, получаем, что

$$\alpha_1 \cdot A \cdot \hat{X}_1 + \ldots + \alpha_l \cdot A \cdot \hat{X}_l = A \cdot 0 = 0 \in \hat{K}^n,$$

и поэтому

$$\alpha_1 \lambda_1 \hat{X}_1 + \ldots + \alpha_l \lambda_l \hat{X}_l = 0. \tag{9.7}$$

Умножая (9.6) на λ_l, имеем

$$\alpha_1 \lambda_l \hat{X}_1 + \ldots + \alpha_l \lambda_l \hat{X}_l = 0. \tag{9.8}$$

Вычитаем (9.8) из (9.7):

$$\alpha_1 (\lambda_1 - \lambda_l) \hat{X}_1 + \ldots + \alpha_{l-1} (\lambda_{l-1} - \lambda_l) \hat{X}_{l-1} = 0 \in \hat{K}^n.$$

Применяя предположение индукции, получаем, что

$$\alpha_1 (\lambda_1 - \lambda_l) = \ldots = \alpha_{l-1} (\lambda_{l-1} - \lambda_l) = 0.$$

Поскольку $\lambda_1 \neq \lambda_l, \ldots, \lambda_{l-1} \neq \lambda_l$, отсюда следует, что $\alpha_1 = \ldots = \alpha_{l-1} = 0$. Следовательно, из (9.6) следует, что $\alpha_l \hat{X}_l = 0 \in \hat{K}^n$. Так как $\hat{X}_l \neq 0$, то $\alpha_l = 0$. Таким образом, $\alpha_1 = \ldots = \alpha_l = 0$, и поэтому собственные векторы $\hat{X}_1, \ldots, \hat{X}_l$ линейно независимы. $\square$

Следствие 9.19.10. *Если $A \in \mathrm{M}_n(K)$, характеристический многочлен $p(t) = |A - tE|$ имеет n различных корней $\lambda_1, \ldots, \lambda_n$ в поле K, то матрица A подобна диагональной матрице:*

$$C^{-1}AC = d(\lambda_1, \ldots, \lambda_n), \quad \text{где } C \in \mathrm{GL}_n(K).$$

Теорема 9.19.11. *Матрица $A \in \mathrm{M}_n(\mathbb{C})$ нильпотентна (т. е. $A^m = (0) \in \mathrm{M}_n(K)$ для некоторого $m \in \mathbb{N}$) тогда и только тогда, когда собственные числа $\lambda_1, \ldots, \lambda_n$ равны нулю.*

Доказательство. а) Если $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$, то $|A - \lambda E| = (-1)^n \lambda^n$. По теореме Гамильтона—Кэли $A^n = (0) \in \mathrm{M}_n(\mathbb{C})$.

б) Если $A^m = (0) \in \mathrm{M}_n(\mathbb{C})$ и $A\hat{X} = \lambda\hat{X}$, где $\lambda \in \mathbb{C}$, $0 \neq \hat{X} \in \hat{\mathbb{C}}^n$, то $\hat{\mathbb{C}}^n \ni 0 = A^m\hat{X} = \lambda^n\hat{X}$, следовательно, $\lambda^n = 0$ и $\lambda = 0$. □

Замечание 9.19.12. Одним из фундаментальных результатов об алгебре матриц $\mathrm{M}_n(\mathbb{C})$ над полем комплексных чисел $\mathbb{C}$ (и о строении отдельно взятого линейного оператора конечномерного линейного пространства ${}_{\mathbb{C}}V$) является теорема о *жордановой нормальной форме*:

1) для каждой матрицы $A \in \mathrm{M}_n(\mathbb{C})$ найдётся такая обратимая матрица $C \in \mathrm{GL}_n(\mathbb{C})$, что

$$C^{-1}AC = J_A = \left(\begin{array}{c|c|c|c} J_1 & 0 & \ldots & 0 \\ \hline 0 & J_2 & \ldots & 0 \\ \hline & & \vdots & \\ \hline 0 & 0 & & J_k \end{array}\right) -$$

жорданова матрица (т. е. $J_1, \ldots, J_k$ — жордановы клетки, см. упражнение 8.6.8);

2) нормальная жорданова форма J_A матрицы A определена однозначно (с точностью до порядка жордановых клеток).

