ЛЕКЦИИ ПО АЛГЕБРЕ

http://vmk.ucoz.net/

Н. Ю. Золотых[1]

3 февраля 2008 г.

[1]Факультет вычислительной математики и кибернетики Нижегородского государственного университета.

Это рабочий вариант записок лекций по курсу геометрии и алгебры, который читает автор на факультете вычислительной математики и кибернетики Нижегородского государственного университета с 1998 г.

Лекции содержат *огромное* количество опечаток. Однако надеюсь, что они (лекции) все равно будут полезны.

Несмотря на то, что за все многочисленные ошибки, неточности и опечатки отвечает автор, на экзамене претензии насчет качества этого документа не принимаются.

Последнюю версию лекций Вы всегда найдете по адресу
http://www.uic.nnov.ru/~zny/algebra/algebra.html

Оглавление http://vmk.ucoz.net/

Глава 1

Счетные и континуальные множества

1.1. Множества

Понятие *множества* вводится с помощью набора аксиом. Мы не будем на них останавливаться, так как это увело бы нас слишком далеко в теорию множеств, а примем так называемую "наивную" точку зрения на понятие множества. Под множеством будем понимать произвольную совокупность некоторых объектов.

Конечные множества можно задать перечислив все его элементы, например: $A = \{a_1, a_2, \ldots, a_n\}$. Под мощностью конечного множества мы будем понимать число элементов в этом множестве. Мощность множества A обозначается $|A|$. Для нашего примера $|A| = n$.

Со школьного курса математики вам известны некоторые *бесконечные* числовые множества: натуральных чисел $\mathbb{N}$, целых чисел $\mathbb{Z}$, рациональных чисел $\mathbb{Q}$, действительных чисел $\mathbb{R}$. Если перенести определение мощности как числа элементов также и на бесконечные множества, то все бесконечные множества окажутся равномощными: все будут иметь "мощность бесконечность". Ниже мы познакомимся с более продуктивным способом определить мощность для бесконечных множеств.

Говорят, что множество B *включено* в множество A и записывают $B \subseteq A$, если любой элемент из B принадлежит также A. В этом случае B называют *подмножеством* множества A. В курсе дискретной математики, вы узнаете, что всего в любом конечном множестве A содержится $2^{|A|}$ его подмножеств (разумеется, считается также пустое множество $\emptyset$ и все множество A). Совокупность всех подмножеств произвольного множества A обычно обозначается 2^A, поэтому справедлива легко запоминаемая формула:

$$\left|2^A\right| = 2^{|A|}. \tag{1.1}$$

Пусть $A = \{1, 2, 3\}$, тогда $2^A = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$. Легко видеть, что $|A| = 3$, $\left|2^A\right| = 2^3 = 8$.

1.2. Отображения

Правило, или закон, по которому каждому элементу x множества A ставится в соответствие единственный элемент y множества B называется *отображением*. Элемент y называется *образом* элемента x, а x — прообразом элемента y. Тот факт, что отображение φ действует из множества A в множество B, обозначают $\varphi : A \to B$. Для отображения φ образ элемента x обозначают φx или $\varphi(x)$.

Если A и B — некоторые подмножества множества $\mathbb{R}$, то отображение называется (числовой) *функцией*. Отображение $\mathbb{N} \to \mathbb{R}$ называется *последовательностью*.

Отображение φ называется *инъективным*, или просто *инъекцией*, если образы любых двух отличных друг от друга элементов x и y из A различны, т.е. $\forall x, y\; x \neq y, x \in A, y \in A\; \varphi x \neq \varphi y$.

Отображение φ называется *сюръективным*, или просто *сюръекцией*, если для произвольного элемента их B найдется по крайней мере один прообраз.

Отображение φ называется *биективным* (*биекцией*), или *взаимно однозначным*, если оно одновременно и инъективно, и сюръективно. Как легко видеть, это условие эквивалентно тому что, любой элемент из B имеет ровно один прообраз.

Упражнение 1.1. Докажите, что для существования инъекции $A \to B$ необходимо и достаточно существование сюръекции $B \to A$.

Заметим, что если существуют инъекция $A \to B$ и сюръекция $A \to B$, то существует и биекция $A \to B$. Это интуитивно ясное утверждение, между тем, имеет достаточно сложное доказательство и мы на нем не останавливаемся.

Пусть множества A и B конечны. Легко видеть, что тогда

- для существования инъекции $A \to B$ необходимо и достаточно, чтобы $|A| \leq |B|$;
- для существования сюръекции $A \to B$ необходимо и достаточно, чтобы $|A| \geq |B|$;
- для существования биекции $A \to B$ необходимо и достаточно, чтобы $|A| = |B|$.

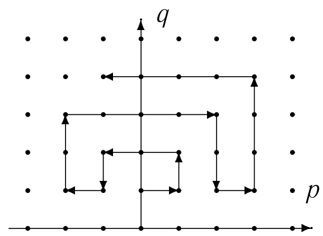
Эти утверждения позволяют теперь дать следующие определения. Для произвольных множеств A и B (в том числе бесконечных) будем говорить, что мощность A меньше или равна мощности B, и записывать $|A| \leq |B|$, если существует инъекция $A \to B$. Будем говорить, что мощность A больше или равна мощности B, и записывать $|A| \geq |B|$, если существует сюръекция $A \to B$. И наконец, будем говорить, что мощности множеств A и B равны (множества A и B равномощны), и записывать $|A| = |B|$, если существует биекция $A \to B$.

Докажем теперь, что $|\mathbb{Z}| = |2\mathbb{Z}|$, где $2\mathbb{Z}$ — множество четных целых чисел. Действительно, биекция $\varphi : \mathbb{Z} \to 2\mathbb{Z}$ задается, например, формулой $\varphi(x) = 2x$. Легко видеть, что любое четное число имеет при этом отображении единственный прообраз. Также нетрудно установить равномощность множеств $\mathbb{Z}$ и $\mathbb{N}$:

$$\begin{array}{ccccccc} \ldots & -2 & -1 & 0 & 1 & 2 & \ldots \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ \ldots & 5 & 3 & 1 & 2 & 4 & \ldots \end{array}$$

В первой строке приведен фрагмент ряда целых чисел. Каждому целому числу x поставлено в соответствие натуральное $\psi(x)$, написанное снизу. Общее правило легко угадывается.

Упражнение 1.2. Привести явную формулу для отображения $\psi(x)$.

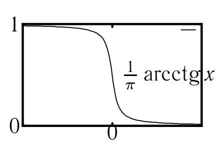


Для построения биективного отображения $\mathbb{Q} \to \mathbb{N}$ рассмотрим кривую, начальный фрагмент которой изображен на рисунке **??**, а дальнейшее поведение легко угадывается. Кривая будет заметать все точки с координатами (p, q), где $p \in \mathbb{Z}$, $q \in \mathbb{N}$. Каждой такой точке поставим в соответствие рациональное число $\dfrac{p}{q}$. Это (вспомогательное) отображение будет сюръективным, но не будет инъективным: например, точкам $(2, 1)$ и $(4, 2)$ будет соответствовать одно и то же число $\dfrac{2}{1} = \dfrac{4}{2}$. Будем двигаться по кривой по направлению, указанному стрелкой, и каждой встретившейся точке (p, q) присваивать порядковый номер $\varphi(p, q)$: $\varphi(0, 1) = 1$, $\varphi(1, 1) = 2$, $\varphi(1, 2) = 3$, $\varphi(0, 2) = 4$, $\varphi(-1, 1) = 5$ и т. д.

Упражнение 1.3. Привести явную формулу для $\varphi(p, q)$.

Для построения биекции $\mathbb{Q} \to \mathbb{N}$ осталось лишь чуть-чуть скорректировать отображение φ: при обходе кривой будем нумеровать лишь точки, для которых НОД$(p, q) = 1$, оставляя без номера все остальные.

Покажем, что множество $\mathbb{R}$ равномощно множеству точек произвольно выбранного на числовой прямой интервала (a, b): равномощность множеств $\mathbb{R}$ и $(0, 1)$ устанавливается, например, взаимно однозначным отображением $\varphi(x) = \dfrac{1}{\pi} \operatorname{arcctg} x$, а равномощность множеств $(0, 1)$ и (a, b) (а так же $[0, 1]$ и $[a, b]$) — биекцией $\psi(x) = a + (b - a)x$ (см. рис. **??**). Функции $\varphi(x)$ и $\psi(x)$ — строго монотонные и, следовательно, как известно из математического анализа, — взаимно однозначные.



Упражнение 1.4. Построить биекцию $[0, 1] \to (0, 1)$.

Подведем итог. Мы рассмотрели несколько бесконечных множеств. Оказалось, что $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$ и $|\mathbb{R}| = |(a, b)|$, однако пока не решен вопрос “$|\mathbb{N}| \overset{?}{=} |\mathbb{R}|$”. Так как

$\mathbb{N} \subseteq \mathbb{R}$, то, конечно, $|\mathbb{N}| \leq |\mathbb{R}|$. Скоро мы увидим, что биекции $\mathbb{N} \to \mathbb{R}$ не существует и, таким образом, $|\mathbb{N}| < |\mathbb{R}|$. Множества, равномощные множеству натуральных чисел называются *счетными*, или множествами мощности $\aleph_0$ (читается ‘алеф-0’). Множества, равномощные множеству действительных чисел называются *континуальными* (мощности континуума), или множествами мощности $\aleph_1$ (читается ‘алеф-1’).

1.3. Свойства счетных множеств

Теорема 1.5. *Любое бесконечное множество A содержит счетное подмножество B.*

Доказательство. Действительно, выберем из множества A произвольный элемент a_1, так как A бесконечно, то мы можем в A выбрать другой элемент $a_2 \neq a_1$, далее выбираем в A элемент a_3, отличный от a_1 и a_2 и т. д. Пусть $B = \{a_1, a_2, a_3, \ldots\}$. Очевидно, B — счетно. □

Теорема 1.6. *Всякое бесконечное подмножество B счетного множества A само счетно.*

Доказательство. Так как A счетно, то мы можем считать, что $A = \{a_1, a_2, a_3, \ldots\}$. Найдем наименьший индекс i_1 (номер в ряду элементов), такой, что $a_{i_1} \in B$, далее выберем наименьший индекс i_2, отличный от i_1, для которого $a_{i_2} \in B$ и т. д. Таким образом, $B = \{a_1, a_2, a_3, \ldots\}$. Очевидно, B — счетно. □

Теорема 1.7. *Объединение конечного или счетного множества счетных множеств — счетно.*

Доказательство. Докажем утверждение для случая счетного множества счетных множеств. Пусть

$$\begin{array}{lcl} A_1 & = & \{a_{11}, a_{12}, a_{13}, \ldots\} \\ A_2 & = & \{a_{21}, a_{22}, a_{23}, \ldots\} \\ A_3 & = & \{a_{31}, a_{32}, a_{33}, \ldots\} \\ \vdots & & \end{array}$$

— семейство счетных множеств. Для установления биекции $A_1 \cup A_2 \cup \ldots \to \mathbb{N}$ необходимо предложить некоторое правило обхода элементов этих множеств, например:

$$\begin{array}{cccccccc}
a_{11} & \rightarrow & a_{12} & & a_{13} & \rightarrow & a_{14} & \dots \\
 & & \downarrow & & \uparrow & & \downarrow & \\
a_{21} & \leftarrow & a_{22} & & a_{23} & & a_{24} & \dots \\
\downarrow & & & & \uparrow & & \downarrow & \\
a_{31} & \rightarrow & a_{32} & \rightarrow & a_{33} & & a_{34} & \dots \\
 & & & & & & \downarrow & \\
a_{31} & \leftarrow & a_{32} & \leftarrow & a_{33} & \leftarrow & a_{44} & \dots \\
\vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

Следуя по стрелке, мы присваиваем элементам множеств A_i порядковые номера и тем самым строим отображение в $\mathbb{N}$. Элементы, встретившиеся неоднократно, конечно, пропускаются (это возможно, так как множества A_i не обязаны не пересекаться). $\square$

1.4. Теорема Кантора

Для произвольного конечного множества A из (1.1) следует неравенство $|A| < \left|2^A\right|$. Оказывается, это неравенство справедливо и для произвольного *бесконечного* множества.

Теорема 1.8 (Кантор). *Для произвольного множества A*

$$|A| < \left|2^A\right|.$$

Доказательство. Неравенство $|A| \leq \left|2^A\right|$ устанавливается сюръекцией $\varphi x = \{x\}$.

Чтобы доказать, что $|A| \neq \left|2^A\right|$, покажем, что допущение о существовании биекции $\psi : A \leq 2^A$ приводит к противоречию. Действительно, пусть ψ — биекция из A в 2^A. Рассмотрим множество

$$I = \{x \in A \mid x \notin \psi\} \tag{1.2}$$

(множество I содержит те и только те элементы, которые не содержатся в своем образе). Так как ψ — биективное отображение, то для любого элемента из 2^A найдется прообраз в A. Пусть для $I \in 2^A$ таким прообразом будет некоторый элемент i: $\psi i = I$. Теперь попробуем решить вопрос '$i \overset{?}{\in} I$'.

1) Если $i \in I$, тогда по (1.2) $i \notin \psi i$. Однако $\psi i = I$, т.е. $i \notin I$.

2) Пусть $i \notin I$, тогда по (1.2) получаем $i \in \psi i$, т.е. $i \in I$.

Оба раза мы приходим к противоречию, следовательно, не верно наше предположение о существовании биекции ψ. $\square$

1.5. Несчетность множества действительных чисел

Основная задача данного раздела — доказать, что $|\mathbb{R}| > |\mathbb{N}|$.

Теорема 1.9 (Несчетность множества $\mathbb{R}$)**.** $\left|2^{\mathbb{N}}\right| = |\mathbb{R}|$.

Доказательство. Сперва установим равномощность множеств $2^{\mathbb{N}}$ и $[0, 1]$. Пусть[1] для любого $A \in 2^{\mathbb{N}}$

$$\varphi(A) = (0, \alpha_1\alpha_2\alpha_3\ldots)_2, \text{ где } \alpha_i = \begin{cases} 0, & \text{если } i \in A, \\ 1, & \text{если } i \notin A \end{cases}.$$

В частности, $\varphi(\emptyset) = 0)$, $\varphi(\mathbb{N}) = 1)$. Очевидно, что φ — сюръекция, однако, следующие примеры:

$$\varphi(\{1, 3\}) = (0, 101(0))_2 = \frac{1}{2} + \frac{1}{8} = \frac{5}{8},$$
$$\varphi(\{1, 4, 5, 6, \ldots\}) = (0, 100(1))_2 = \frac{1}{2} + \frac{1}{16} + \frac{1}{32} + \ldots = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$$

— показывают, что φ не является инъекцией: двум разным подмножествам множества натуральных чисел соответствует одно и то же число, или, что эквивалентно, существует $\alpha \in [0, 1]$, обладающее двумя прообразами. Это возможно, если в двоичной системе счисления α представимо как бесконечная дробь с периодом (1) или (0). Легко видеть, что множество N таких чисел счетно: оно есть объединение счетного числа конечных множеств, состоящих из чисел, период (1) которых начинается с первого, второго и т. д. места после запятой. Счетным является также множество M прообразов всех чисел из N ($M = \left\{x \in 2^{\mathbb{N}} : \varphi(x) \in N\right\}$). Пусть ψ — некоторая биекция из M в N. Скорректируем теперь отображение φ. Для любого $A \in 2^{\mathbb{N}}$ положим

$$\theta(A) = \begin{cases} \psi(A), & \text{если } A \in M, \\ \varphi(A), & \text{если } A \notin M. \end{cases}$$

Легко видеть, что θ — биекция из $2^{\mathbb{N}}$ в $[0, 1]$. Так как $|[0, 1]| = |\mathbb{R}|$, то $\left|2^{\mathbb{N}}\right| = |\mathbb{R}|$. $\square$

Теперь из теоремы Кантора получаем

Утверждение 1.10. $|\mathbb{R}| > |\mathbb{N}|$.

---

[1] Через

$$(0, \alpha_1\alpha_2\alpha_3\ldots)_2 = \frac{\alpha_1}{2} + \frac{\alpha_2}{4} + \frac{\alpha_3}{8} + \ldots + \frac{\alpha_i}{2^i} + \ldots$$

мы обозначили представление действительного числа в двоичной системе счисления.

1.6. Декартово произведение

Декартовым произведением $A_1 \times A_2 \times \ldots \times A_n$ множеств $A_1, A_2, \ldots, A_n$ называется множество всех упорядоченных наборов $(a_1, a_2, \ldots a_n)$, таких, что $a_i \in A$ $(i = 1, 2, \ldots n)$. *Декартовой n-ой степенью* A^n множества A называется множество всех упорядоченных наборов $(a_1, a_2, \ldots a_n)$, таких, что $a_i \in A$ $(i = 1, 2, \ldots n)$.

Упражнение 1.11. Доказать, что $|\mathbb{Z}^n| = |\mathbb{N}|$, $|\mathbb{R}^n| = |\mathbb{R}|$ для любого натурального n.

Глава 2

Векторная алгебра

2.1. Векторы на плоскости и в пространстве

Вектором называется упорядоченная пара точек. Вектор обозначается $\overrightarrow{\mathcal{AB}}$, где $\mathcal{A}$ — точка, называемая *началом* вектора, а $\mathcal{B}$ — точка, называемая его *концом*. Вектор $\overrightarrow{\mathcal{AB}}$ изображается с помощью стрелки, идущей из $\mathcal{A}$ в $\mathcal{B}$, см. рис. 2.1.

Если начало вектора совпадает с его концом, то вектор называется *нулевым*.

Длиной, или *модулем*, вектора $\overrightarrow{\mathcal{AB}}$ называется расстояние между точками $\mathcal{A}$ и $\mathcal{B}$. Длина обозначается $|\overrightarrow{\mathcal{AB}}|$, или $|\mathcal{AB}|$, или просто $\mathcal{AB}$. Нулевой вектор имеет нулевую длину.

Векторы $\overrightarrow{\mathcal{AB}}$ и $\overrightarrow{\mathcal{CD}}$ называются *коллинеарными*, если параллельны (или совпадают) прямые $\mathcal{AB}$ и $\mathcal{CD}$. Нулевой вектор коллинеарен любому другому вектору. Три или более вектора называются *компланарными*, если существует плоскость которой они параллельны.

Два вектора называются *равными*, если они коллинеарны, имеют одинаковое направление и одинаковую длину, см. рис. 2.2. Все нулевые векторы равны друг другу.

Из этого определения следует, что каковы бы ни были вектор $\overrightarrow{\mathcal{AB}}$ и точка $\mathcal{A}'$ найдется, причем единственная, точка $\mathcal{B}'$, такая, что $\overrightarrow{\mathcal{AB}} = \overrightarrow{\mathcal{A}'\mathcal{B}'}$. В этом случае говорят, что вектор $\overrightarrow{\mathcal{AB}}$ отложен из точки $\mathcal{A}'$.

Вектор может обозначаться одиночной буквой (как правило, малой латинской), например, $a = \overrightarrow{\mathcal{AB}}$. Нулевой вектор обозначается o.