Эта теорема обычно является одним из центральных результатов курса линейной алгебры. Она также доказывается в более общем виде в разделе о строении конечнопорождённых модулей над кольцами главных идеалов.

Конечно, теорема Гамильтона—Кэли над полем $\mathbb{C}$ является следствием теоремы о жордановой нормальной форме. В то же время имеются элегантные доказательства теоремы о жордановой нормальной форме, использующие теорему Гамильтона—Кэли.

Мы оставляем этот сюжет для следующих частей наших «начал алгебры» (или его можно рассматривать как достаточно трудную задачу).

Список литературы*

Учебники

[1] Александров П. С. Введение в теорию групп. — М.: Учпедгиз, 1951.

[2] Александров П. С. Курс аналитической геометрии и линейной алгебры. — М.: Наука, 1979.

[3] Андреева Е., Фалина И. Системы счисления и компьютерная арифметика. — М.: Лаборатория Базовых Знаний, 2000.

[4] Аничков Д. С. Теоретическая и практическая арифметика. — М., 1764, 1775, 1786, 1793.

[5] Аничков Д. С. Начальные основания алгебры, или арифметики литеральной. — М., 1781.

[6] Арнольд И. В. Теоретическая арифметика. — М.: Учпедгиз, 1939.

[7] Артин Э. Геометрическая алгебра. — М.: Мир, 1970.

[8] Архангельский А. В. Конечномерные векторные пространства. — М.: Изд-во Моск. ун-та, 1982.

[9] Афанасьев П. А. Арифметика. — М., 1814.

*За 300 лет истории математики и математического образования в России на русском языке накопилось достаточно много интересных и глубоких базовых учебных текстов по алгебре как российских авторов, так и переводов. Привести все эти публикации не представляется возможным в ограниченном объёме. Мы ограничились некоторой выборкой с целью показать связь эпох в преподавании алгебры.

[10] Афанасьев П. А. Алгебра по руководствам Франкера, Лакруа и других новейших математиков. — М., 1816.

[11] Барсов А. Д. Новая алгебра. — М., 1797.

[12] Барсов А. Д. Новейшая арифметика. — М., 1797.

[13] Барти Т., Биркгоф Г. Современная прикладная алгебра. — М.: Мир, 1986.

[14] Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы. — М.: Наука, 1987.

[15] Бахтурин Ю. А. Основные структуры современной алгебры. — М.: Наука, 1990.

[16] Безу Е. Курс математики. — М., 1798, 1801, 1809.

[17] Беклемишев Д. В. Курс аналитической геометрии и линейной алгебры. — М.: Наука, 1984.

[18] Беклемишев Д. В. Дополнительные главы линейной алгебры. — М.: Наука, 1983.

[19] Беллман Р. Введение в теорию матриц. — М.: Наука, 1969.

[20] Бертран Ж. Алгебра. — СПб., 1899, 1901.

[21] Богомолов А. М., Салий В. Н. Алгебраические основы теории дискретных систем. — М.: Наука, 1997.

[22] Боревич З. И. Определители и матрицы. — М.: Наука, 1988.

[23] Бохер М. Введение в высшую алгебру. — М.: Гостехиздат, 1933.

[24] Бугров Я. С., Никольский С. М. Элементы линейной алгебры и аналитической геометрии. — М.: Наука, 1980.

[25] Букреев Б. Я. Элементы теории определителей. — Киев, 1907.

[26] Букреев Б. Я. Элементы алгебраического анализа. — Киев, 1912.

[27] Бурбаки Н. Алгебра. Алгебраические структуры. Линейная и полилинейная алгебра. — М.: Физматгиз, 1962.

[28] Бурбаки Н. Алгебра. Многочлены и поля. Упорядоченные группы. — М.: Физматгиз, 1965.

[29] Бэр Р. Линейная алгебра и проективная геометрия. — М.: ИЛ, 1955.

[30] Бюшгенс С. С. Высшая алгебра. — 1915.

[31] Ван дер Варден Б. Л. Современная алгебра. — М.—Л.: ОНТИ, 1937.

[32] Ван дер Варден Б. Л. Алгебра. — М.: Наука, 1976.

[33] Ващенко-Захарченко М. Е. Теория определителей и теория форм. — Киев, 1877.

[34] Ващенко-Захарченко М. Е. Алгебраический анализ или высшая алгебра. — Киев, 1887.

[35] Вейдлер И. Ф. Арифметика. — М., 1765, 1787, 1795.

[36] Вейдлер И. Ф. Аналитика специоза или алгебра. — М., 1795.

[37] Виленкин Н. Я. Комбинаторика. — М.: Наука, 1969.

[38] Винберг Э. Б. Курс алгебры. — М.: Факториал, 1999, 2001, 2002.

[39] Виноградов С. П. Основы теории детерминантов. — М.: Т-во И. Д. Сытина, 1915; ОНТИ, 1935.

[40] Воеводин В. В. Численные методы алгебры. Теория и алгоритмы. — М.: Наука, 1966.

[41] Воеводин В. В. Линейная алгебра. — М.: Наука, 1974.

[42] Воеводин В. В., Кузнецов В. А. Матрицы и вычисления. — М.: Наука, 1984.

[43] Гантмахер Ф. Р. Теория матриц. — М.: Наука, 1967.

[44] Гельфанд И. М. Лекции по линейной алгебре. — М.: Наука, 1971; МЦНМО, 1998.

[45] Глазман И. М., Любич Ю. И. Конечномерный анализ. — М.: Наука, 1969.

[46] Глухов М. М., Елизаров В. П., Нечаев А. А. Алгебра. — М.: Гелиос АРВ, 2003.

[47] Годунов С. К. Современные аспекты линейной алгебры. — Новосибирск: Научная книга, 1997.

[48] Головина Л. И. Линейная алгебра и некоторые ее приложения. — М.: Наука, 1975.

[49] Голуб Д., Ван Лоун Ч. Матричные вычисления. — М.: Мир, 1999.

[50] Граве Д. А. Курс алгебраического анализа. — Киев, 1910.

[51] Граве Д. А. Элементы высшей алгебры. — Киев, 1914.

[52] Граве Д. А. Начала алгебры. — Петроград, 1915.

[53] Грузинцев А. П. Теория определителей. — Харьков, 1907.

[54] Грэхем Р., Кнут Д., Паташник О. Конкретная математика. Основание информатики. — М.: Мир, 1998.

[55] Деларю Д. Лекции алгебраического анализа. Вып. 1. — Харьков, 1866.

[56] Демидович Б. П., Марон И. А. Основы вычислительной математики. — М.: Физматлит, 1963; Наука, 1970.

[57] Джонсон Ч., Хорн Р. Матричный анализ. — М.: Мир, 1989.

[58] Ефимов Н. В., Розендорн Э. Р. Линейная алгебра и многомерная геометрия. — М.: Наука, 1970.

[59] Зуланке Р., Онищик А. Л. Алгебра и геометрия. Т. 1. Введение. — М.: МЦНМО, 2004.

[60] Ильин В. А., Ким Г. Д. Линейная алгебра и аналитическая геометрия. — М.: Изд-во Моск. ун-та, 1998.

[61] Ильин В. А., Куркина А. В. Высшая математика. — М.: Проспект, 2002.

[62] Ильин В. А., Позняк Э. Г. Линейная алгебра. — М.: Наука, 1974, 1984; Физматлит, 2002.

[63] Каган В. Ф. Основания теории определителей. — Одесса: Гос. изд. Украины, 1922.

[64] Калужнин Л. А. Введение в общую алгебру. — М.: Наука, 1973.

[65] Кемени Д., Снелл Д., Томпсон Д. Введение в конечную математику. — М.: ИЛ, 1963.

[66] Кнут Д. Искусство программирования для ЭВМ. Т. 1—3. — М.: Мир, 1976—1978; Вильямс, 2000.