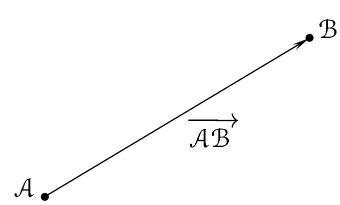


Рис. 2.1: Вектор

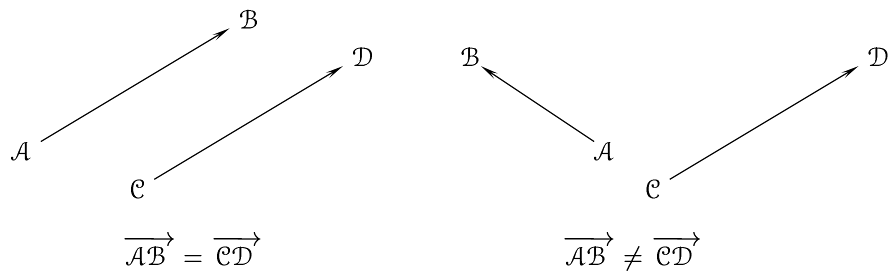


Рис. 2.2: Равные и неравные векторы

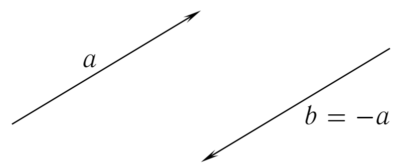


Рис. 2.3: Противоположные векторы

Два коллинеарных вектора, имеющих одинаковую длину и направленных в разные стороны называются *противоположными*. Вектор, противоположный a, обозначается $-a$. Таким образом, $\overrightarrow{\mathcal{AB}} = -\overrightarrow{\mathcal{BA}}$. Вектор, противоположный нулевому вектору, — это сам нулевой вектор: $-o = o$.

2.1.1. Линейные операции

Определим операцию *сложения векторов*. *Суммой* векторов a и b называется вектор c, который можно получить, отложив вектор b из конца вектора a и соединив начало вектора a с концом b (*правило треугольника*). Сумма векторов a и b обозначается $a + b$.

Другой способ найти сумму — использовать так называемое *правило параллелограмма*: векторы a и b откладываются из одной точки и строится паралле-

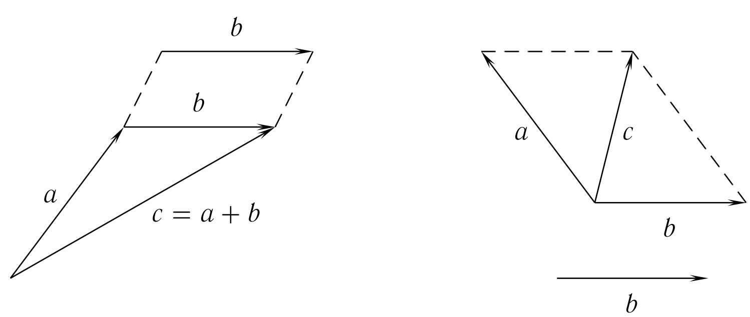


Рис. 2.4: Сумма векторов. Правило треугольника и правило параллелограмма.

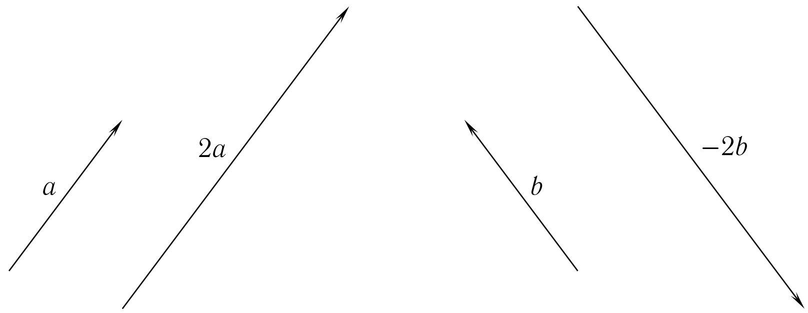


Рис. 2.5:

лограмм (если векторы a и b неколлинеарны), смежные стороны которого суть a и b. Диагональ параллелограмма, начинающаяся из начальной точки векторов a и b очевидно есть сумма векторов a и b.

Определим операцию *умножения вектора на скаляр*. Пусть a — вектор, а α — вещественное число (скаляр). *Произведением вектора b на скаляр α* называется вектор, коллинеарный вектору b, длина которого равна $|\alpha||b|$, а направление совпадает с направлением вектора b, если $\alpha > 0$, и противоположно, если $\alpha < 0$. Произведение вектора b на скаляр α обозначается $\alpha \cdot b$. Знак операции «$\cdot$» часто опускается. Если $a = o$ или $\alpha = 0$, то считают, что $\alpha a = o$.

Операции сложение векторов и умножения вектора на скаляр называются *линейными операциями*. Рассмотрим некоторые свойства этих операций.

Утверждение 2.1 (Свойства линейных операций)**.** *Для произвольных векторов a, b, c и произвольных скаляров α, β справедливо*

1. $a + b = b + a$ (коммутативность сложения)
2. $(a + b) + c = a + (b + c)$ (ассоциативность сложения)
3. $a + o = a$
4. $a + (-a) = o$
5. $1a = a$
6. $\alpha(\beta a) = (\alpha\beta)a$
7. $(\alpha + \beta)a = \alpha a + \beta a$
8. $\alpha(a + b) = \alpha a + \alpha b$

Доказательство. Доказательства свойств 1, 2 понятно из рис. 2.7, 2.6. Остальные свойства очевидны. $\square$

Разумеется, здесь можно было бы упомянуть и о других свойствах линейных операций геометрических векторов, например, $(-1) \cdot a = -a$, $0 \cdot a = o$ и др.

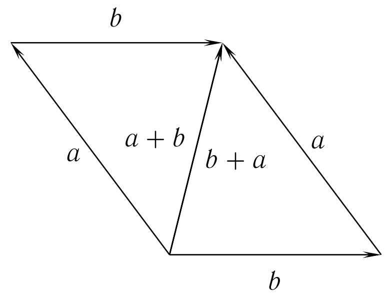


Рис. 2.6: Доказательство коммутативности сложения геометрических векторов

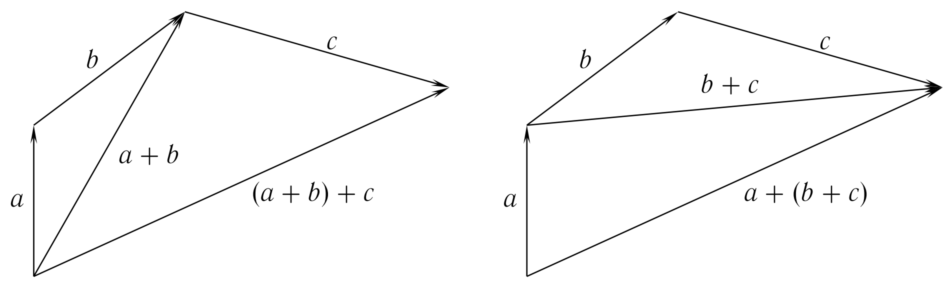


Рис. 2.7: Доказательство ассоциативности сложения геометрических векторов

Почему мы остановились именно на упомянутых выше свойствах, будет ясно из дальнейшего.

Введем операцию *вычитания векторов*. *Разностью* векторов a и b называется такой вектор c, что $c+b=a$. Непосредственно можно проверить, что вектор $c=a+(-b)$ удовлетворяет этому определению ($c+b=a+(-b)+b=a+o=a$) и, следовательно, является разностью. Также легко видеть, что по векторам a и b их разность определяется единственным образом. Действительно, пусть векторы c и c' — оба удовлетворяют определению разности. Тогда $c+b=c'+b$. Прибавляя к обеим частям этого равенства вектор $-b$, получим $c=c'$.

Разность векторов a и b обозначается $a-b$. Итак,

$$a-b=a+(-b),$$

см. рис. 2.8. Если a и b отложены из одной начальной точки, то из рисунка видно, что $a-b$ можно изобразить как вектор, соединяющий конец вектора b с концом вектора a.

Часто удобно все векторы откладывать из одной точки $\mathcal{O}$, называемой в этом случае *полюсом*. Такие векторы называются *радиус-векторами*.

Если $\mathcal{A}$ — некоторая точка, то вектор $\overrightarrow{\mathcal{O}\mathcal{A}}$ называется *радиус-вектором точки* $\mathcal{A}$. Очевидно, что если полюс фиксирован, то между множеством всех радиус-векторов и множеством всех точек (на плоскости или в пространстве) существует взаимнооднозначное соответствие. Этот факт позволяет иногда отождествлять радиус-векторы с соответствующими им точками.

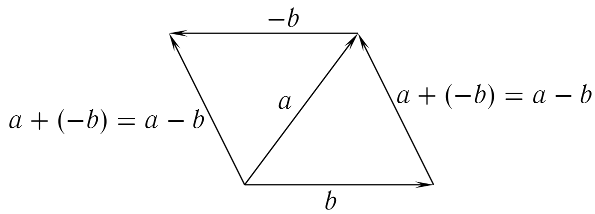


Рис. 2.8: Разность векторов

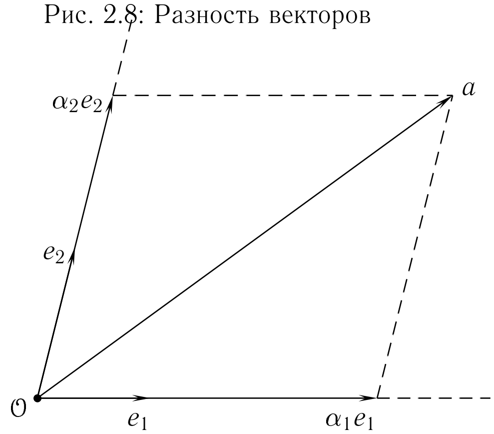


Рис. 2.9:

2.1.2. Базисы и аффинные системы координат

Рассмотрим два некомпланарных вектора на плоскости. Будем говорить, что они образуют *базис* на плоскости.

Теорема 2.2. *Пусть e_1, e_2 образуют базис на плоскости. Тогда для любого вектора a на плоскости существуют, причем единственные, скаляры α_1 и α_2, такие, что*

$$a = \alpha_1 e_1 + \alpha_2 e_2. \tag{2.1}$$

Доказательство. Существование. Отложим векторы из одной точки $\mathcal{O}$, см. рис. 2.9. Через конец вектора a проведем прямую параллельно вектору e_2 до пересечения с прямой, на которой лежит e_1. Аналогично, через конец вектора a проведем прямую параллельно вектору e_1 до пересечения с прямой, на которой лежит e_2. Получаем, что $a = \alpha_1 e_1 + \alpha_2 e_2$ для некоторых чисел α_1, α_2.
Единственность. Предположим, что имеется два разложения:

$$a = \alpha_1 e_1 + \alpha_2 e_2 = \beta_1 e_1 + \beta_2 e_2,$$

откуда $(\alpha_1 - \beta_1)e_1 + (\alpha_2 - \beta_2)e_2 = o$. Если $\alpha_1 \neq \beta_1$, то

$$e_1 = -\frac{\alpha_2 - \beta_2}{\alpha_1 - \beta_1} e_2,$$

т. е. векторы e_1 и e_2 коллинеарны, что невозможно, так как они образуют базис. К аналогичному выводу приходим, предположив, что $\alpha_2 \neq \beta_2$. Следовательно, $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, т. е. разложение по базису единственно. □

Равенство (2.1) называется *разложением вектора a по базису e_1, e_2*, а скаляры α_1 и α_2 — *координатами вектора a в базисе e_1, e_2*.

Тот факт, что координаты вектора a в базисе e_1, e_2 суть α_1 и α_2, будем записывать следующим образом:

$$a(\alpha_1, \alpha_2) \quad \text{или} \quad [a] = (\alpha_1, \alpha_2)$$

или (чтобы уточнить, в каком базисе заданы координаты)

$$[a]_{e_1,e_2} = (\alpha_1, \alpha_2).$$

Теорема 2.3. *Координаты суммы векторов равны сумме соответствующих координат: если $[a] = (\alpha_1, \alpha_2)$, $[b] = (\beta_1, \beta_2)$, то $[a+b] = (\alpha_1+\beta_1,\ \alpha_2+\beta_2)$. Координаты произведения вектора на число равны произведению координат на это число: если $[a] = (\alpha_1, \alpha_2)$, то $[\beta a] = (\beta\alpha_1,\ \beta\alpha_2)$.*

Доказательство. Найдем сумму векторов $a = \alpha_1 e_1 + \alpha_2$ и $b = \beta_1 e_1 + \beta_2 e_2$. Имеем $a+b = (\alpha_1+\beta_1)e_1+(\alpha_2+\beta_2)e_2$. Следовательно, $[a+b] = (\alpha_1+\beta_1,\ \alpha_2+\beta_2)$.

Умножим вектор $a = \alpha_1 e_1 + \alpha_2$ на скаляр β. Получаем $a = \beta\alpha_1 e_1 + \beta\alpha_2 e_2$. Следовательно, $[\beta a] = (\beta\alpha_1,\ \beta\alpha_2)$. □

Совокупность некоторой фиксированной точки $\mathcal{O}$ (полюса) и базиса e_1, e_2 на плоскости называется (*аффинной*) *системой координат*. При этом точка $\mathcal{O}$ называется *началом* системы координат, а прямые, проходящие через $\mathcal{O}$ параллельно базисным векторам, — ее *осями*. На каждой из осей задается *направление*, определяемое направлением соответствующего базисного вектора. Оси системы координат на плоскости традиционно называются *осью абсцисс* (или осью $\mathcal{O}x$) и *осью ординат* (или осью $\mathcal{O}y$).

Координатами точки $\mathcal{A}$ в некоторой системе координат $\mathcal{O}$, e_1, e_2 называются координаты ее радиус-вектора:

$$[\mathcal{A}] = [\,\overrightarrow{\mathcal{O}\mathcal{A}}\,]$$

Имеем $\overrightarrow{\mathcal{A}\mathcal{B}} = \overrightarrow{\mathcal{O}\mathcal{A}} - \overrightarrow{\mathcal{O}\mathcal{B}}$, поэтому *координаты вектора равны координатам его конца минус координаты начала.*

2.1.3. Ортонормированный базис и декартова система координат

Вектор длины 1 называется *единичным вектором*. Легко видеть, что $\dfrac{a}{|a|}$ есть единичный вектор, коллинеарный вектору a и сонаправленный с ним.

Углом между векторами $\overrightarrow{\mathcal{O}\mathcal{A}}$, $\overrightarrow{\mathcal{O}\mathcal{B}}$ называется величина угла $\mathcal{A}\mathcal{O}\mathcal{B}$. Говорят, что векторы a и b *ортогональны*, или *перпендикулярны*, есди угол между a и b равен $\pi/2$. Обозначение для ортогональных векторов: $a \perp b$. Считают, что нулевой вектор ортогонален любому вектору.

Базис e_1, e_2 называется *ортогональным*, если векторы e_1 и e_2 ортогональны. Базис e_1, e_2 называется *ортонормированным*, если векторы e_1 и e_2 суть ортогональные единичные векторы. Если базис e_1, e_2 ортонормированный, то система

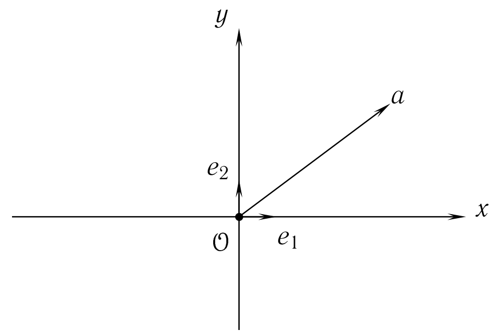


Рис. 2.10: Прямоугольная система координат на плоскости

координат $\mathcal{O}$, e_1, e_2 называется *прямоугольной*, или *декартовой* системой координат; см. рис. 2.10.

Рассмотрим три некомпланарных вектора e_1, e_2, e_3 в пространстве. Будем говорить, что они образуют *базис* в пространстве.

Теорема 2.4. *Пусть векторы e_1, e_2, e_3 образуют базис в пространстве. Тогда для любого вектора a найдутся, причем единственные, числа α_1, α_2 и α_3, такие, что*

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3. \tag{2.2}$$

Доказательство. Существование. Отложим векторы e_1, e_2, e_3, a из одной точки $\mathcal{O}$; см. рис. 2.11. Проведем через конец вектора a прямую, параллельную вектору e_3. Прямая пересечет плоскость, в которой лежат векторы e_1 и e_2, в некоторой точке. Пусть b — ее радиус-вектор. Так как векторы e_1 и e_2 образуют базис в этой плоскости, то по теореме 2.2 вектор b выражается через них, т. е. $b = \alpha_1 e_1 + \alpha_2 e_2$ для некоторых чисел α_1, α_2. Далее, проведем через конец вектора a плоскость параллельную векторам e_1 и e_2. Эта плоскость пересечет прямую, на которой лежит вектор e_3 в некоторой точке. Ее радиус-вектор, очевидно, коллинеарен вектору e_3 и поэтому равен $\alpha_3 e_3$ для некоторого α_3. Теперь имеем

$$a = b + \alpha_3 e_3 = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3.$$

Единственность доказыватся аналогично доказательству единственности из теоремы 2.2 и вытекает из некопланарности базисных векторов. □

Представление (2.2) называется *разложением вектора a по базису e_1, e_2, e_3*. Числа α_1, α_2 и α_3 в равенстве $a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$ называются *координатами вектора a в базисе e_1, e_2, e_3*.

Для координат вектора a в базисе e_1, e_2, e_3 будем использовать следующее обозначение:

$$a(\alpha_1, \alpha_2, \alpha_3) \quad \text{или} \quad [a] = (\alpha_1, \alpha_2, \alpha_3)$$

или (когда нужно явно указать базис)

$$[a]_{e_1,e_2,e_3} = (\alpha_1, \alpha_2, \alpha_3).$$

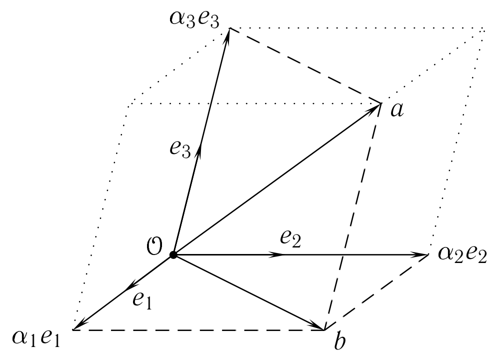


Рис. 2.11: Разложение вектора по базису в пространстве

Совокупность точки $\mathcal{O}$ и базиса e_1, e_2, e_3 называется (*аффиннной*) *системой координат*. В этом случае $\mathcal{O}$ называется *началом системы координат*, а прямые, проходящие через $\mathcal{O}$ параллельно базисным векторам, — ее *осями*. Оси системы координат в пространстве традиционно называются *осью абсцисс* (или осью Ox), *осью ординат* (или осью O), и *осью аппликат* (или осью O) соответственно. На осях выбирается направление согласно направлению базисных векторов.

Координатами точки $\mathcal{A}$ в некоторой системе координат $\mathcal{O}$, e_1, e_2, e_3 называются координаты ее радиус-вектора:

$$[\mathcal{A}] = [\,\overrightarrow{\mathcal{O}\mathcal{A}}\,]$$

Как и для векторов на плоскости, доказывается следующее утверждение.