[67] Кон П. Универсальная алгебра. — М.: Мир, 1968.

[68] Кострикин А. И. Введение в алгебру. Ч. I. Основы алгебры. — М.: Физматлит, 2000.

[69] Кострикин А. И. Введение в алгебру. Ч. II. Линейная алгебра. — М.: Физматлит, 2000.

[70] Кострикин А. И. Введение в алгебру. Ч. III. Основные структуры алгебры. — М.: Физматлит, 2000.

[71] Кострикин А. И., Манин Ю. И. Линейная алгебра и геометрия. — М.: Наука, 1986.

[72] Котельников С. К. Первые основания математических наук части первой, содержащей в себе арифметику. — СПб.: 1766.

[73] Коши О. Л. Алгебраический анализ. — Лейпциг, 1821, 1864.

[74] Кузьмин Р. О., Фаддеев Д. К. Алгебра и арифметика комплексных чисел. — М.: Учпедгиз, 1939.

[75] Куликов Л. Я. Алгебра и теория чисел. — М.: Высшая школа, 1979.

[76] Курант Р., Роббинс Г. Что такое математика? Элементарный очерк идей и методов. — 1947, 1966; М.: МЦНМО, 2001.

[77] Курош А. Г. Курс высшей алгебры. — М.: Наука, 1975.

[78] Курош А. Г. Лекции по общей алгебре. — М.: Наука, 1973.

[79] Лаврентьев М. А., Шабат Б. В. Методы теории функций комплексного переменного. — М.: Физматгиз, 1958.

[80] Лакруа С. Курс математики. — СПб., 1827.

[81] Лакруа С. Основания алгебры. — М., 1837.

[82] Ланкастер П. Теория матриц. — М.: Наука, 1982.

[83] Ларин С. В. Числовые системы. — М.: Академия, 2001.

[84] Лебединцев К. Ф. Курс алгебры. — 1911.

[85] Ленг С. Алгебра. — М.: Мир, 1968.

[86] Лидл Р., Пильц Г. Прикладная абстрактная алгебра. — Екатеринбург: Изд. УрГУ, 1996.

[87] Липский В. Комбинаторика для программистов. — М.: Мир, 1988.

[88] Лобачевский Н. Алгебра или вычисление конечных. — Казань, 1834.

[89] Ляпин Е. С. Курс высшей алгебры. — Учпедгиз, 1953, 1955.

[90] Ляпин Е. С., Евсеев А. Е. Алгебра и теория чисел. I, II. — М.: Просвещение, 1974, 1978.

[91] Магницкий Л. Арифметика. — М., 1703.

[92] Мальцев А. И. Алгебраические системы. — М.: Наука, 1970.

[93] Мальцев А. И. Основы линейной алгебры. — М.: Наука, 1970.

[94] Маргулис Б. Е. Системы линейных уравнений. — М.: Физматгиз, 1960.

[95] Мельников Ю. Б., Мельникова Н. В. Лекции по алгебре. — Екатеринбург: Уральское изд-во, 2003.

[96] Мешков А. Курс высшей алгебры. — СПб., 1862.

[97] Милованов М. В., Тышкевич Р. И., Феденко А. С. Алгебра и аналитическая геометрия. — Минск: ВШ, 1984.

[98] Мишина А. П., Проскуряков И. В. Высшая алгебра. Линейная алгебра, многочлены, общая алгебра. — М.: Физматлит, 1962. — Сер. «Справочная математическая библиотека».

[99] Млодзеевский Б. К. Высшая алгебра. — М., 1895, 1900, 1906.

[100] Млодзеевский Б. К. Теория детерминантов. — М., 1900, 1906.

[101] Млодзеевский Б. К. Курс высшей алгебры. — М., 1911.

[102] Млодзеевский Б. К. Основы высшей алгебры. — М., 1922, 1923.

[103] Муравьев Н. Е. Начальные основания математики. — СПб., 1752.

[104] Мысовских И. П. Лекции по методам вычислений. — М.: Физматгиз, 1962.

[105] Нетто Е. Начала теории определителей. — Одесса, 1912.

[106] Нечаев В. И. Числовые системы. — М.: Просвещение, 1975.

[107] Ньютон И. Всеобщая арифметика, или книга об арифметических синтезе и анализе. — Изд. АН СССР, 1948.

[108] Окунев Л. Я. Основы современной алгебры. — Учпедгиз, 1941.

[109] Окунев Л. Я. Высшая алгебра. — М.—Л.: ОГИЗ, 1944.

[110] Остроградский М. В. Лекции алгебраического и трансцендентного анализа. — СПб., 1837.

[111] Перевощиков Д. М. Основания алгебры. — СПб., 1854.

[112] Петрова В. Т. Лекции по алгебре и геометрии 1, 2. — М.: Владос, 1999.

[113] Понтрягин Л. С. Алгебра. — М.: Наука, 1987.

[114] Привалов И. И. Введение в теорию функций комплексного переменного. — М.: Физматгиз, 1960.

[115] Проскуряков И. В. Числа и многочлены. — М.: Просвещение, 1965.

[116] Родосский К. А. Алгоритм Евклида. — М.: Наука, 1988.

[117] Рублев А. Н. Курс линейной алгебры и аналитической геометрии. — М.: ВШ, 1972.

[118] Румовский С. Сокращения математики, часть первая. — СПб., 1760.

[119] Сачков В. Н. Введение в комбинаторные методы дискретной математики. — М.: Наука, 1982.

[120] Себржинский В. И. Основания алгебры. — М., 1820.

[121] Селиванов Д. Ф. Курс высшей алгебры. — СПб., 1892.

[122] Серре Ж. А. Курс высшей алгебры. — СПб.: Изд. т-ва Вольф, 1883, 1910.

[123] Скорняков Л. А. Элементы алгебры. — М.: Наука, 1980.

[124] Скорняков Л. А. Системы линейных уравнений. — М.: Наука, 1986.

[125] Смирнов В. И. Курс высшей математики. Т. I—V. — М.: Физматгиз, 1958—1959.

[126] Сохоцкий Ю. В. Высшая алгебра. — СПб., 1882, 1911.

[127] Стренг Г. Линейная алгебра и её применения. — М.: Мир, 1980.

[128] Сушкевич А. К. Основы высшей алгебры. — М.: Гостехиздат, 1932, 1937, 1941.

[129] Тихомандрицкий М. А. Краткий курс высшей алгебры. — Харьков, 1887, 1892.

[130] Тышкевич Р. И., Феденко А. С. Линейная алгебра и аналитическая геометрия. — Минск: ВШ, 1968.

[131] Фаддеев Д. К. Лекции по алгебре. — М.: Наука, 1984.

[132] Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. — М.: Наука, 1963.

[133] Федорчук В. В. Курс аналитической геометрии и линейной алгебры. — М.: Изд-во Моск. ун-та, 1990.

[134] Феферман С. Числовые системы. Основания алгебры и анализа. — М.: Наука, 1971.

[135] Франкер Л. Высшая алгебра. — М., 1824.

[136] Франкер Л. Полный курс чистой математики. — СПб., 1827.

[137] Фукс Б. А., Шабат Б. В. Функции комплексного переменного и некоторые их приложения. — М.: Физматгиз, 1959.

[138] Фусс Н. Начальные основание алгебры. — 1821.

[139] Халмош П. Конечномерные векторные пространства. — М.: Физматгиз, 1963.

[140] Холл М. Комбинаторика. — М.: Мир, 1970.

[141] Хорн Р., Джонсон Ч. Матричный анализ. — М.: Мир, 1989.

[142] Чезаро Э. Элементарный учебник алгебраического анализа и исчисления бесконечно малых. Ч. I. — Одесса: Mathesis, 1913; ОНТИ, 1936.

[143] Шапиро Г. М. Высшая алгебра. — Учпедгиз, 1938.

[144] Шафаревич И. Р. Основные понятия алгебры. — М.: ВИНИТИ, 1986; РХД, 1999.

[145] Шевцов Г. С. Линейная алгебра. Теория и прикладные аспекты. — М.: Финансы и статистика, 2003.