Теорема 2.5. *Если* $[a] = (\alpha_1, \alpha_2, \alpha_3)$, $[b] = (\beta_1, \beta_2, \beta_3)$, *то* $[a+b] = (\alpha_1+\beta_1,\ \alpha_2+\beta_2,\ \alpha_3+\beta_3)$ *и* $[\beta a] = (\beta\alpha_1,\ \beta\alpha_2,\ \beta\alpha_3)$.

Базис e_1, e_2, e_3 называется *ортогональным*, если $e_1 \perp e_2$, $e_1 \perp e_3$, $e_2 \perp e_3$. Ортогональный базис e_1, e_2, e_3 называется *ортонормированным*, если векторы e_1, e_2, e_3 имеют единичную длину. Если базис ортонормированный, то соответствующая система координат называется *прямоугольной* (или *декартовой*).

2.2. Скалярное, векторное и смешанное произведения

2.2.1. Скалярное произведение

Скалярным (или *внутренним*) *произведением* векторов a и b называется

$$(a, b) = |a| \cdot |b| \cdot \cos\vartheta,$$

где ϑ — угол между a и b.

Теорема 2.6. *Для любых векторов* a, b, c *на плоскости или в пространстве и любого скаляра* α *справедливо*

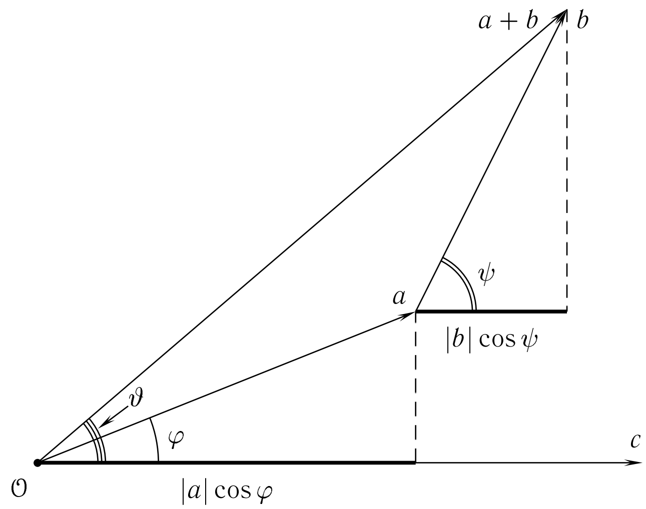


Рис. 2.12:

1. $(a, b) = (b, a)$ (*симметричность*)

2. $(a + b, c) = (a, c) + (b, c)$

3. $(\alpha a, b) = \alpha(a, b)$

4a. $(a, a) \geq 0$

4b. $(a, a) = 0$ *тогда и только тогда, когда* $a = o$

5a. $(a, b) = 0$ *тогда и только тогда, когда* a *и* b *ортогональны*

5b. $(a, b) > 0$ *тогда и только тогда, когда угол между* a *и* b *острый*

5c. $(a, b) < 0$ *тогда и только тогда, когда угол между* a *и* b *тупой*

Доказательство. Свойства 1, 3–5 очевидны.

Докажем свойство 2. В левой части доказываемого равенства имеем $(a + b, c) = |a + b| \cdot |c| \cdot \cos\vartheta$, где ϑ — угол между $a + b$ и c. В правой части — $(a, c) + (b, c) = |a| \cdot |c| \cdot \cos\varphi + |b| \cdot |c| \cdot \cos\psi = (|a|\cos\varphi + |b|\cos\psi)|c|$, где φ — угол между a и c, а ψ — угол между b и c. Таким образом, обе части содержат одинаковый множитель $|c|$. Сравним оставшиеся множители. Из рис. 2.12, на котором углы φ и ψ изображены острыми, видно, что

$$|a|\cos\varphi + |b|\cos\psi = |a + b|\cos\vartheta$$

Аналогично рассматриваются случаи, когда один из углов, φ или ψ, тупой и когда оба угла тупые. □

Свойства 2 и 3 называются *свойствами линейности* скалярного произведения (по первому аргументу). Из свойства 1 (симметричность) следует также линейность по второму аргументу:

$$(a, b + c) = (a, b) + (a, c), \qquad (a, \alpha b) = \alpha(a, b).$$

Пусть e_1, e_2, e_3 — ортонормированный базис, тогда, легко видеть,

$$(e_1, e_1) = (e_2, e_2) = (e_3, e_3) = 1, \qquad (e_1, e_2) = (e_1, e_3) = (e_2, e_3) = 0. \tag{2.3}$$

Теорема 2.7. *Пусть e_1, e_2 — ортонормированный базис на плоскости и $[a] = (\alpha_1, \alpha_2)$, $[b] = (\beta_1, \beta_2)$, тогда*

$$(a, b) = \alpha_1\beta_1 + \alpha_2\beta_2.$$

Пусть e_1, e_2, e_3 — ортонормированный базис в пространстве и $[a] = (\alpha_1, \alpha_2, \alpha_3)$, $[b] = (\beta_1, \beta_2, \beta_3)$, тогда

$$(a, b) = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3.$$

Доказательство. Имеем $(a, b) = (\alpha_1 e_1 + \alpha_2 e_2, \beta_1 e_1 + \beta_2 e_2)$, откуда, пользуясь свойствами линейности скалярного произведения и формулами (2.3), получаем

$$(a, b) = \alpha_1\beta_1(e_1, e_1) + \alpha_1\beta_2(e_1, e_2) + \alpha_2\beta_1(e_2, e_1) + \alpha_2\beta_2(e_2, e_2) = \alpha_1\beta_1 + \alpha_2\beta_2$$

Доказательство для случая векторов в пространстве аналогично. □

Следствие 2.8. *Пусть e_1, e_2 — ортонормированный базис на плоскости и $[a] = (\alpha_1, \alpha_2)$ тогда*

$$|a| = \sqrt{\alpha_1^2 + \alpha_2^2}.$$

Пусть e_1, e_2, e_3 — ортонормированный базис в пространстве и $[a] = (\alpha_1, \alpha_2, \alpha_3)$ тогда

$$|a| = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}.$$

Следствие 2.9. *Пусть e_1, e_2 — ортонормированный базис на плоскости и $[A] = (\alpha_1, \alpha_2)$, $[B] = (\beta_1, \beta_2)$, тогда*

$$|AB| = \sqrt{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2}.$$

Пусть e_1, e_2, e_3 — ортонормированный базис в пространстве и $[A] = (\alpha_1, \alpha_2, \alpha_3)$, $[B] = (\beta_1, \beta_2, \beta_3)$, тогда

$$|AB| = \sqrt{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 + (\alpha_3 - \beta_3)^2}.$$

Скалярное произведение удобно использовать для вычисления углов между векторами. Так как $(a, b) = |a| \cdot |b| \cdot \cos\vartheta$, где ϑ — угол между a и b, то

$$\cos\vartheta = \frac{(a, b)}{|a| \cdot |b|}.$$

Отсюда, в частности, получаем

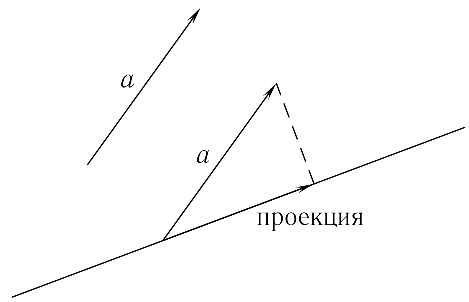


Рис. 2.13: Ортогональная проекция вектора на прямую

Следствие 2.10. *Пусть* e_1, e_2, e_3 *— ортонормированный базис в пространстве,* $[a] = (\alpha_1, \alpha_2, \alpha_3)$, $[b] = (\beta_1, \beta_2, \beta_3)$ *и* ϑ *— угол между* a *и* b, *тогда*

$$\cos\vartheta = \frac{\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3}{\sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2} \cdot \sqrt{\beta_1^2 + \beta_2^2 + \beta_3^2}}.$$

Аналогичное утверждение про векторы на плоскости.

2.2.2. Ортогональная проекция

Ортогональной проекцией вектора a на прямую называется вектор, который можно получить следующим образом. Через начало и конец вектора a опускаются перпендикуляры на прямую. Точки пересечения этих перпендикуляров с прямой являются соответственно началом и концом проекции. Разумеется, вектор a можно сразу расположить так, чтобы его начало лежало на прямой. Тогда достаточно провести перпендикуляр лишь через его конец; см. рис. 2.13.

Пусть b — некоторый ненулевой вектор, параллельный прямой (*направляющий вектор прямой*). Тогда проекция вектора a на прямую, определяемую направлением b, равна (см. рис. 2.14)

$$\mathrm{pr}_v\, a = |a| \cdot \cos\vartheta \cdot \frac{b}{|b|} = |a| \cdot |b| \cdot \cos\vartheta \cdot \frac{b}{|b|^2} = \frac{(a,b)}{(b,b)} b.$$

Утверждение 2.11. *Пусть* e_1, e_2, e_3 *— ортонормированный базис,* a *— вектор, причем* $[a] = (\alpha_1, \alpha_2, \alpha_3)$, *тогда*

$$\alpha_1 = (a, e_1), \quad \alpha_2 = (a, e_2), \quad \alpha_3 = (a, e_3).$$

Аналогичное утверждение для векторов на плоскости.

Доказательство. В ортонормированном базисе координаты, очевидно, равны $\alpha_j = |a| \cos\varphi_j$, где φ_j — угол между векторами a и e_j, откуда $\alpha_j = (a, e_j)$ $(j = 1, 2, 3)$.

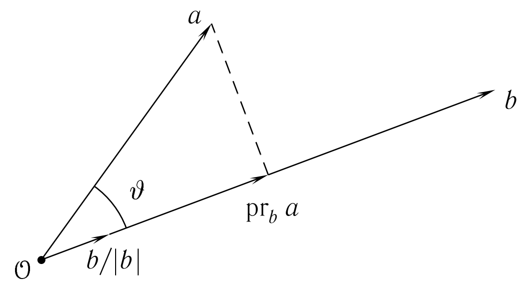


Рис. 2.14: Ортогональная проекция вектора на прямую

Можно дать другое доказательство. Имеем

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3.$$

Домножая обе части этого равенства скалярно на e_j и пользуясь линейностью скалярного произведения, получаем

$$(a, e_j) = \alpha_1(e_1, e_j) + \alpha_2(e_2, e_j) + \alpha_3(e_3, e_j).$$

Так как $(e_i, e_j) = 0$ при $i \neq j$ и $(e_j, e_j) = 1$, то из трех слагаемых в правой части остается только одно: $(a, e_j) = \alpha_j$. $\square$

Пусть e_1, e_2, e_3 — ортонормированный базис и, как и в доказательстве теоремы, φ_j — угол между векторами a и e_j $(j = 1, 2, 3)$. Величины $\cos\varphi_1$, $\cos\varphi_2$, $\cos\varphi_3$ называются *направляющими косинусами*.

Упражнение 2.12. Доказать, что направляющие косинусы удовлетворяют равенству

$$\cos^2\varphi_1 + \cos^2\varphi_2 + \cos^2\varphi_3 = 1.$$

Аналогичное утверждение для векторов на плоскости.

2.2.3. Ориентация векторов в пространстве

Упорядоченная тройка некомпланарных векторов a, b, c называется *правой*, если после того, как отложить все три вектора из одной точки, кратчайший поворот от a к b при наблюдении из конца вектора c происходит в направлении против часовой стрелки; см. рис. 2.15. В противном случае тройка некомпланарных векторов называется *левой*.

Легко видеть, что если тройка a, b, c — правая, то тройки b, c, a, c, a, b — также правые, а тройки b, a, c, a, c, b, c, b, a — левые. Если тройка a, b, c — левая, то тройки b, c, a, c, a, b — также левые, а тройки b, a, c, a, c, b, c, b, a — правые.

Базис e_1, e_2, e_3 называется *правым*, если тройка e_1, e_2, e_3 правая, и *левым* в противном случае.

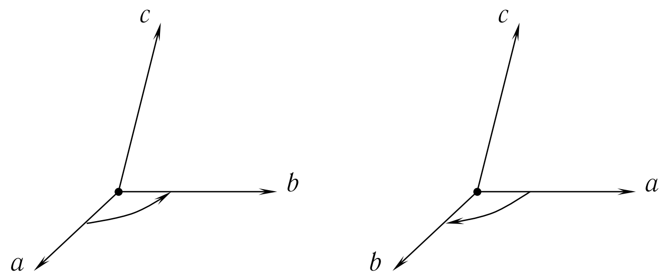


Рис. 2.15: Правая и левая тройки векторов a, b, c

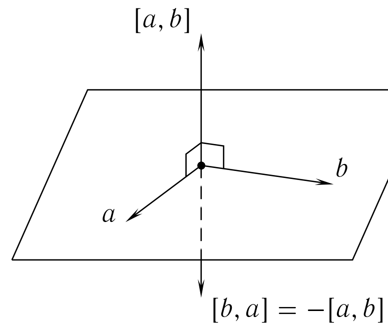


Рис. 2.16: Векторное произведение

2.2.4. Векторное и смешанное произведения

Пусть a и b — неколлинеарные векторы в пространстве. Построим вектор c, такой, что

1. $|c| = |a| \cdot |b| \cdot \sin\vartheta$, где ϑ — угол между a и b;
2. $c \perp a$, $c \perp b$;
3. тройка a, b, c — правая.

Очевидно, что три приведенных правила по заданным a и b однозначно определяют вектор c. Если a и b коллинеарны, то пусть $c = o$. Вектор c называется *векторным* (или *внешним*) произведением векторов a и b; см. рис. 2.16. Обозначение: $c = [b, b]$.

Из определения сразу следует, что модуль векторного произведения двух неколлинеарных векторов равен площади параллелограмма, построенного на этих векторах, если их отложить из одной точки; см. рис. 2.17.

Смешанным произведением трех векторов a, b, c в пространстве называется

$$(a, b, c) = ([a, b], c).$$

Теорема 2.13. *Пусть a, b, c — некомпланарные векторы в пространстве. Абсолютное значение их смешанного произведения равно объему парал-*

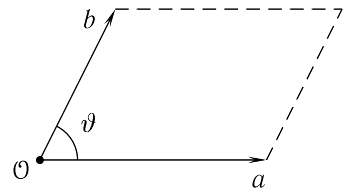


Рис. 2.17: Параллелограмм, построенный на векторах a, b

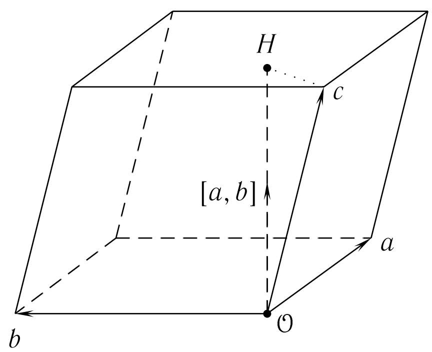


Рис. 2.18: Абсолютное значение смешанного произведения есть объем параллелепипеда, построенного на заданных векторах.

лелепипеда, построенного на векторах a, b, c, если их отложить из одной точки.

Доказательство. Объем параллелепипеда, построенного на векторах a, b, c равен $V = Sh$, где S — площадь основания, h — высота; см. рис. 2.18. Основанием является параллелограмм, построенный на векторах a, b, поэтому $S = \big|[a,b]\big|$. Высоту можно представить как проекцию вектора c на прямую, перпендикулярную плоскости, в которой лежат векторы a и b. Эта прямая коллинеарна вектору $[a,b]$, поэтому $h = OH = \big|\mathrm{pr}_{[a,b]}\, a\big|$. Теперь имеем

$$V = S \cdot h = \big|[a,b]\big| \cdot OH = \big|[a,b]\big| \cdot |c| \cdot |\cos\vartheta| = \big|([a,b],\ c)\big|,$$

где ϑ — угол между векторами $[a,b]$ и c. □

Теорема 2.14 (Свойства смешанного произведения)**.** *Для любых векторов a, b, c, c' и любого скаляра α справедливо*

1. $(a,b,c) = (b,c,a) = (c,a,b) = -(a,c,b) = -(b,a,c) = -(c,b,a)$
2. $(a,b,c+c') = (a,b,c) + (a,b,c')$
3. $(a,b,\alpha c) = \alpha(a,b,c)$

4a. $(a,b,c) = 0$ *тогда и только тогда, когда векторы a, b, c компланарны*

4b. $(a, b, c) > 0$ *тогда и только тогда, когда тройка* a, b, c *правая*

4c. $(a, b, c) < 0$ *тогда и только тогда, когда тройка* a, b, c *левая*

Доказательство. Докажем свойство 1. Оно очевидно, если векторы a, b, c компланарны. Пусть теперь векторы a, b, c не компланарны. Параллелепипед, построенный на этих векторах, не изменится, если векторы рассматривать в другом порядке. Следовательно, не изменится и абсолютная величина смешанного произведения, однако может измениться знак. Однако знак зависит от ориентации рассматриваемой тройки векторов.

Свойства 2, 3 следуют из соответсвующих свойств скалярного произведения:

$$(a, b, c + c') = ([a, b], c + c') = ([a, b], c) + ([a, b], c') = (a, b, c) + (a, b, c'),$$

$$(a, b, \alpha c) = ([a, b], \alpha c]) = \alpha([a, b], c) = \alpha(a, b, c).$$

Свойства 4a, 4b, 4c сразу вытекают из определения. □

Свойства 2 и 3 называются свойствами *линейности* смешанного произведения (по третьему аргументу). Из свойства 1 очевидным образом вытекает линейность по первому и второму аргументам.

Теорема 2.15 (Свойства векторного произведения). *Для любых векторов* a, b, c *в пространстве и любого скаляра* α

1. $[a, b] = -[b, a]$ (*антикоммутативность*)
2. $[a + b, c] = [a, c] + [b, c]$
3. $[\alpha a, b] = \alpha[a, b]$
4. $[a, b] = o$ *тогда и только тогда, когда* a *и* b *коллинеарны*

Доказательство. Свойства 1, 3–5 элементарным образом следуют из определения. Докажем свойство 2. Для этого рассмотрим вектор

$$d = [a + b, c] - [a, c] - [b, c]$$

и покажем, что он нулевой. Действительно, используя линейность скалярного и смешанного произведений, получаем

$$(d, d) = (d, [a + b, c] - [a, c] - [b, c]) = (d, a + b, c) - (d, a, c) - (d, b, c) = 0,$$

откуда $d = o$. □

2.2.5. Выражение векторного и смешанного произведений через координаты сомножителей

Пусть векторы e_1, e_2, e_3 образуют правый ортонормированный базис. Нетрудно проверить, что

$$\begin{array}{lll} [e_1, e_2] = e_3, & [e_2, e_3] = e_1, & [e_3, e_1] = e_2, \\ [e_2, e_1] = -e_3, & [e_3, e_2] = -e_1, & [e_1, e_3] = -e_2, \\ [e_1, e_1] = o, & [e_2, e_2] = o, & [e_3, e_3] = o. \end{array}$$

Теорема 2.16. *Пусть e_1, e_2, e_3 — правый ортонормированный базис, в котором $[a] = (\alpha_1, \alpha_2, \alpha_3)$, $[b] = (\beta_1, \beta_2, \beta_3)$, тогда*

$$[a, b] = (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)e_3.$$

Доказательство. Так как $a = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$ и $b = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3$, то

$$[a, b] = [\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3, \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3] =$$

$$\begin{aligned} &= [\alpha_1 e_1, \beta_1 e_1] + [\alpha_1 e_1, \beta_2 e_2] + [\alpha_1 e_1, \beta_3 e_3] + \\ &+ [\alpha_2 e_2, \beta_1 e_1] + [\alpha_2 e_2, \beta_2 e_2] + [\alpha_2 e_2, \beta_3 e_3] + \\ &+ [\alpha_3 e_3, \beta_1 e_1] + [\alpha_3 e_3, \beta_2 e_2] + [\alpha_3 e_3, \beta_3 e_3] = \end{aligned}$$

$$= 0 + \alpha_1\beta_2 e_3 - \alpha_1\beta_3 e_2 - \alpha_2\beta_1 e_3 + 0 + \alpha_2\beta_3 e_1 + \alpha_3\beta_1 e_2 - \alpha_3\beta_2 e_1 + 0 =$$

$$= (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)e_3.$$

□

Теорема 2.17. *Пусть e_1, e_2, e_3 — базис в пространстве*

$$[a] = (\alpha_1, \alpha_2, \alpha_3), \quad [b] = (\beta_1, \beta_2, \beta_3), \quad [c] = (\gamma_1, \gamma_2, \gamma_3),$$

тогда

$$(a, b, c) = (\alpha_1\beta_2\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1 - \alpha_2\beta_1\gamma_3 - \alpha_1\beta_3\gamma_2) \cdot (e_1, e_2, e_3).$$

Если базис правый ортонормированный, то

$$(a, b, c) = \alpha_1\beta_2\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3.$$

Доказательство. Аналогично доказательству теоремы для выражения скалярного произведения через координаты векторов. □

2.2.6. Определители второго и третьего порядка

Рассмотрим 4 числа α_1, α_2, β_1, β_2. Из них можно составить таблицу

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix},$$

называемую *матрицей второго порядка*. По этой матрице можно вычислить число $\alpha_1\beta_2 - \alpha_2\beta_1$, называемое *определителем*, или *детерминантом*, второго порядка. Определитель обозначается

$$\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}$$

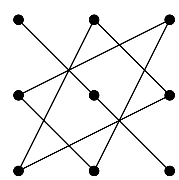
Матрицей третьего порядка называется таблица

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix}.$$

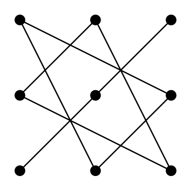
Ее определителем называется число

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} = \alpha_1\beta_2\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3.$$

Обратим внимание, что обе формулы: для определителя второго порядка и для определителя третьего порядка — представляют собой алгебраические суммы произведений элементов матриц, причем каждое произведение составлено из элементов матрицы, которые берутся из разных строк и разных столбцов. Для определения знака, с которым входят члены определителя третьего порядка удобно использовать следующие диаграммы:



$+$



$-$

Легко проверить, что

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} = \alpha_1 \cdot \begin{vmatrix} \beta_2 & \beta_3 \\ \gamma_2 & \gamma_3 \end{vmatrix} - \alpha_2 \cdot \begin{vmatrix} \beta_1 & \beta_3 \\ \gamma_1 & \gamma_3 \end{vmatrix} + \alpha_3 \cdot \begin{vmatrix} \beta_1 & \beta_2 \\ \gamma_1 & \gamma_2 \end{vmatrix}.$$

Используя определители, можно компактно записать формулы для выражения векторного и смешанного произведений через координаты векторов.

Следствие 2.18. *Пусть e_1, e_2, e_3 — правый ортонормированный базис и*

$$[a] = (\alpha_1, \alpha_2, \alpha_3), \quad [b] = (\beta_1, \beta_2, \beta_3),$$

тогда

$$[a, b] = e_1 \begin{vmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{vmatrix} - e_2 \begin{vmatrix} \alpha_1 & \alpha_3 \\ \beta_1 & \beta_3 \end{vmatrix} + e_3 \begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix} = \begin{vmatrix} e_1 & e_2 & e_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix}.$$

Следствие 2.19. *Пусть e_1, e_2, e_3 — базис в пространстве и*

$$[a] = (\alpha_1, \alpha_2, \alpha_3), \quad [b] = (\beta_1, \beta_2, \beta_3), \quad [c] = (\gamma_1, \gamma_2, \gamma_3),$$

тогда

$$(a, b, c) = \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} \cdot (e_1, e_2, e_3).$$

В частности, если базис правый ортонормированный, то смешанное произведение равно определителю, составленному из координат векторов.

Следствие 2.20 (Геометрический смысл определителя третьего порядка)**.** *Абсолютная величина определителя, составленного из координат трех векторов в ортонормированном базисе, равна объему параллелепипеда, построенного на этих векторах.*

Следствие 2.21 (Критерий компланарности векторов)**.** *Векторы a, b, c компланарны тогда и только тогда, когда определитель, составленный из их координат* (*в любом базисе*), *равен нулю.*

Оказывается, аналогичные утверждения можно доказать и для определителей второго порядка.

Утверждение 2.22 (Геометрический смысл определителя второго порядка)**.** *Абсолютная величина определителя, составленного из координат двух векторов на плоскости в ортонормированном базисе, равна площади параллелограмма, построенного на этих векторах.*

Доказательство. Пусть e_1, e_2 — ортонормированный базис, в котором векторы a и b имеют координаты (α_1, α_2), (β_1, β_2) соответственно. Можно считать, что плоскость, в которой располагаются векторы, находится в пространстве и e_3 — единичный вектор, ортогональный этой плоскости. Тогда e_1, e_2, e_3 — ортонормированный базис в пространстве. В нем векторы a и b имеют координаты $(\alpha_1, \alpha_2, 0)$, $(\beta_1, \beta_2, 0)$ соответственно. Имеем

$$[a, b] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \alpha_1 & \alpha_2 & 0 \\ \beta_1 & \beta_2 & 0 \end{vmatrix} = e_3 \begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}.$$

Площадь параллелограмма, построенного на векторах a, b, равна модулю этого векторного произведения, т. е. абсолютному значению определителя

$$\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}.$$

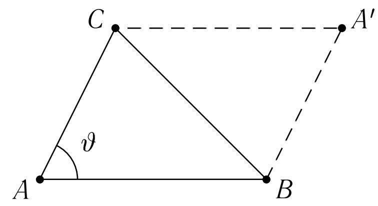


Рис. 2.19:

Заметим, что можно дать простое «планиметрическое» доказательство, не использующее понятие векторного произведения и не выводящее за пределы плоскости, в которой расположены векторы a, b. □

Утверждение 2.23 (Критерий коллинеарности векторов)**.** *Векторы a, b на плоскости коллинеарны тогда и только тогда, когда определитель, составленный из их координат (в любом базисе), равен нулю.*

Доказательство. Равенство нулю определителя

$$\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix} = \alpha_1\beta_2 - \alpha_2\beta_1 = 0,$$

очевидно, эквивалентно равенству

$$\frac{\alpha_1}{\beta_1} = \frac{\alpha_2}{\beta_2}$$

(случаи, когда в знаменателях стоят нули, рассматриваются отдельно), которое, в свою очередь, эквивалентно коллинеарности векторов $a(\alpha_1, \alpha_2)$, $b(\beta_1, \beta_2)$. □

Пример 2.24. Найдем векторное произведение $[a, b]$, если $a(1, 2, 3)$, $b(3, 4, 2)$ (в правом ортонормированном базисе)

$$[a, b] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 2 & 3 \\ 3 & 4 & 2 \end{vmatrix} = (2 \cdot 2 - 3 \cdot 4)e_1 - (1 \cdot 2 - 3 \cdot 3)e_2 + (1 \cdot 4 - 3 \cdot 2)e_3 = -8e_1 + 7e_2 - 2e_3.$$

Таким образом, координаты $[a, b]$ суть $(-8, 7, -2)$.

Пример 2.25. Найдем площадь S треугольника ABC, если $A(0, 1, 2)$, $B(2, 4, 3)$, $C(-1, 3, 0)$ (в прямоугольной системе координат); см. рис. 2.19.

Треугольники ABC и $A'BC$ равны. Следовательно, площадь треугольника ABC равна половине площади параллелограмма $ABA'C$:

$$S = \frac{1}{2} AB \cdot AC \cdot \sin\vartheta = \frac{1}{2}\left|[\overrightarrow{AB}, \overrightarrow{AC}]\right|.$$

Так как $[\overrightarrow{AB}] = (2, 3, 1)$, $[\overrightarrow{AC}] = (-1, 2, -2)$, то

$$[\overrightarrow{AB}, \overrightarrow{AC}] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 2 & 3 & 1 \\ -1 & 2 & -2 \end{vmatrix} = -8e_1 + 3e_2 + 7e_3.$$

Таким образом, координаты произведения $[\overrightarrow{AB}, \overrightarrow{AC}]$ суть $(-8, 3, 7)$. Его модуль равен $\sqrt{(-8)^2 + 3^2 + 7^2} = \sqrt{122}$. Итак, $S = \dfrac{\sqrt{122}}{2}$.

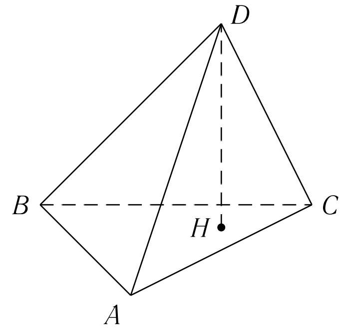


Рис. 2.20:

Пример 2.26. Найдем смешанное произведение векторов $a(1,2,3)$, $c(4,5,7)$, $b(-2,-3,-4)$ (базис правый ортонормированный)

$$(a,b,c) = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ -2 & -3 & -4 \end{vmatrix} = 1\cdot5\cdot(-4)+2\cdot7\cdot(-2)+3\cdot4\cdot(-3)-3\cdot5\cdot(-2)-2\cdot4\cdot(-4)-1\cdot7\cdot(-3) = -1.$$

Пример 2.27. Найдем объем V тетраэдра $ABCD$, если $A(1,1,1)$, $B(1,2,3)$, $C(0,2,1)$, $D(3,2,1)$ (система координат прямоугольная); см. рис. 2.20. Объем равен $V = \dfrac{1}{3}Sh$, где $S = \dfrac{1}{2}\left|[\overrightarrow{AB}, \overrightarrow{AC}]\right|$ — площадь треугольника ABC, $h = DH = \left|\mathrm{pr}_{[\overrightarrow{AB}, \overrightarrow{AC}]} \overrightarrow{AD}\right|$ — высота тетраэдра. Следовательно,

$$V = \frac{1}{6} \cdot \left|\left(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AD}\right)\right|$$

Так как $\overrightarrow{AB}\,(0,1,2)$, $\overrightarrow{AC}\,(-1,1,0)$, $\overrightarrow{AD}\,(2,1,0)$, то

$$(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AD}) = \begin{vmatrix} 0 & 1 & 2 \\ -1 & 1 & 0 \\ 2 & 1 & 0 \end{vmatrix} = -6.$$

Следовательно, $V = 1$.

Пример 2.28. Найдем объем V призмы $ABCA'B'C'$, если $A(1,1,1)$, $B(2,4,3)$, $C(3,-2,2)$, $A'(3,4,2)$ (система координат прямоугольная); см. рис. 2.21. Так как

$$V = \frac{1}{2} \cdot \left|\left(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AA'}\right)\right|,$$

$\overrightarrow{AB}\,(1,3,2)$, $\overrightarrow{AC}\,(2,-3,1)$, $\overrightarrow{AA'}\,(2,3,1)$, то

$$(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AD}) = \begin{vmatrix} 1 & 3 & 2 \\ 2 & -3 & 1 \\ 2 & 3 & 1 \end{vmatrix} = 18.$$

Поэтому $V = 9$.

Выражение $[[a,b],c]$ называется *двойным векторным произведением.*

Утверждение 2.29. *$[[a,b],c] = (a,c)b - (b,c)a$.*

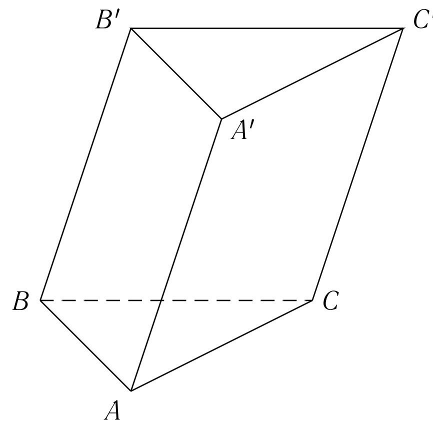


Рис. 2.21:

Доказательство. Простой, но длинный способ доказательства утверждения — ввести произвольный правый ортонормированный базис, перейти к координатам векторов и воспользоваться формулами для выражения скалярного и векторного произведений через координаты. Можно сократить такое доказательство, если выбрать базис специальным образом. Пусть e_1, e_2, e_3 — правый ортонормированный базис, такой, что e_1 коллинеарен вектору v, а e_2 компланарен векторам b и c. Имеем

$$a = \alpha_1 e_1, \quad b = \beta_1 e_1 + \beta_2 e_2, \quad c = \gamma_1 e_1 + \gamma_2 e_2 + \gamma_3 e_3.$$

Откуда $[a, b] = \alpha_1 \beta_2 e_3$ и, следовательно,

$$[[a, b], c] = [\alpha_1 \beta_2 e_3, \gamma_1 e_1 + \gamma_2 e_2 + \gamma_3 e_3] = \alpha_1 \beta_2 \gamma_1 e_2 - \alpha_1 \beta_2 \gamma_2 e_1.$$

С другой стороны,

$$(a, c)b - (b, c)a = \alpha_1 \gamma_1 (\beta_1 e_1 + \beta_2 e_2) - (\beta_1 \gamma_1 + \beta_2 \gamma_2) \alpha_1 e_1 = \alpha_1 \beta_2 \gamma_1 e_2 - \alpha_1 \beta_2 \gamma_2 e_1.$$

Мы видим, что левая и правая части доказываемого равенства совпадают. □

Упражнение 2.30. Доказать, что

$$|[a, b]|^2 = \begin{vmatrix} (a, a) & (a, b) \\ (b, a) & (b, b) \end{vmatrix}.$$

Упражнение 2.31. Доказать, что

$$(a, b, c)^2 = \begin{vmatrix} (a, a) & (a, b) & (a, c) \\ (b, a) & (b, b) & (b, c) \\ (c, a) & (c, b) & (c, c) \end{vmatrix}.$$

Пример 2.32. Найдем объем параллелепипеда $\mathcal{ABCDA'B'C'D'}$, если $\mathcal{AB} = 1$, $\mathcal{AD} = 2$, $\mathcal{AA'} = \sqrt{3}$, $\angle \mathcal{BAD} = \pi/4$, $\angle \mathcal{A'AB} = \pi/6$, $\angle \mathcal{A'AD} = \pi/3$.

Обозначим $a = \overrightarrow{\mathcal{AB}}$, $v = \overrightarrow{\mathcal{AD}}$, $c = \overrightarrow{\mathcal{AA'}}$ Легко вычисляются попарные скалярные произведения этих векторов, так как известны их длины и углы между ними: $(a, a) = 1$, $(b, b) = 4$, $(c, c) = 3$,

$(a, b) = \sqrt{2}$, $(a, c) = 3/2$, $(b, c) = \sqrt{3}$. Осюда, воспользовавшись результатом упражнения 2.31, находим квадрат смешанного произведения:

$$(a, b, c)^2 = \begin{vmatrix} 1 & \sqrt{2} & 3/2 \\ \sqrt{2} & 4 & \sqrt{3} \\ 3/2 & \sqrt{3} & 3 \end{vmatrix} = 3\sqrt{6} - 6.$$

Таким образом, объем параллелепипеда равен $\sqrt{3\sqrt{6} - 6}$.

Глава 3

Комплексные числа

См. пособие Н.Ю.Золотых «Комплексные числа».

Глава 4

Многочлены

Всюду в этой главе через F обозначено некоторое произвольное числовое поле, а через K — некоторое произвольное числовое кольцо.

4.1. Основные определения и простейшие свойства

Многочленом, или *полиномом*, называется выражение вида

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \tag{4.1}$$

где $a_j \in F$ ($j = 0, 1, \ldots, n$), а x — символ, называемый *независимой переменной*. Величины a_j называются *коэффициентами* многочлена, а выражения a_jx^{n-j} — *членами* (или *мономами*) многочлена $f(x)$, при этом $n - j$ называется *степенью* монома. Если $a_0 \neq 0$, то n называется *степенью* многочлена, а a_0x^n — его старшим членом. Многочлен $f(x) = 0$ называется нулевым; его степень не определена. Многочлены 1-й, 2-й и 3-й степени называются *линейными*, *квадратными* и *кубическими* соответственно. Многочлены нулевой степени вместе с нулевым многочленом называют *константами*. В записи (4.1) члены с нулевым коэффициентом обычно опускают. Помимо записи (4.1), в которой члены записаны в порядке убывания степеней, часто используется запись с упорядочением членов по возрастанию степеней и др. записи. Два многочлена (4.1) и

$$g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m \tag{4.2}$$

равны, если $m = n$ и $a_j = b_j$ ($j = 0, 1, \ldots, n$).

Таким образом, мы принимаем алгебраическую точку зрения на многочлены. Возможна также другая — «функциональная» — точка зрения, по которой многочлены рассматриваются как функции $f : F \to F$. Эквивалентность этих точек зрения на многочлены над числовыми полями мы установим в теореме 4.41.

Множество всех многочленов с коэффициентами из множества A обозначим $A[x]$.

Многочлены из $F[x]$ можно складывать и умножать. При этом снова получается многочлен из $F[x]$. Сложение и умножение многочленов выполняется по обычным правилам преобразования алгебраических выражений. Для определения суммы многочленов $f(x)$ и $g(x)$, определенных согласно (4.1) и (4.2), предположим, что $m = n$ (чтобы это условие выполнялось припишем, если необходимо, к $f(x)$ или $g(x)$ нужное количество членов с нулевыми коэффициентами). Тогда *суммой* многочленов $f(x)$ и $g(x)$ называется многочлен

$$f(x) + g(x) = (a_0 + b_0)x^n + (a_1 + b_1)x^{n-1} + \ldots + (a_{n-1} + b_{n-1})x + (a_n + b_n).$$

Произведением многочленов $f(x)$ и $g(x)$, определенных согласно (4.1) и (4.2) (при любых соотношениях между m и n, называется

$$f(x)\,g(x) = a_0b_0x^{n+m} + (a_0b_1 + a_1b_0)x^{n+m-1} + \ldots + (a_{n-1}b_m + a_nb_{m-1})x + a_nb_m,$$

т. е.

$$f(x)\,g(x) = c_0x^{n+m} + c_1x^{n+m-1} + \ldots + c_{n+m-1}x + c_{n+m}, \quad \text{где} \quad c_k = \sum_{i+j=k} a_ib_j. \tag{4.3}$$

Операции сложения и умножения обладают следующими легко проверяемыми свойствами: для любых многочленов $f(x)$, $g(x)$ и $h(x)$ справедливо

1) $f(x) + g(x) = g(x) + f(x)$ (коммутативность сложения),
2) $f(x)g(x) = g(x)f(x)$ (коммутативность умножения),
3) $f(x) + \big(g(x) + h(x)\big) = \big(g(x) + f(x)\big) + h(x)$ (ассоциативность сложения),
4) $f(x)\big(g(x)h(x)\big) = \big(g(x)f(x)\big)h(x)$ (ассоциативность умножения),
5) $f(x)\big(g(x) + h(x)\big) = f(x)g(x) + f(x)h(x)$ (дистрибутивность).