[146] Шилов Г. Е. Введение в теорию линейных пространств. — М.—Л.: ГТТИ, 1952.

[147] Шилов Г. Е. Математический анализ. Конечномерные линейные пространства. — М.: Наука, 1969.

[148] Шмидт О. Ю. Абстрактная теория групп. — Киев, 1916; М., 1933.

[149] Шрейер О., Шпернер Е. Введение в линейную алгебру в геометрическом изложении. — М.: ОНТИ, 1934.

[150] Шрейер О., Шпернер Е. Теория матриц. — М.: ОНТИ, 1936.

[151] Эйлер Л. Универсальная арифметика. — СПб., 1768, 1787—1788.

[152] Эйлер Л. Оснований алгебры Эйлера части первой первые три отделения. — СПб., 1812.

[153] Яблонский С. В. Введение в дискретную математику. — М.: Наука, 1986.

[154] Яглом И. М. Комплексные числа. — М.: Физматлит, 1963.

Задачники

[1] Бахвалов Н. С., Лапин А. В., Чижонков Е. В. Численные методы в задачах и упражнениях. — М.: ВШ, 2000.

[2] Беклемишева Л. А., Петрович А. Ю., Чубаров И. А. Сборник задач по аналитической геометрии и линейной алгебре. — М.: Наука, 1987.

[3] Буренин К. П., Малинин А. Ф. Руководство алгебры и собрание алгебраических задач. — М., 1890.

[4] Бутузов В. Ф., Крутицкая И. Д., Шишкин А. А. Линейная алгебра в вопросах и задачах. — М.: Физматлит, 2001, 2003.

[5] Гашков С. Б., Чубариков В. Н. Арифметика, алгоритмы, сложность вычислений. — М.: ВШ, 2000.

[6] Гюнтер Н. М., Кузьмин Р. О. Сборник задач по высшей математике. Т. II. — М.—Л.: ГИТТЛ, 1949.

[7] Журавский А. М. Сборник задач по высшей алгебре. — Л.—М.: ГТТИ, 1933.

[8] Золотаревская Д. И. Сборник задач по линейной алгебре. — М.: УРСС, 2004.

[9] Икрамов Х. Д. Задачник по линейной алгебре. — М.: Наука, 1975.

[10] Ким Г. Д., Крицков Л. В. Алгебра и аналитическая геометрия. Теоремы и задачи. Т. I, II. — М.: Зерцало-М, 2003.

[11] Кострикин А. И. (ред.). Сборник задач по алгебре. — М.: Факториал, 1995; М.: Физматлит, 2001.

[12] Кречмар В. А. Задачи по алгебре. — М.: Наука, 1972.

[13] Куликов Л. Я., Москаленко А. И., Фомин А. А. Сборник задач по алгебре и теории чисел. — М.: Просвещение, 1993.

[14] Лефор Г. Алгебра и анализ. Задачи. — М.: Наука, 1973.

[15] Минорский В. П. Сборник задач по высшей математике. — М.: Физматлит, 1959, 2003.

[16] Окунев Л. Я. Сборник задач по высшей алгебре. — М.: Просвещение, 1964.

[17] Полиа Г., Сеге Г. Задачи и теоремы из анализа. Ч. I, II. — М.: Гостехиздат, 1956; М.: Наука, 1978.

[18] Прасолов В. В. Задачи и теоремы линейной алгебры. — М.: Наука, 1991, 1996.

[19] Проскуряков И. В. Сборник задач по линейной алгебре. — М.: Наука, 1984.

[20] Смирнов Ю. М. (ред.). Сборник задач по аналитической геометрии и линейной алгебре. — М.: Логос, 2005.

[21] Фаддеев Д. К., Соминский И. С. Сборник задач по высшей алгебре. — М.: Наука, 1968, 1977.

[22] Феденко А. С. (ред.). Сборник задач по алгебре и аналитической геометрии. — Минск, 1999.

[23] Шапошников Н. А. Курс алгебры и собрание алгебраических задач. — 1876/77.

Указатель обозначений

Предметный указатель