Докажем, например, ассоциативность умножения многочленов. Если $f(x)$ и $g(x)$ определены согласно (4.1) и (4.2) соответственно и

$$h(x) = c_0x_l + c_1x_{l-1} + \ldots + c_{l-1}x + c_l,$$

то

$$f(x)\big(g(x)h(x)\big) = d_0x_{n+m+l} + d_1x_{n+m+l-1} + \ldots + d_{n+m+l-1}x + d_{n+m+l},$$

где, согласно (4.3) (мы пользуемся дистрибутивностью в поле F),

$$d_s = \sum_{i+t=s} a_i \left(\sum_{j+k=t} b_jc_k\right) = \sum_{i+j+k=s} a_ib_jc_k = \sum_{t+k=s}\left(\sum_{i+j=t} a_ib_j\right)c_k.$$

Аналогично, если

$$\big(f(x)\,g(x)\big)h(x) = e_0x_{n+m+l} + e_1x_{n+m+l-1} + \ldots + e_{n+m+l-1}x + e_{n+m+l},$$

то

$$e_s = \sum_{t+k=s}\left(\sum_{i+j=t} a_ib_j\right)c_k = d_s \quad (s = 1, 2, \ldots, m + n + l),$$

откуда $f(x)\big(g(x)h(x)\big) = \big(g(x)f(x)\big)h(x)$.

Легко видеть, что в $F[x]$ константа 0 (и только она) является *нейтральным* элементом относительно сложения, т. е. для любого $f(x) \in F[x]$ справедливо $f(x) + 0 = f(x)$. Для многочлена $f(x) \in F[x]$ *противоположным* называется такой многочлен $h(x)$, что $f(x) + h(x) = 0$. Противоположный многочлен обозначается $-f(x)$. Легко видеть, что для многочлена $f(x)$, опредленного согласно (4.1), противоположным является

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

и только он. Операция вычитания вводится как операцию обратная к сложению, а именно *разностью* или многочленом, полученным в результате *вычитания*, многочленов $g(x)$ и $f(x)$ называется такой многочлен $h(x) = g(x) - f(x)$, что $f(x) + h(x) = g(x)$. Легко видеть, что $g(x) - f(x) = g(x) + (-f(x))$.

В $F[x]$ константа 1 (и только она) является *нейтральным* элементом относительно умножения, т. е. для любого $f(x) \in F[x]$ справедливо $1f(x) = f(x)$. Для многочлена $f(x) \in F[x]$ *обратным* называется такой многочлен $h(x)$, что $f(x)h(x) = 1$. Обратный многочлен (если он существует) обозначается $1/f(x)$. Из утверждения вытекает, что $\deg(1/f(x)) = \deg f(x)$, откуда получаем, что обратный многочлен существует тогда и только тогда, когда $\deg f(x) = 0$, т. е. $f(x)$ — ненулевая константа. Операция деления вводится как операция, обратная умножению, а именно *частным* или многочленом, полученным в результате *деления нацело*, $g(x)$ на $f(x)$ называется такой многочлен $h(x) = g(x)/f(x)$, что $f(x)h(x) = g(x)$. Легко привести примеры, когда деление многочленов нацело невозможно.

Из определения суммы многочленов получаем, что либо $f(x) + g(x) = 0$, либо

$$\deg(f(x) + g(x)) \leq \max\{\deg f(x), \deg g(x)\}.$$

Из определения произведения многочленов получаем, что если $f(x) \neq 0$ и $g(x) \neq 0$, то

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

Отсюда получаем, что в $F[x]$ нет делителей нуля, т. е. из равенства $f(x)g(x) = 0$ следует, что $f(x) = 0$ или $g(x) = 0$.

Лемма 4.1 (Лемма о сокращении)**.** *Если $f(x)h(x) = g(x)h(x)$ и $h \neq 0$, то $f(x) = g(x)$.*

Доказательство. Из равенства $f(x)h(x) = g(x)h(x)$ получаем $\big(f(x) - g(x)\big)h(x) = 0$. Так как в $F[x]$ нет делителей нуля, то $f(x) - g(x) = 0$. □

Замечание 4.2. Многое, что было сказано про многочлены с коэффициентыми из поля F, справедливо и для многочленов с коэффициентами из кольца K. В частности, на множестве $K[x]$ также определены операции сложения, вычитания и умножения, причем результаты этих операций принадлежат $K[x]$ (т. е. $K[x]$ замкнуто относительно этих операций). $K[x]$ содержит единственный нейтральный элемент — константу 0. Заметим, что нейтральный элемент относительно умножения — константа 1 — в $K[x]$ может отсутствовать, так как ее может не быть в кольце K (например, 1 нет в кольце четных чисел).

4.2. Деление многочлена с остатком

Теорема 4.3. *Для любого многочлена $f(x) \in F[x]$ и любого ненулевого многочлена $g(x) \in F[x]$ существуют и единственны многочлены $q(x)$ и $r(x)$ из $F[x]$, такие, что*

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

где $r(x) = 0$ или $\deg r(x) < \deg g(x)$. Многочлен $q(x)$ называется частным, *а $r(x)$ —* остатком *при делении $f(x)$ на $g(x)$.*

Доказательство. Существование Если $f(x) = 0$ или $n < m$, где $n = \deg f(x)$, $m = \deg g(x)$, то положим $p(x) = 0$, $r(x) = f(x)$. В противном случае положим

$$f_1(x) = f(x) - \frac{a_0}{b_0} x^{n-m} g(x), \qquad (\alpha_1)$$

где a_0, b_0 — коэффициенты при старших членах многочленов $f(x)$, $g(x)$ соответственно;

$$f_2(x) = f_1(x) - \frac{a_{01}}{b_0} x^{n_1-m} g(x), \qquad (\alpha_2)$$

где $n_1 = \deg f_1(x)$, а a_{01} — коэффициент при старшем члене многочлена $f_1(x)$;

$$f_3(x) = f_2(x) - \frac{a_{02}}{b_0} x^{n_2-m} g(x), \qquad (\alpha_3)$$

где $n_2 = \deg f_2(x)$, а a_{02} — коэффициент при старшем члене многочлена $f_2(x)$ и т. д. Вычисления будем продолжать до тех пор, пока не будет получен многочлен

$$f_s(x) = f_{s-1}(x) - \frac{a_{0s-1}}{b_0} x^{n_{s-1}-m} g(x), \qquad (\alpha_s)$$

такой, что $f_s(x) = 0$ или $n_s = \deg f_s(x) < m$. Суммируя равенства (α_1), $(\alpha_2), \ldots, (\alpha_s)$, получаем

$$f_s(x) = f(x) - \frac{1}{b_0}(a_0 x^{n-m} + a_{01} x^{n_1-m} + \ldots + a_{0s-1} x^{n_{s-1}-m})\,g(x).$$

Пусть

$$r(x) = f_s(x), \quad q(x) = \frac{1}{b_0}(a_0 x^{n-m} + a_{01} x^{n_1-m} + \ldots + a_{0s-1} x^{n_{s-1}-m}).$$

Легко видеть, что $r(x)$ и $q(x)$ удовлетворяют требуемым свойствам.
Единственность Предположим, что нашлись многочлены $r(x)$, $r_1(x)$, $q(x)$, $q_1(x)$, такие, что

$$f(x) = q(x)\,g(x) + r(x) = q_1(x)\,g(x) + r_1(x), \qquad (4.4)$$

причем $\deg r(x) < \deg g(x)$, $\deg r(x) < \deg g(x)$. Докажем тогда, что $q(x) = q_1(x)$ и $r(x) = r_1(x)$. Действительно, из (4.4) получаем

$$(q(x) - q_1(x))\,g(x) = r_1(x) - r(x). \qquad (4.5)$$

Если $q(x) \neq q_1(x)$, то $\deg((q(x) - q_1(x))\,g(x)) \geq \deg g(x)$, что не возможно, так как $\deg(r_1(x) - r(x)) < \deg g(x)$. Если же $q(x) = q_1(x)$, то из (4.5) получаем, что $r_1(x) = r(x)$. □

Замечание 4.4. Приведенное доказательство существования является конструктивным, т. е. доказательством путем описания соответствующего алгоритма построения искомых объектов. При ручных вычислениях приведенный алгоритм обычно реализуется с помощью схемы деления «уголком».

$$
\begin{array}{l|l}
f(x) & g(x) \\ \cline{2-2}
\underline{\dfrac{a_0}{b_0}x^{n-m}g(x)} & \dfrac{a_0}{b_0}x^{n-m} + \dfrac{a_{01}}{b_0}x^{n_1-m} + \ldots + \dfrac{a_{0,s-1}}{b_0}x^{n_{s-1}-m} \\
\quad f_1(x) & \\
\quad \underline{\dfrac{a_{01}}{b_0}x^{n_1-m}g(x)} & \\
\qquad f_2(x) & \\
\qquad \underline{\dfrac{a_{02}}{b_0}x^{n_2-m}g(x)} & \\
\qquad\quad \ddots & \\
\qquad\qquad f_{s-1}(x) & \\
\qquad\qquad \underline{\dfrac{a_{0,s-1}}{b_0}x^{n_{s-1}-m}g(x)} & \\
\qquad\qquad\quad f_s(x) &
\end{array}
$$

Пример 4.5. Разделим с остатком $x^4 + 2x^3 - 2x + 1$ на $2x^2 + x + 1$:

$$
\begin{array}{rrrr|l}
x^4 + 2x^3 & & -\ 2x & +\ 1 & 2x^2 + x + 1 \\ \cline{5-5}
x^4 + \frac{1}{2}x^3 & + \frac{1}{2}x^2 & & & \frac{1}{2}x^2 + \frac{3}{4}x - \frac{5}{8} \\ \cline{1-4}
\frac{3}{2}x^3 & - \frac{1}{2}x^2 & -\ 2x & +\ 1 & \\
\frac{3}{2}x^3 & + \frac{3}{4}x^2 & + \frac{3}{4}x & & \\ \cline{1-4}
 & -\frac{5}{4}x^2 & - \frac{11}{4}x & +\ 1 & \\
 & -\frac{5}{4}x^2 & - \frac{5}{8}x & - \frac{5}{8} & \\ \cline{2-4}
 & & -\frac{17}{8}x & + \frac{13}{8} &
\end{array}
$$

Итак, получены частное $q(x) = \frac{1}{2}x^2 + \frac{3}{4}x - \frac{5}{8}$ и остаток $r(x) = -\frac{17}{8}x + \frac{13}{8}$.

Замечание 4.6. Обратим внимание, что операция деления с остатком определена на множестве $F[x]$, а не $K[x]$. Легко привести примеры, когда частное и остаток от деления двух многочленов из $K[x]$ уже не принадлежат $K[x]$. Однако если если старший коэффициент b_0 делителя обратим в кольце K (т. е. элемент $1/b_0$ существует и принадлежит K), то частное и остаток принадлежат $K[x]$.

Упражнение 4.7. Пусть $f(x), g(x) \in \mathbb{Z}[x]$, $n = \deg f(x)$, $m = \deg g(x)$ и b_0 — коэффициент при старшей степени многочлена $g(x)$. Докажите, что все многочлены, получающиеся в процессе работы и на выходе алгоритма деления $b_0^{n-m+1}f(x)$ на $g(x)$, имеют целые коэффициенты. Деление многочлена $b_0^{n-m+1}f(x)$ на $g(x)$ называется *псевдоделением* $f(x)$ на $g(x)$. Как по частному и остатку, полученным при псевдоделении, получить частное и остаток от деления $f(x)$ на $g(x)$?

4.3. Делимость. Наибольший общий делитель. Алгоритм Евклида

Пусть $f(x)$ и $g(x)$ — многочлены из $F[x]$, причем $g(x) \neq 0$. Будем говорить, что $g(x)$ является *делителем* $f(x)$, или, просто, что $f(x)$ *делится* на $g(x)$, и писать $f(x) \,\vdots\, g(x)$, если для некоторого $q(x) \in F[x]$ имеем $f(x) = q(x)g(x)$, т. е. остаток при делении $f(x)$ на $g(x)$ равен нулю. Если $f(x)$ не делится на $g(x)$, то будем писать $f(x) \not\vdots\, g(x)$.

Упражнение 4.8. Докажите следующие свойства:

1) Если $f(x) \,\vdots\, g(x)$ и $g(x) \,\vdots\, h(x)$, то $f(x) \,\vdots\, h(x)$.
2) Если $f(x) \,\vdots\, g(x)$ и $h(x) \,\vdots\, g(x)$, то $(f(x) + h(x)) \,\vdots\, g(x)$.
3) Если $f(x) \,\vdots\, g(x)$, то $f(x)h(x) \,\vdots\, g(x)$ для любого $h(x)$.
4) Любой многочлен делится на произвольную ненулевую константу.
5) Если $f(x) \,\vdots\, g(x)$, то $f(x) \,\vdots\, cg(x)$ для любой ненулевой константы c.
6) Для того, чтобы многочлен $f(x)$ делился на многочлен $g(x)$ той же степени, необходимо и достаточно, чтобы $f(x) = cg(x)$ для некоторой константы c.
7) Для того, чтобы $f(x) \,\vdots\, g(x)$ и $g(x) \,\vdots\, f(x)$, необходимо и достаточно, чтобы $f(x) = cg(x)$ для некоторой константы c.

Многочлен $d(x)$ называется *общим делителем* многочленов $f(x)$ и $g(x)$, если $f(x) \,\vdots\, d(x)$ и $g(x) \,\vdots\, d(x)$. Общий делитель $d(x)$ называется *наибольшим* (сокращенно, НОД), если он делится на любой другой общий делитель многочленов $f(x)$ и $g(x)$.

Теорема 4.9. *Для любых $f(x) \in F[x]$ и $g(x) \in F[x]$, одновременно не равных нулю, их наибольший общий делитель $d(x)$ существует и определен однозначно с точностью до множителя c, где c — произвольная ненулевая константа.*

Доказательство. Вначале докажем, что если НОД существует, то он определен с точностью до множителя c. Действительно, если $d(x)$, $d_1(x)$ — наибольшие общие делители многочленов $f(x)$ и $g(x)$, то $d(x) \,\vdots\, d_1(x)$ и $d_1(x) \,\vdots\, d(x)$, поэтому по пункту 7) упражнения 4.8 имеем $d(x) = cd_1(x)$ для некоторой константы c.

Приведем конструктивное доказательство существования наибольшего общего делителя. Опишем хорошо известный *алгоритм Евклида* нахождения НОД. Если $f(x) \neq 0$, а $g(x) = 0$, то, очевидно, в качестве наибольшего общего делителя можно взять $f(x)$, поэтому, не нарушая общности, можно считать, что $g(x) \neq 0$ (напомним, что $f(x)$ и $g(x)$ не равны нулю одновременно).

На нулевой итерации разделим $f(x)$ на $g(x)$, в частном получим $q_1(x)$, в остатке — $r_1(x)$. Если $r_1(x) \neq 0$, то перейдем к первой итерации, на которой разделим $g(x)$ на $r_1(x)$, в частном получим $q_2(x)$, в остатке — $r_2(x)$. На i-й итерации разделим $r_{i-1}(x)$ на r_i, в в частном получим $q_{i+1}(x)$, в остатке — $r_{i+1}(x)$. Вычисления продолжаются до тех пор, пока на некоторой, скажем, s-й, итерации вычисленный в результате очередного деления остаток r_{s+1} не будет нулевым. Докажем, что $r_s(x)$ является наибольшим общим делителем многочленов $f(x)$ и $g(x)$.

Имеем

$$f(x) = q_1(x)g(x) + r_1(x), \tag{β_0}$$

$$g(x) = q_2(x)r_1(x) + r_2(x), \tag{β_1}$$

$$r_1(x) = q_3(x)r_2(x) + r_3(x), \tag{β_2}$$

$$r_{s-3}(x) = q_{s-1}(x)r_{s-2}(x) + r_{s-1}(x), \tag{β_{s-2}}$$

$$r_{s-2}(x) = q_s(x)r_{s-1}(x) + r_s(x), \tag{β_{s-1}}$$

$$r_{s-1}(x) = q_{s+1}(x)r_s(x). \tag{β_s}$$

Из равенства (β_s) следует, что $r_{s-1} \,\vdots\, r(s)$. Поэтому в правой части равенства (β_{s-1}) первое слагаемое делится на $r_s(x)$. Так как второе слагаемое, очевидно, также делится на $r_s(x)$, то вся правая часть равенства (β_{s-1}) делится на $r_s(x)$, поэтому на $r_s(x)$ делится и левая часть этого равенства, т. е. r_{s-2}. В правой части равенства (β_{s-2}) на $r_s(x)$ также делятся оба слагаемых и, следовательно, $r_{s-2} \,\vdots\, r_s$. Рассматривая эти равенства далее снизу вверх (легко провести индукцию), приходим к выводу, что на $r_s(x)$ делятся правые части в (β_0) и (β_1), т. е. $f(x) \,\vdots\, r_s(x)$ и $g(x) \,\vdots\, r_s(x)$, т. е. $r_s(x)$ — общий делитель многочленов $f(x)$ и $g(x)$.

Теперь покажем, что любой общий делитель $d(x)$ многочленов $f(x)$ и $g(x)$ является также делителем многочлена $r_s(x)$. Так как $f(x) \,\vdots\, d(x)$ и $g(x) \,\vdots\, d(x)$, то из (β_0) получаем, что $r_1(x) \,\vdots\, d(x)$. Далее, так как $g(x) \,\vdots\, d(x)$ и $r_1(x) \,\vdots\, d(x)$, то из (β_1) получаем, что $r_2(x) \,\vdots\, d(x)$. Рассматривая далее эти равенства сверху вниз (легко провести индукцию), приходим к выводу, что $r_s(x) \,\vdots\, d(x)$. □

Замечание 4.10. В алгоритме Евклида многочлены $r_j(x)$ $(j = 1, 2, \ldots, s)$ можно умножать на произвольные константы. Легко видеть, что доказательство теоремы 4.9 распространяется и на такую модификацию алгоритма.

Теорема 4.11. *Пусть $f(x)$, $g(x)$, $d(x)$ — ненулевые многочлены из $F[x]$ и $d(x)$ — НОД многочленов $f(x)$ и $g(x)$. Тогда найдутся такие $u(x)$, $v(x)$ из $F[x]$, что*

$$u(x)f(x) + v(x)g(x) = d(x), \tag{4.6}$$

причем $\deg u(x) < \deg g(x)$, *а* $\deg v(x) < \deg f(x)$. *Многочлены $u(x)$ и $v(x)$ называются* коэффициентами Безу.

Доказательство. Очевидно, что достаточно научиться находить коэффициенты Безу по крайней мере для одного из возможных НОД, например, для НОД, выдаваемого алгоритмом Евклида. Применим к $f(x)$ и $g(x)$ алгоритм Евклида. В ходе его работы получим последовательности частных $q_1, q_2, \ldots, q_{s+1}$ и остатков $r_1(x), r_2(x), \ldots, r_s(x)$. На первой итерации запишем два тривиальных равенства:

$$f(x) = 1 \cdot f(x) + 0 \cdot g(x), \tag{γ_{-1}}$$

$$g(x) = 0 \cdot f(x) + 1 \cdot g(x), \tag{γ_0}$$

и вычтем из первого второе, умноженное на q_1. Тогда согласно (β_0) в левой части получим $r_1(x)$. В правой части соберем множители у $f(x)$ и у $g(x)$:

$$r_1(x) = 1 \cdot f(x) + (-q_1(x)) \cdot g(x). \tag{γ_1}$$

На второй итерации вычтем из равенства (γ_0) равенство (γ_1), умноженное на $q_2(x)$. Согласно (β_1) в левой части получим $r_2(x)$. В правой части снова соберем множители у $f(x)$ и у $g(x)$:

$$r_2(x) = (-q_2(x)) \cdot f(x) + (1 + q_1(x)q_2(x)) \cdot g(x). \qquad (\gamma_2)$$

На третьей итерации из (γ_1) вычтем (γ_2), умноженное на $q_3(x)$. Согласно (β_2) в левой части получим $r_3(x)$. В правой части снова соберем множители у $f(x)$ и у $g(x)$:

$$r_3(x) = (1 + q_2(x)q_3(x)) \cdot f(x) + (-q_1(x) - q_3(x) - q_1(x)q_2(x)q_3(x)) \cdot g(x). \qquad (\gamma_3)$$

Будем выполнять такие преобразования далее. На k-й итерации ($k = 1, 2, \ldots, s$), вычитая из равенства (γ_{k-2}) равенство (γ_{k-1}), умноженное на $q_k(x)$, получим

$$r_k(x) = u_k(x)f(x) + v_k(x)g(x), \qquad (\gamma_k)$$

где $u_k(x)$, $v_k(x)$ — некоторые многочлены. На s-й итерации получим

$$d(x) = r_s(x) = u_s(x)f(x) + v_s(x)g(x). \qquad (\gamma_s)$$

Можно считать, что многочлены $u_s(x)$ и $v_s(x)$ ненулевые[1]. Если $\deg u_s(x) < \deg g(x)$ и $\deg v_s(x) < \deg f(x)$, то $u_s(x)$ и $v_s(x)$ — искомые коэффициенты Безу. В противном случае выполним следующую процедуру.

Пусть, для определенности, $\deg u_s(x) \geq \deg g(x)$. Разделим $u_s(x)$ на $g(x)$: $u_s(x) = q(x)g(x) + r(x)$. Теперь из (γ_s) получаем

$$d(x) = r(x)f(x) + (v_s(x) + q(x)f(x))g(x).$$

Обозначим $u(x) = r(x)$, $v(x) = v_s(x) + q(x)f(x)$. Имеем $u(x)f(x) + v(x)g(x) = d(x)$ и $\deg u(x) < \deg g(x)$. Покажем, что $\deg v(x) < \deg f(x)$, тем самым завершив доказательство теоремы. Предположим противное: $\deg v(x) \geq \deg f(x)$, тогда $\deg(u(x)f(x)) < \deg g(x) + \deg f(x)$, но $\deg(v(x)g(x)) \geq \deg g(x) + \deg f(x)$, поэтому

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

что противоречит тому, что $d(x)$ — НОД многочленов $f(x)$ и $g(x)$. ☐

Алгоритм нахождения коэффициентов Безу, описанный при доказательстве теоремы 4.11, называется *расширенным алгоритмом Евклида*.

Упражнение 4.12. *Континуантой* $\kappa(a_1, a_2, \ldots, a_n)$ называется сумма всевозможных произведений элементов $a_1, a_2, \ldots, a_n$, одно из которых содержит все эти элементы, а другие получаются из него выбрасыванием одной или нескольких пар сомножителей с соседними номерами. При этом член, получаемый выбрасыванием всех сомножителей (при четном n), считается равным 1. Докажите, что для $u_k(x)$ и $v_k(x)$ из (γ_k) справедливо $u_k(x) = (-1)^{k+1}\kappa(q_2(x), q_3(x), \ldots q_k(x))$, $v_k(x) = (-1)^k\kappa(q_1(x), q_2(x), \ldots, q_k(x))$.

---

[1] Если $u(x)f(x) + v(x)g(x) = d(x)$ и один из многочленов $u(x)$ или $v(x)$ — нулевой: для определенности, $v(x) = 0$, то $d(x) \vdots f(x)$, но так как $f(x) \vdots d(x)$ и $g(x) \vdots d(x)$, то $u(x)$ — константа и многочлены $f(x)$, $g(x)$, $d(x)$ отличаются друг от друга константными множителями. Поэтому многочлены $u(x)$ и $v(x)$ можно заменить на ненулевые константы, так, что равенство $u(x)f(x) + v(x)g(x) = d(x)$ останется выполненным.

Пример 4.13. Найдем НОД и коэффициенты Безу многочленов $f(x) = x^4 - 3$, $g(x) = x^3 + 2x^2 + x + 1$. При делении $f(x)$ на $g(x)$ получаем частное и остаток:

$$q_1(x) = x - 2, \qquad r_1(x) = 3x^2 + x - 1.$$

При делении $g(x)$ на $r_1(x)$ получаем

$$q_2(x) = \frac{1}{3}x + \frac{5}{9}, \qquad r_2(x) = \frac{7}{9}x + \frac{14}{9}.$$

При делении $r_1(x)$ на $\frac{9}{7} \cdot r_2(x) = x + 2$ получаем

$$q_3(x) = 3x - 5, \qquad r_3(x) = 9.$$

Остаток при делении $r_2(x)$ на $r_3(x)$ равен нулю, следовательно, $cr_3(x)$, где c — произвольная ненулевая константа, является наибольшим общим делителем многочленов $f(x)$ и $g(x)$. Итак, НОД многочленов $f(x)$ и $g(x)$ с точностью до ненулевого константного множителя равен 1.

Для нахождения коэффициентов Безу воспользуемся алгоритмом из доказательства теоремы 4.11. Имеем

$$f(x) = 1 \cdot f(x) + 0 \cdot g(x), \tag{4.7}$$

$$g(x) = 0 \cdot f(x) + 1 \cdot g(x). \tag{4.8}$$

Вычитая из равенства (4.7) равенство (4.8), умноженное на $q_1(x) = x - 2$, получаем

$$r_1(x) = 3x^2 + x - 1 = 1 \cdot f(x) + (-x + 2) \cdot g(x). \tag{4.9}$$

Вычитая из равенства (4.8) равенство (4.9), умноженное на $q_2(x) = \frac{1}{3}x + \frac{5}{9}$, получаем

$$r_2(x) = \frac{7}{9}x + \frac{14}{9} = \left(-\frac{1}{3}x - \frac{5}{9}\right) \cdot f(x) + \left(\frac{1}{3}x^2 - \frac{1}{9}x - \frac{1}{9}\right) \cdot g(x),$$

откуда

$$\frac{9}{7} \cdot r_2(x) = x + 2 = \left(-\frac{3}{7}x - \frac{5}{7}\right) \cdot f(x) + (3x^2 - x - 1) \cdot g(x). \tag{4.10}$$

Вычитая из равенства (4.9) равенство (4.10), умноженное на $q_3(x) = 3x - 5$, получаем

$$r_3(x) = 9 = \left(\frac{9}{7}x^2 - \frac{18}{7}\right) \cdot f(x) + \left(-\frac{9}{7}x^3 + \frac{18}{7}x^2 - \frac{9}{7}x + \frac{9}{7}\right) \cdot g(x),$$

откуда

$$1 = \left(\frac{1}{7}x^2 - \frac{2}{7}\right) \cdot f(x) + \left(-\frac{1}{7}x^3 + \frac{2}{7}x^2 - \frac{1}{7}x + \frac{1}{7}\right) \cdot g(x).$$

Итак,

$$u(x) = \frac{1}{7}x^2 - \frac{2}{7}, \qquad v(x) = -\frac{1}{7}x^3 + \frac{2}{7}x^2 - \frac{1}{7}x + \frac{1}{7}.$$

Иногда удобно использовать другой алгоритм нахождения коэффициентов Безу, основанный на общем *методе неопределенных коэффициентов*. Этот способ предполагает, что НОД $d(x)$ многочленов $f(x)$ и $g(x)$ уже известен. Во-первых, вместо $f(x)$ и $g(x)$ рассмотрим многочлены $f_1(x) = f(x)/d(x)$ и $g_1(x) = g(x)/d(x)$. Из (4.6) получаем

$$1 = u(x)f_1(x) + v(x)g_1(x). \tag{4.11}$$

Далее запишем $u(x)$ и $v(x)$ с неопределенными коэффициентами, учитывая, что $\deg u(x) < \deg g(x)$, и приравняем коэффициенты при одинаковых степенях в левой и правой частях равенства (4.11). Из полученной линейной системы определим коэффициенты многочленов $u(x)$ и $v(x)$. Другой способ составить систему для определения этих коэффициентов заключается в следующем. Выберем $\deg u(x) + \deg g(x)$ различных чисел (по числу неопределенных коэффициентов). Для облегчения вычислений желательно, чтобы среди этих чисел были корни многочленов $f(x)$ и $g(x)$. Последовательно подставляя эти числа вместо x в (**??**), получим линейную систему, из которой определим неизвестные коэффициенты.

Пример 4.14. Найдем коэффициенты Безу многочленов $f(x) = x^6 + 2x^5 + 4x^4 + 4x^3 + 2x^2 - 1$ и $g(x) = x^5 + x^4 + x^2 - 5x + 2$, если известен их НОД $d(x) = x^3 + 2x - 1$. Имеем $f_1(x) = f(x)/d(x) = x^3 + 2x^2 + 2x + 1$, $g_1(x) = g(x)/d(x) = x^2 + x - 2$. Равенство (4.11) примет вид

$$1 = (a_0x + a_1)(x^3 + 2x^2 + 2x + 1) + (b_0x^2 + b_1x + b_2)(x^2 + x - 2).$$

Последовательно подставляя в него вместо x числа $-2, -1, 0, 1, 2$, получим систему

$$\begin{cases} 6a_0 - 3a_1 = 1, \\ -2b_0 + 2b_1 - 2b_2 = 1, \\ a_1 - 2b_2 = 1, \\ 6a_0 + 6a_1 = 1, \\ 42a_0 + 21a_1 + 16b_0 + 8b_1 + 4b_2 = 1, \end{cases}$$

из которой находим $a_0 = 1/6$, $a_1 = 0$, $b_0 = -1/6$, $b_1 = -1/6$, $b_2 = -1/2$. Таким образом,

$$u(x) = \frac{1}{6}x, \qquad v(x) = -\frac{1}{6}x^2 - \frac{1}{6}x - \frac{1}{2}.$$

Упражнение 4.15. Пусть $f(x), g(x), h(x) \in F[x]$. Докажите, что для того, чтобы многочлен $h(x)$ можно было представить в виде $h(x) = u_1(x)f(x) + v_1(x)g(x)$, где $u_1(x)$ и $v_1(x)$ — некоторые многочлены из $F[x]$, необходимо и достаточно, чтобы $h(x)$ делился на НОД многочленов $f(x)$ и $g(x)$. Как $u_1(x)$ и $v_1(x)$ можно выразить через коэффициенты Безу многочленов $f(x)$ и $g(x)$?

4.4. Взаимно простые многочлены

Многочлены $f(x)$ и $g(x)$ назовем *взаимно простыми*, если их НОД равен ненулевой константе.

Следствие 4.16. *Для того, чтобы многочлены $f(x)$ и $g(x)$ из $F[x]$ были взаимно простыми необходимо и достаточно, чтобы в $F[x]$ существовали такие $u(x)$ и $v(x)$, что $u(x)f(x) + v(x)g(x) = 1$.*

Доказательство. Необходимость является прямым следствием теоремы 4.11. *Достаточность* Пусть $d(x)$ — НОД многочленов $f(x)$ и $g(x)$, поэтому каждое из слагаемых в левой части равенства $u(x)f(x) + v(x)g(x) = 1$ делится на $d(x)$, а значит на $d(x)$ делится и правая часть: $1 \,\vdots\, d(x)$, откуда выводим, что $d(x)$ — ненулевая константа. □

Следствие 4.17. *Пусть поле F является подполем поля F'. Для того, чтобы многочлены $f(x)$ и $g(x)$ из $F[x]$ были взаимно простыми необходимо и достаточно, чтобы эти многочлены, рассматриваемые как элементы множества $F'[x]$, были также взаимно простыми.*

Доказательство. Вытекает из следствия 4.16. □

Утверждение 4.18. *Если многочлен $f(x)$ взаимно прост с каждым из многочленов $g(x)$ и $h(x)$, то он взаимно прост и с их произведением $g(x)h(x)$.*

Доказательство. Так как многочлен $f(x)$ взаимно прост с каждым из многочленов $g(x)$ и $h(x)$, то по следствию 4.16 найдутся такие $u(x)$, $u_1(x)$, $v(x)$, $v_1(x)$, что

$$u(x)\ f(x) + \ v(x)\ g(x) = 1,$$
$$u_1(x)\ f(x) + v_1(x)\ h(x) = 1.$$

Складывая эти два равенства и осуществляя очевидные преобразования, получаем:

$$\Big(u(x)u_1(x)f(x)u(x)v_1(x)h(x) + v(x)g(x)u_1(x)\Big)f(x) + \Big(v(x)v_1(x)\Big)g(x)h(x) = 1,$$

откуда по следствию 4.16 получаем, что многочлены $f(x)$ и $g(x)h(x)$ взаимно просты. □

Утверждение 4.19. *Если $f(x)g(x) \vdots h(x)$, причем многочлены $f(x)$ и $h(x)$ взаимно просты, то $g(x) \vdots h(x)$.*

Доказательство. Так как многочлены $f(x)$ и $h(x)$ взаимно просты, то по следствию 4.16 найдутся такие $u(x)$, $v(x)$, что $u(x)f(x) + v(x)h(x) = 1$. Умножая обе части этого равенства на $g(x)$, получаем

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

Очевидно, второе слагаемое в левой части делится на $h(x)$. Первое слагаемое левой части делится на $h(x)$ по условию. Следовательно, $g(x) \vdots h(x)$. □

Утверждение 4.20. *Если $f(x) \vdots g(x)$ и $f(x) \vdots h(x)$, причем $g(x)$ и $h(x)$ взаимно просты, то $f(x) \vdots g(x)h(x)$.*

Доказательство. Так как $f(x) \vdots g(x)$, то для некоторого $q(x)$ имеем $f(x) = q(x)g(x)$. Но $f(x) \vdots h(x)$. Так как $g(x)$ и $h(x)$ взаимно просты, то по утверждению 4.19 $g(x) \vdots h(x)$, откуда $f(x) \vdots g(x)h(x)$. □

Рассмотрим систему многочленов $f_1(x), f_2(x), \ldots, f_s(x)$ из $F[x]$. Многочлен $d(x)$ называется их *общим делителем*, если $f_j \vdots d(x)$ $(j = 1, 2, \ldots, s)$. Общий делитель называется *наибольшим*, если он делится на любой другой их общий делитель.

Упражнение 4.21.

1) Пусть $d(x)$ — НОД многочленов $f_1(x), \ldots, f_{s-1}$. Докажите, что НОД многочленов $d(x)$ и $f_s(x)$ является наибольшим общим делителем многочленов $f_1(x), f_2(x), \ldots, f_s(x)$.

2) Пусть $d(x)$ — НОД многочленов $f_1(x), \ldots, f_s$. Докажите, что тогда найдутся такие $u_1(x), u_2(x), \ldots u_s(x)$, что
$$u_1(x)f_1(x) + u_2(x)f_2(x) + \ldots + u_s(x)f_s(x) = d(x).$$

Многочлены $f_1(x), f_2(x), \ldots, f_s(x)$ называются *взаимно простыми* (*в совокупности*), если их наибольший общий делитель равен ненулевой константе. Многочлены $f_1(x), f_2(x), \ldots, f_s(x)$ называются *попарно взаимно простыми*, если при любых i, j, $1 \le i < j \le s$ многочлены $f_i(x)$ и $f_j(x)$ взаимно просты.

Очевидно, что если многочлены попарно взаимно просты, то они взаимно просты в совокупности. Легко привести примеры, показывающие, что обратное утверждение в общем случае не верно.

Упражнение 4.22. Докажите, что для того, чтобы многочлены $f_1(x), f_2(x), \ldots f_s(x)$ были взаимно простыми необходимо и достаточно, чтобы существовали такие $u_1(x), u_2(x), \ldots u_s(x)$, что

$$u_1(x)f_1(x) + u_2(x)f_2(x) + \ldots + u_s(x)f_s(x) = 1.$$

4.5. Корни многочлена

Значением многочлена $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in F[x]$ в точке $c \in F$ называется число из F, обозначаемое $f(c)$ и равное

$$f(c) = a_0c^n + a_1c^{n-1} + \ldots + a_{n-1}c + a_n \in F. \tag{4.12}$$

Если $f(c) = 0$, то c называется *корнем* многочлена $f(x)$. Заметим, что если $f(x) \in K[x]$ и $c \in K$, то $f(c) \in K$.

Очевидно, значение от суммы, разности и произведения многочленов совпадает с суммой, разностью и произведением значений соответствующих многочленов в той же точке.

Утверждение 4.23 (Теорема Безу)**.** *Остатком при делении многочлена $f(x) \in F[x]$ на линейный многочлен $x - c \in F[x]$ является константа $f(c) \in F$.*

Доказательство. При делении $f(x)$ на $x - c$ в частном получим многочлен $q(x)$, а в остатке — константу r:

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

Полагая в левой и правой части $x = c$, получаем: $f(c) = q(c)(c - c) + r = r$. □

Следствие 4.24.

$$f(c) = 0 \quad \Leftrightarrow \quad f(x) \,\vdots\, (x - c).$$

Будем говорить, что корень c многочлена $f(x)$ имеет *кратность* k, если

$$f(x) \,\vdots\, (x - c)^k, \qquad f(x) \,\not\vdots\, (x - c)^{k+1}.$$

Корень кратности 1, назовем *простым*.

Схема Горнера — это метод деления многочлена $f(x)$ на линейный множитель $x - c$.

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = \big(q_0x^{n-1} + q_1x^{n-2} + \ldots + q_{n-2}x + q_{n-1}\big)(x - c) + r.$$

Раскрывая скобки в правой части равенства и приравнивая коэффициенты из разных частей при равных степенях, получаем

$$\begin{aligned} a_0 &= q_0 \\ a_1 &= q_1 - q_0 c \\ a_2 &= q_2 - q_1 c \\ &\dots\dots\dots\dots \\ a_{n-1} &= q_{n-1} - q_{n-2} c \\ a_n &= r - q_{n-1} c \end{aligned}$$

откуда

$$\begin{aligned} q_0 &= a_0 \\ q_1 &= a_1 + q_0 c \\ q_2 &= a_2 + q_1 c \\ &\dots\dots\dots\dots \\ q_{n-1} &= a_{n-1} + q_{n-2} c \\ r &= a_n + q_{n-1} c \end{aligned} \tag{4.13}$$

Схема Горнера заключается в вычислении коэффициентов частного и остатка по формулам (4.13). При ручных вычислениях обычно используют таблицу следующего вида:

	a_0	a_1	a_2	$\dots$	a_{n-1}	a_n
c	$q_0 = a_0$	$q_1 = a_1 + q_0 c$	$q_2 = a_2 + q_1 c$	$\dots$	$q_{n-1} = a_{n-1} + q_{n-2} c$	$r = a_n + q_{n-1} c$

Пример 4.25. Разделим $f(x) = x^4 - 8x^3 + 21x^2 - 21x + 7$ на $x - 2$.

	1	-8	21	-21	7
2	1	-6	9	-3	1

Получили частное $q(x) = x^3 - 6x^2 + 9x - 3$ и остаток $f(2) = 1$.

Из схемы Горнера следует

$$f(c) = \Big(\dots\big((a_0 c + a_1)c + a_2\big)c + \dots + a_{n-1}\Big)c + a_n.$$

Заметим, что ту же формулу можно получить простым преобразованием правой части равенства (4.12).

Производной многочлена

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

называется многочлен

$$f'(x) = a_0 n x^{n-1} + a_1 (n-1) x^{n-2} + \dots + 2a_{n-2} x + a_{n-1}.$$

Индуктивно определим *производную k-го порядка* (или, просто, *k-ю производную*). Пусть $f^{(k-1)}(x)$ — производная $(k-1)$-го порядка. Производной k-го порядка называется многочлен $f^{(k)}(x) = \big(f^{(k-1)}(x)\big)'$.

Замечание 4.26. Определение производной многочлена, очевидно, согласуется с общим определением производной функции, даваемым в математическом анализе с помощью предельного перехода. Мы дали именно «формальное» определение, по крайней мере, по двум причинам. Во-первых, мы стараемся последовательно проводить «алгебраическую» точку зрения, рассматривая многочлены как формальные алгебраические выражения. Во-вторых, далее мы построим теорию многочленов над достаточно произвольными областями (не обязательно числовыми полями), в которых предельный переход не возможен, но, тем не менее, понятие производной, существует.

Упражнение 4.27. Пусть $f(x), g(x) \in F[x]$. Докажите следующие свойства:

1) $(f(x) \pm g(x))' = f'(x) \pm g'(x)$,
2) $(f(x) \cdot g(x))' = f'(x)g(x) + f(x)g'(x)$,
3) $(f^k(x))' = kf^{k-1}(x)f'(x)$.

Теорема 4.28. *Корень кратности $k \geq 2$ многочлена $f(x)$ является корнем кратности $k-1$ многочлена $f'(x)$. Простой корень многочлена $f(x)$ не является корнем многочлена $f'(x)$.*

Доказательство. Пусть c — корень кратности k многочлена $f(x)$:

$$f(x) = q(x)(x-c)^k, \qquad q(x) \not\vdots\, (x-c).$$

Тогда

$$f'(x) = q'(x)(x-c)^k + q(x)k(x-c)^{k-1} = \big(q'(x)(x-c) + q(x)k\big)(x-c)^{k-1}.$$

Так как

$$q'(x)(x-c) \vdots (x-c), q(x)k \not\vdots\, (x-c), \text{то} \big(q'(x)(x-c) + q(x)k\big) \not\vdots\, (x-c),$$

откуда следует доказываемое. ☐

Следствие 4.29. *Для того, чтобы корень многочлена $f(x)$ имел кратность k необходимо и достаточно, чтобы он являлся корнем последовательных производных многочлена $f(x)$ вплоть до $(k-1)$-го порядка и не являлся корнем k-й производной.*

Утверждение 4.30. *Для любой константы $c \in F$ многочлен*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in F[x],$$

можно представить в виде

$$f(x) = b_0(x-c)^n + b_1(x-c)^{n-1} + \ldots + b_{n-1}(x-c) + b_n, \tag{4.14}$$

называемом разложением $f(x)$ по степеням $x - c$. *При этом коэффициенты $b_0, b_1, \ldots, b_n$ определяются единственным образом.*

Доказательство. Чтобы получить разложение (4.14) разделим $f(x)$ с остатком на $x - c$. В остатке, очевидно, получим b_n, а в частном некоторый многочлен $f_1(x)$. Разделим $f_1(x)$ с остатком на $x - c$. В остатке, очевидно, получим b_{n-1}, а в частном некоторый многочлен $f_2(x)$ и т. д. Таким образом, существование разложения (4.14) доказано. Единственность легко доказывается методом от противного. ☐

Пример 4.31. Разложим $f(x) = x^4 - 8x^3 + 21x^2 - 21x + 7$ по степеням $g(x) = x - 2$. Для этого применим схему Гонера сначала к многочленам $f(x)$ и $g(x)$, затем к их частному $f_1(x)$ и $g(x)$ и т. д.:

	1	−8	21	−21	7
2	1	−6	9	−3	**1**
2	1	−4	1	−**1**	
2	1	−2	−**3**		
2	1	**0**			
2	**1**				

(4.15)

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

Пример 4.32. С помощью схемы Горнера разложим многочлен $f(x) = (x-2)^4 + 2(x-2)^3 + 2(x-2)^2 - 2$ по степеням x.
1-*й способ*. Необходимо решить задачу, обратную задаче из примера 4.31. Заполняя таблицу, аналогичную (4.15), *снизу вверх*, получаем:

	1	**−6**	**14**	**−16**	**6**
2	1	−4	6	−4	-2
2	1	−2	2	0	
2	1	0	2		
2	1	2			
2	1				

Итак, $f(x) = x^4 - 6x^3 + 14x^2 - 16x + 6$.
2-*й способ*. Вместо многочлена $f(x)$ рассмотрим многочлен $g(y) = y^4 + 2y^3 + 2y^2 - 2$, который разложим по степеням $y+2$. Получим $g(y) = (y+2)^4 - 6(y+2)^3 + 14(y+2)^2 - 16(y+2) + 6$, откуда $f(x) = x^4 - 6x^3 + 14x^2 - 16x + 6$.

Утверждение 4.33 (Формула Тейлора)**.** *Для любой константы $c \in F$ и любого многочлена $f(x) \in F[x]$ степени n справедливо*

$$f(x) = f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n-1)}(c)}{(n-1)!}(x-c)^{n-1} + \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

Доказательство. Взяв k-ю производную от многочлена

$$f(x) = b_0 + b_1(x-c) + b_2(x-c)^2 + \ldots + b_n(x-c)^n$$

и подставив $x = c$, получим $b_k = f^{(k)}(c)/k!$. □

Пример 4.34. Для многочлена $f(x) = x^4 - 8x^3 + 21x^2 - 21x + 7$ в примере 4.31 найдено разложение $f(x) = (x-2)^4 - 3(x-2)^2 - (x-2) + 1$, откуда можно сразу получить значения производных в точке 2:

$$f'(2) = -1 \cdot 1! = -1, \quad f''(2) = 0 \cdot 2! = 0, \quad f'''(2) = -3 \cdot 3! = -18, \quad f^{IV}(2) = 1 \cdot 4! = 24.$$

4.6. «Основная теорема алгебры»

Если любой многочлен $f(x) \in F[x]$, степени не меньшей 1, имеет по крайней мере один корень из F, то поле F называется *алгебраически замкнутым*. Поле $\mathbb{R}$ алгебраически замкнутым не является. Например, многочлен $x^2 + 1$ не имеет вещественных корней. Именно это послужило основной причиной построения поля комплексных чисел. Поле $\mathbb{C}$ алгебраически замкнуто, а именно, справедлива

Теорема 4.35 («Основная теорема алгебры»)**.** *Любой многочлен $f(x) \in \mathbb{C}[x]$ степени не меньше 1 имеет по крайней мере один комплексный корень.*

Известно несколько доказательтв «основной теоремы алгебры». Здесь мы приведем одно доказательство, называемое «дамой с собачкой».

Доказательство. Не нарушая общности, можно считать, что старший коэффициент многочлена равен 1. Кроме того, будем считать, что свободный коэффициент не равен нулю, так как в противном случае число 0 является корнем многочлена. Итак,

$$f(x) = x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{C}[x], \quad a_n \neq 0.$$

Положим $A = \max\{|a_1|, |a_2|, \ldots, |a_n|\}$, $a = \min\left\{1, \dfrac{a_n}{n \max\{A, 1\}}\right\}$.

Во-первых, покажем, что если $\alpha \in \mathbb{C}$ и $|\alpha| \geq A+1$, то $|\alpha^n| > |a_1\alpha^{n-1} + \ldots + a_n|$. Действительно,

$$\begin{aligned} |a_1\alpha^{n-1} + a_2\alpha^{n-2} + \ldots + a_n| &\leq |a_1||\alpha|^{n-1} + |a_2||\alpha|^{n-2} + \ldots + |a_n| \leq \\ &\leq A(|\alpha|^{n-1} + |\alpha|^{n-2} + \ldots + |a_n|) = \\ &= A \cdot \frac{|\alpha|^n - 1}{|\alpha| - 1} \leq \\ &\leq A \cdot \frac{|\alpha|^n - 1}{(A+1) - 1} = \\ &= |\alpha|^n - 1. \end{aligned}$$

Во-вторых, покажем, что если $\alpha \in \mathbb{C}$ и $|\alpha| < a$, то $|f(\alpha) - a_n| < a_n$. Действительно,

$$\begin{aligned} |f(\alpha) - a_n| &\leq |\alpha| \cdot |\alpha^{n-1} + a_1\alpha^{n-2} + \ldots + a_{n-1}| < \\ &< a \cdot (|\alpha|^{n-1} + |a_1||\alpha|^{n-2} + \ldots + |a_{n-1}|) \leq \\ &\leq a \cdot \max\{A, 1\} \cdot (|\alpha|^{n-1} + |\alpha|^{n-2} + \ldots + 1) \leq \\ &\leq a \cdot \max\{A, 1\} \cdot n \leq \\ &\leq a_n. \end{aligned}$$

Для завершения доказательства исследуем линии, по которым проходит точка $f(\alpha)$, когда α пробегает на комплексной плоскости окружность $|\alpha| = r$. При каждом фиксированном $r > 0$ линия $f(\alpha)$ — некоторая непрерывная замкнутая кривая.

Если $r = A+1$, то эта кривая ограничивает точку 0. Действительно, легко видеть, что кривая α^n обращается (по окружности $|\alpha| = r^n$) вокруг нуля n раз. Однако так как $|\alpha^n| > |a_1\alpha^{n-1} + \ldots + a_n|$, то кривая $f(\alpha) = \alpha^n + a_1\alpha^{n-1} + \ldots + a_n$ также n раз обращается вокруг точки 0 (уже не обязательно по окружности)[2].

Если $r = a/2$, то, ввиду неравенства $|f(\alpha) - a_n| < a$, кривая $f(\alpha)$ лежит ближе, чем на расстоянии a_n, от точки a_n. Следовательно, точка 0 лежит вне этой кривой.

Если r пробегает все вещественные значения от $A+1$ до $a/2$, то рассматриваемая кривая плавно меняется. Но так как при $r = A+1$ точка 0 лежит снаружи этой кривой, а при $r = a/2$ — внутри, то при некотором r кривая пересечет точку 0, т. е. для некоторого α получим $f(\alpha) = 0$. ☐

Следствие 4.36. *Для любого многочлена*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{C}[x] \quad (a_0 \neq 0)$$

найдутся такие числа $c_1, c_2, \ldots, c_s \in \mathbb{C}$ *и* $k_1, k_2, \ldots, k_s \in \mathbb{N}$, *что*

$$\begin{gathered} f(x) = a_0(x - c_1)^{k_1}(x - c_2)^{k_2} \ldots (x - c_s)^{k_s}, \\ \text{где} \quad k_1 + \ldots + k_s = n, \quad c_i \neq c_j \quad (i \neq j;\ i, j = 1, 2, \ldots, s). \end{gathered} \tag{4.16}$$

[2] Этому можно дать следующую интерпретацию: α^n — это «дама», $f(\alpha)$ — «собачка», $a_1\alpha^{n-1} + \ldots + a_n$ — «поводок». «Дама» гуляет вокруг точки 0. «Поводок» на столько короток, что «собачка» не может добежать до нуля.

Для любого многочлена представление вида (4.16) *с указанными свойствами единственно с точностью до перестановки множителей* $(c-c_j)^{k_j}$ $(j=1,2,\ldots,s)$. *Запись вида* (4.16) *называется* разложением многочлена $f(x)$ на линейные множители.

Доказательство. Существование. Пусть $f(x)\in\mathbb{C}[x]$ и $\deg f(x) >= 1$. Согласно теореме 4.35 найдется $c_1\in\mathbb{C}$, такое, что $f(c_1)=0$. По теореме Безу имеем:

$$f(x)=(x-c_1)f_1(x),$$

где $f_1(x)\in\mathbb{C}[x]$. Если $\deg f_1(x) >= 1$, то по теореме 4.35 найдется $c_2\in\mathbb{C}$, такое, что $f(c_2)=0$, поэтому

$$f(x)=(x-c_1)(x-c_2)f_2(x),$$

где $f_2(x)\in\mathbb{C}[x]$. Продолжая эти рассуждения далее (легко провести индукцию), получим

$$f(x)=c_0(x-c_1)(x-c_2)\ldots(x-c_n),$$

где c_0 — константа. Легко видеть, что коэффициент при старшей степени равен c_0, поэтому $c_0=a_0$. После переобозначения получаем (4.16).
Единственность. Предположим, что

$$f(x)=a_0(x-c_1)^{k_1}(x-c_2)^{k_2}\ldots(x-c_s)^{k_s}=b_0(x-d_1)^{l_1}(x-d_2)^{l_2}\ldots(x-d_t)^{l_t}. \quad (4.17)$$

Во-первых, $a_0=b_0$.

Во-вторых, докажем, что $\{c_1,c_2,\ldots,c_s\}=\{d_1,d_2,\ldots,d_t\}$. Предположим противное: для определенности, $c_1\neq d_j$ $(j=1,2,\ldots,t)$. Подставив c_1 в обе части равенства (4.17), получаем слева 0, а справа ненулевое значение. Поэтому $s=t$ и мы можем считать, что $c_j=d_j$. Теперь (4.17) имеет вид

$$a_0(x-c_1)^{k_1}(x-c_2)^{k_2}\ldots(x-c_s)^{k_s}=a_0(x-c_1)^{l_1}(x-c_2)^{l_2}\ldots(x-c_s)^{l_s}. \quad (4.18)$$

Наконец, докажем, что $k_j=l_j$ $(j=1,2,\ldots,$. Предположим противное: для определенности, $k_1>l_1$. Тогда по лемме из (4.18) получаем

$$(x-c_1)^{k_1-l_1}(x-c_2)^{k_2}\ldots(x-c_s)^{k_s}=(x-c_2)^{l_2}\ldots(x-c_t)^{l_t}. \quad (4.19)$$

Подставив c_1 в обе части равенства (4.19), получаем слева 0, а справа ненулевое значение. □

Замечание 4.37. В представлении (4.16) многочлена $f(x)$ в виде произведения множителей вида $(x-c_j)^{k_j}$ $(j=1,2,\ldots,t)$ показатель степени k_j является кратностью многочлена $f(x)$.

Следствие 4.38. *Всякий ненулевой многочлен* $f(x)\in\mathbb{C}[x]$ *степени* n *имеет с учетом кратности ровно* n *комплексных корней.*

Заметим, что полученное следствие справедливо для любого алгебраически замкнутого поля.

4.7. Интерполяционный многочлен

Рассмотрим таблицу чисел

$$\begin{array}{|c|c|c|c|}\hline x_0 & x_1 & \dots & x_n \\ \hline y_0 & y_1 & \dots & y_n \\ \hline \end{array}, \qquad x_i \neq y_j, \qquad (i \neq j). \tag{4.20}$$

Многочлен $f(x) \in \mathbb{C}[x]$ степени не выше n, удовлетворяющий условиям

$$f(x_j) = y_j \quad (j = 0, 1, \dots, n) \tag{4.21}$$

назовем *интерполяционным многочленом*, относящимся к интерполяционной таблице (4.20).

Теорема 4.39. *Для любой таблицы интерполяции* (4.20) *интерполяционный многочлен существует и единственен.*

Доказательство. Существование. Легко видеть, что многочлен

$$f(x) = \sum_{j=0}^{n} y_j \cdot \frac{(x - x_1) \cdot \ldots \cdot (x - x_{j-1}) \cdot (x - x_{j+1}) \cdot \ldots \cdot (x - x_n)}{(x_j - x_1) \cdot \ldots \cdot (x_j - x_{j-1}) \cdot (x_j - x_{j+1}) \cdot \ldots \cdot (x_j - x_n)}, \tag{4.22}$$

называемый *интерполяционным многочленом в форме Лагранжа*, является интерполяционным для таблицы (4.20).

Единственность. Пусть $f(x)$, $g(x)$ — два интерполционных многочлена, соответствующих одной таблице интерполяции (4.20). Рассмотрим многочлен $h(x) = f(x) - g(x)$. Имеем $h(x_j) = 0$ $(j = 0, 1, \dots, n)$. Таким образом, многочлен $h(x)$ степени, не превосходящей n, имеет не менее $n + 1$ корней. По следствию 4.38 получаем, что $h(x) = 0$, т. е. $f(x) = g(x)$. □

Пример 4.40. Построим интерполяционный многочлен по таблице

x_j	1	2	3
y_j	-6	-6	-4

По формуле (4.22) имеем

$$f(x) = -6 \cdot \frac{(x-2)(x-3)}{(1-2)(1-3)} - 6 \cdot \frac{(x-1)(x-3)}{(2-1)(2-3)} - 4 \cdot \frac{(x-1)(x-2)}{(3-1)(3-2)} = x^2 - 3x - 4.$$

Следствие 4.41. *Пусть F — некоторое подполе поля $\mathbb{C}$. Для любых многочленов $f(x)$ и $g(x)$ из $F[x]$*

$$f(x) = g(x) \qquad \Leftrightarrow \qquad \forall c \in F \quad f(c) = g(c).$$

Доказательство. Если $f(x) = g(x)$, то, очевидно, $f(c) = g(c)$ для всех $c \in F$. Обратно, пусть $f(c) = g(c)$ для всех $c \in F$, но $f(x) \neq g(x)$. Это противоречит теореме 4.39. Действительно, положим $n = \max\{\deg f(x),\ \deg g(x)\}$ и рассмотрим в F $n + 1$ попарно различных чисел $x_0, x_1, \dots, x_n$. Интерполяционный многочлен степени, не превосходящей n, принимающий в точках x_j значения $f(x_j) = g(x_j)$ существует и единственен, поэтому $f(x) = g(x)$. □

Замечание 4.42. Мы показали, что два многочлена, рассматриваемых с алгебраической точки зрения: т. е. как алгебраические выражения, равны тогда и только тогда, когда совпадают соответствующие им функции. Таким образом, «алгебраическая» точка зрения на многочлены как на формальные выражения совпадает с «функциональной» точкой зрения на них как на функции $f : F \to F$, если F — числовое поле.

Рассмотрим еще один способ нахождения интерполяционного многочлена — *метод Ньютона*. Интерполяционный многочлен для таблицы (4.20) ищется в виде

$$f(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \ldots + c_n(x - x_0)(x - x_1)\ldots(x - x_{n-1}). \qquad (4.23)$$

Последовательно полагая $x = x_j$ $(j = 0, 1, \ldots, n)$, находим $c_0, c_1, \ldots c_n$.

Пример 4.43. Методом Ньютона построим интерполяционный многочлен по таблице интерполяции из примера 4.40. Ищем многочлен в виде

$$f(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1). \qquad (4.24)$$

Полагая в (4.24) $x = x_0$, получаем $-6 = c_0$.
Полагая $x = x_1$, получаем $-6 = -6 + c_1(2 - 1)$, откуда $c_1 = 0$.
Полагая $x = x_2$, получаем $-4 = -6 + 0 \cdot (2 - 1) + c_2(3 - 1)(3 - 2)$, откуда $c_2 = 1$.
Итак, $f(x) = -6 + (x - 1)(x - 2) = x^2 - 3x - 4$.

Для эффективного вычисления коэффициентов интерполяционного многочлена $f(x)$ в форме Ньютона введем так называемые разделенные разности. Пусть $f(x)$ — интерполяционный многочлен, построенный по таблице (4.20), а $z_0, z_1, \ldots, z_n$ — некоторые числа из F. *Разделенной разностью первого порядка* называется величина

$$f(z_0, z_1) = \frac{f(z_0) - f(z_1)}{z_0 - z_1},$$

разделенной разностью второго порядка называется

$$f(z_0, z_1, z_2) = \frac{f(z_0, z_1) - f(z_1, z_2)}{z_0 - z_2}$$

и т. д. Вообще, *разделенная разность k-го порядка* определяется через разделенную разность $(k - 1)$-го порядка следующим образом:

$$f(z_0, z_1, \ldots, z_k) = \frac{f(z_0, z_1, \ldots, z_{k-1}) - f(z_1, z_2, \ldots, z_k)}{z_0 - z_k}.$$

Заметим, что так как $f(x_0) = y_0$, т. е. x_0 является корнем многочлена $f(x) - y_0$, то $f(x, x_0)$ представляет собой многочлен и степень этого многочлена на 1 меньше степени $f(x)$. Аналогично, $f(x, x_0, x_1)$ также является многочленом и степень его на 1 меньше степени $f(x, x_0)$ и т. д. Наконец, $f(x, x_0, x_1, \ldots, x_n) = 0$.

Из определения разделенных разностей получаем

$$f(x) = f(x_0) + (x - x_0)f(x, x_1),$$

$$f(x, x_0) = f(x_0, x_1) + (x - x_1)f(x, x_0, x_1),$$
$$f(x, x_0, x_1) = f(x_0, x_1, x_2) + (x - x_2)f(x, x_0, x_1, x_2)$$

и т. д., откуда

$$f(x) = f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \ldots$$
$$\ldots + (x - x_0)(x - x_1)\ldots(x - x_{n-1})f(x_0, x_1, \ldots, x_n).$$

В силу единственности представления многочлена $f(x)$ в виде (4.23), получаем, что коэффициенты c_j в (4.23) суть разделенные разности, а именно:

$$c_j = f(x_0, x_1, \ldots, x_j) \qquad (j = 0, 1, \ldots, n).$$

Вычислять разделенные разности удобно в таблице следующего вида:

$$\begin{array}{c|lllll}
x_0 & f(x_0)=y_0 \\
 & & f(x_0,x_1) \\
x_1 & f(x_1)=y_1 & & f(x_0,x_1,x_2) \\
 & & f(x_1,x_2) & & \ddots \\
x_2 & f(x_2)=y_2 & & & & f(x_0,x_1,\dots,x_n) \\
\vdots & \quad\vdots \\
 & & f(x_{n-1},x_n) \\
x_n & f(x_n)=y_n
\end{array}$$

Пример 4.44. Построим интерполяционный многочлен $f(x)$ по таблице

x_j	-2	-1	0	1	2
y_j	3	8	17	24	47

Составим таблицу разделенных разностей:

$$\begin{array}{c|ccccc}
-2 & \mathbf{3} \\
 & & \mathbf{5} \\
-1 & 8 & & \mathbf{2} \\
 & & 9 & & \mathbf{-1} \\
0 & 17 & & -1 & & \mathbf{1} \\
 & & 7 & & 3 \\
1 & 24 & & 8 \\
 & & 23 \\
2 & 47
\end{array}$$

Таким образом,

$$f(x) = (x+2)(x+1)x(x-1) - (x+2)(x+1)x + 2(x+2)(x+1) + 5(x+2) + 3.$$

Для раскрытия скобок в (4.23), т. е. нахождения самих коэффициентов интерполяционного многочлена, существует эффективная процедура. Действительно, рассматривая (4.23), заметим, что коэффициент c_0 равен остатку от деления $f(x)$ на $x - x_0$; коэффициент c_1 равен остатку от деления полученного на предыдущем шаге частного на $x - x_1$ и т. д. Таким образом, для раскрытия скобок в (4.23) можно применить схему, аналогичную методу, проиллюстрированному в примере 4.32 (1-й способ).

Пример 4.45. Для многочлена $f(x)$ из примера 4.44 составим следующую таблицу, которую будем заполнять снизу вверх:

	1	**1**	**−2**	**7**	**17**
-2	1	-1	0	7	3
-1	1	-2	2	5	
0	1	-2	2		
1	1	-1			
2	1				

Итак, $f(x) = x^4 + x^3 - 2x^2 + 7x + 17$.

4.8. Формулы Виета

Рассмотрим задачу нахождения коэффициентов многочлена по его корням. Пусть

$$f(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = (x - c_1)(x - c_2)\ldots(x - c_n).$$

Раскрывая скобки в правой части и собирая коэффициенты при каждой степени, получаем *формулы Виета*:

$$\begin{gathered}
a_1 = -(c_1 + c_2 + \ldots + c_n),\\
a_2 = c_1c_2 + c_1c_3 + \ldots + c_{n-1}c_n,\\
a_3 = -(c_1c_2c_3 + c_1c_2c_4 + \ldots + c_{n-2}c_{n-1}c_n),\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_j = (-1)^j \sum_{1\le i_1<i_2<\ldots<i_j\le n} a_{i_1}a_{i_2}\ldots a_{i_j}, \qquad (4.25)\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
a_{n-1} = (-1)^{n-1}(c_1c_2\ldots c_{n-1} + c_1\ldots c_{n-2}c_n + \ldots + c_2c_3\ldots c_n),\\
a_n = (-1)^n c_1c_2\ldots c_n.
\end{gathered}$$

Пример 4.46. Найдем многочлен $f(x)$ минимальной степени со старшим коэффициентом 1, у которого 1, 2 — простые корни, а 3 — корень кратности 2. Очевидно этим многочленам является $f(x) = (x-1)(x-2)(x-3)^2$. По формулам (4.25) получаем

$$\begin{gathered}
a_1 = -(1+2+3+3) = -9,\\
a_2 = 1\cdot 2 + 1\cdot 3 + 1\cdot 3 + 2\cdot 3 + 2\cdot 3 + 3\cdot 3 = 29,\\
a_3 = -(1\cdot 2\cdot 3 + 1\cdot 2\cdot 3 + 1\cdot 3\cdot 3 + 2\cdot 3\cdot 3) = -39,\\
a_4 = 1\cdot 2\cdot 3\cdot 3 = 18.
\end{gathered}$$

Поэтому $f(x) = x^4 - 9x^3 + 29x^2 - 39x + 18$.

4.9. Многочлены с действительными коэффициентами

Утверждение 4.47. *Пусть $f(x) \in \mathbb{R}[x]$. Если $c \in \mathbb{C}$ — корень многочлена $f(x)$, то $\overline{c}$ также является корнем этого многочлена и имеет ту же кратность.*

Доказательство. Пусть $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$. Так как $f(c) = 0$, то

$$a_0c^n + a_1c^{n-1} + \ldots + a_{n-1}c + a_n = 0,$$

откуда

$$\overline{a_0c^n + a_1c^{n-1} + \ldots + a_{n-1}c + a_n} = \overline{0},$$

поэтому

$$\overline{a}_0\overline{c}^n + \overline{a}_1\overline{c}^{n-1} + \ldots + \overline{a}_{n-1}\overline{c} + \overline{a}_n = 0.$$

Но $a_j \in \mathbb{R}$ $(j = 1, 2, \ldots, n)$, следовательно, $\overline{a}_j = a_j$, поэтому

$$a_0\overline{c}^n + a_1\overline{c}^{n-1} + \ldots + a_{n-1}\overline{c} + a_n = 0,$$

т. е. $f(\overline{c}) = 0$.

Чтобы показать, что корень $\overline{c}$ имеет ту же кратность, что и c достаточно применить те же рассуждения к производным многочлена $f(x)$ и воспользоваться следствием 4.29. □

Следствие 4.48. *Для любого многочлена*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{R}[x] \quad (a_0 \neq 0)$$

найдутся числа $c_1, c_2, \ldots, c_s \in \mathbb{R}$, $p_1, p_2, \ldots, p_t \in \mathbb{R}$, $q_1, q_2, \ldots, q_t \in \mathbb{R}$, $k_1, k_2, \ldots, k_s \in \mathbb{N}$, $l_1, l_2, \ldots, l_t \in \mathbb{N}$, *такие, что*

$$f(x) = a_0(x - c_1)^{k_1} \ldots (x - c_s)^{k_s} (x^2 + p_1x + q_1)^{l_1} \ldots (x^2 + p_tx + q_t)^{l_t}, \qquad (4.26)$$

где

$$k_1 + \ldots + k_s + 2l_1 + \ldots + 2l_t = n,$$
$$c_i \neq c_j \quad (i \neq j;\ i, j = 1, 2, \ldots, s),$$
$$x^2 + p_ix + q_i \neq x^2 + p_jx + q_j \quad (i \neq j;\ i, j = 1, 2, \ldots, t),$$

причем многочлены $x^2 + p_jx + q_j$ *вещественных корней не имеют. Для любого многочлена* $f(x) \in \mathbb{R}[x]$ *представление вида* (4.26) *с указанными свойствами единственно с точностью до перестановки множителей.*

Доказательство. Существование. Согласно следствию 4.36 и утверждению 4.47

$$f(x) = a_0(x - c_1)^{k_1} \ldots (x - c_s)^{k_s} (x - d_1)^{l_1} (x - \overline{d}_1)^{l_1} \ldots (x - d_t)^{l_t} (x - \overline{d}_t)^{l_t},$$

где $c_1, c_2, \ldots, c_s \in \mathbb{R}$, $d_1, d_2, \ldots, d_t \in \mathbb{C}$. Теперь разложение (4.26) следует из равенства

$$(x - d_j)(x - \overline{d}_j) = x^2 + p_jx + q_j,$$

в котором $p_j = -d_j - \overline{d}_j = -2\operatorname{Re} d_j \in \mathbb{R}$, $q_j = d_j\overline{d}_j = |d_j|^2 \in \mathbb{R}$.

Единственность. Чтобы по представлению (4.26) получить разложение на линейные множители (4.16) достаточно разложить на линейные множители каждый из квадратных многочленов $x^2 + p_jx + q_j$. Каждый из этих многочленов имеет вещественные коэффициенты и пару комплексно сопряженных корней. Если бы для одного и того же многочлена существовали различные разложения вида (4.26) с указанными свойствами, то мы из них получили бы различные представления (4.16), что противоречит следствию 4.36. □

Пример 4.49. Найдем разложение вида (4.26) для многочлена $f(x) = x^4 + 4$. Корнями этого многочлена являются все значения корня 4-й степени из 4, т.е. $\pm 1 \pm i$, поэтому разложение на линейные множители (4.16) имеет вид $f(x) = (x - 1 - i)(x - 1 + i)(x + 1 - i)(x + 1 + i)$. Перемножая множители, соответствующие комплексно сопряженным корням, получаем требуемое разложение $f(x) = (x^2 - 2x + 2)(x^2 + 2x + 2)$.

Следствие 4.50. *Произвольный многочлен нечетной степени с вещественными коэффициентами имеет по крайней мере один вещественный корень.*

Доказательство. По следствию 4.48, так как степень многочлена нечетна, то в разложении вида (4.26) обязательно найдется по крайней мере один линейный множитель. □

4.10. Неприводимые многочлены

Пусть F — поле. Многочлен $f(x) \in F[x]$ ненулевой степени называется *неприводимым* над полем F, если он не имеет делителей степени большей 1. В противном случае $f(x)$ назвается *приводимым*. Очевидно, любой многочлен степени 1 неприводим. Далее, если $f(x)$ неприводим, то и $cf(x)$ неприводим для любой константы c.

Замечание 4.51. Обращаем внимание, что свойство многочлена быть неприводимым зависит от того, над каким полем рассматривается этот многочлен. Например, многочлен $f(x) = x^2 + 1 = (x - i)(x + i)$, очевидно, неприводим над $\mathbb{R}$ и $\mathbb{Q}$, но приводим над $\mathbb{C}$. Многочлен $f(x) = x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ неприводим над $\mathbb{Q}$, но приводим над $\mathbb{C}$ и $\mathbb{R}$.

Легко видеть, что если многочлен степени выше 1 имеет корень из F, то он приводим над F. Обратное в общем случае не верно. Например, многочлен $x^4 + 4$ ни рациональных, ни даже вещественных корней не имеет, но раскладывается на рациональные множители, т.е. является приводимым над $\mathbb{Q}$ и $\mathbb{R}$ (см. пример 4.49).

Замечание 4.52. Понятие неприводимого многочлена можно сфорулировать не над полем, а над кольцом. В дальнейшем, например, нам понадобится понятие неприводимого многочлена над кольцом $\mathbb{Z}$. Заметим однако, что многие свойства, устанавливаемые в этом разделе для неприводимых над полем многочленов, не верны для многочленов, неприводимых над кольцом.

Следствие 4.53. *Неприводимыми над $\mathbb{C}$ многочленами являются все многочлены первой степени из $\mathbb{C}[x]$ и только они. Неприводимыми над $\mathbb{R}$ многочленами являются все многочлены первой степени из $\mathbb{R}[x]$ и все многочлены второй степени из $\mathbb{R}[x]$ с отрицательным дискриминантом и только они.*

Доказательство. Первая часть утверждения вытекает из следствия 4.36. Вторая часть утверждения вытекает из следствия 4.48. □

Утверждение 4.54. *Пусть $f(x)$ — произвольный, а $h(x)$ — неприводимый многочлен. Тогда $f(x) \,\vdots\, h(x)$ или многочлены $f(x)$ и $h(x)$ взаимно просты.*

Доказательство. Пусть $d(x)$ — НОД многочленов $f(x)$ и $h(x)$. Если $\deg d(x) = 1$, то $f(x)$ и $h(x)$ взаимно просты. Пусть $\deg d(x) > 1$. Так как $h(x) \,\vdots\, d(x)$, а $h(x)$ — неприводимый, то $h(x) = cd(x)$, где c — ненулевая константа. Так как $f(x) \,\vdots\, d(x)$, то $f(x) \,\vdots\, h(x)$. □

Утверждение 4.55. *Пусть $f(x)g(x) \,\vdots\, h(x)$, причем $h(x)$ — неприводимый. Тогда $f(x) \,\vdots\, h(x)$ или $g(x) \,\vdots\, h(x)$.*

Доказательство. Покажем, что если $f(x) \not\vdots\, h(x)$, то $g(x) \,\vdots\, h(x)$. Действительно, по утверждению 4.54, если $f(x) \not\vdots\, h(x)$, то $f(x)$ и $h(x)$ взаимно просты. Теперь $g(x) \,\vdots\, h(x)$ следует из утверждения 4.19. □

Теорема 4.56. *Для любого многочлена $f(x) \in F[x]$ ненулевой степени существуют такие неприводимые многочлены $p_1(x), p_2(x), \ldots, p_s(x)$, что*

$$f(x) = p_1(x)p_2(x)\ldots p_s(x). \tag{4.27}$$

Разложение вида (4.27) единственно с точностью до константных множителей и перестановки многочленов $p_j(x)$.

Доказательство. Существование. Если $f(x)$ — неприводимый, то разложение получено. В противном случае найдутся многочлены $g_1(x)$ и $g_2(x)$ степени меньшей $\deg f(x)$, к которым применим те же рассуждения, и т. д. (можно применить индукцию по степени многочлена), пока не получим требуемого разложения. *Единственность.* Покажем, что если

$$p_1(x)p_2(x)\dots p_s(x) = q_1(x)q_2(x)\dots q_t(x) \tag{4.28}$$

— два разложения $f(x)$ на неприводимые множители, то $s = t$ и при соответствующей нумерации $p_j(x) = c_j q_j(x)$, где c_j — ненулевая константа $(j = 1, 2, \dots, s)$.

Из (4.28) следует, что $p_1(x)p_2(x)\dots p_s(x) \,\vdots\, q_1(x)$. Так как $q_1(x)$ — неприводимый, то по утверждению 4.55 найдется j, такое, что $p_j \,\vdots\, q_1$. Не нарушая общности, будем считать, что $j = 1$, т. е. $p_1(x) \,\vdots\, q_1(x)$. Так как $p_1(x)$ неприводим, то $p_1(x) = c_1 q_1(x)$ для некоторой ненулевой константы c_1. Переобозначим $p_2(x) = c_1 p_2(x)$. Сокращая по лемме 4.1 равенство на $q_1(x)$, приходим к равенству

$$p_2(x)p_3(x)\dots p_s(x) = q_2(x)q_3(x)\dots q_t(x),$$

к которому применяем те же рассуждения, и т. д. (можно применить индукцию), пока не получим $p_s(x) = q_s(x)$. □

Следствие 4.57. *Для любого многочлена $f(x) \in F[x]$ ненулевой степени существуют неприводимые многочлены $p_1(x), p_2(x), \dots, p_s(x)$ и натуральные числа $k_1, k_2, \dots, k_s$, такие, что*

$$f(x) = p_1^{k_1}(x)p_2^{k_2}(x)\dots p_s^{k_s}(x), \tag{4.29}$$

$p_i(x) \not\vdots\, p_j(x)$ $(i \neq j)$. Разложение вида (4.29) с указанными свойствами единственно с точностью до константных множителей и перестановки многочленов $p_j(x)$.

Запись вида (4.29) называется разложением многочлена $f(x)$ на неприводимые множители. *Величина k_j называется* кратностью *множителя $p_j(x)$.*

Следствие 4.58. *Для любого многочлена $f(x) \in F[x]$ ненулевой степени существуют неприводимые многочлены $p_1(x), p_2(x), \dots, p_s(x)$ со старшим коэффициентом 1 и натуральные числа $k_1, k_2, \dots, k_s$, такие, что*

$$f(x) = a_0 p_1^{k_1}(x)p_2^{k_2}(x)\dots p_s^{k_s}(x), \tag{4.30}$$

$p_i(x) \neq p_j(x)$ $(i \neq j)$, где a_0 — коэффициент при старшем члене многочлена $f(x)$. Разложение вида (4.30) с указанными свойствами единственно с точностью до перестановки множителей.

Замечание 4.59. Следствия 4.36 и 4.48 являются частными случаями только что установленного следствия 4.58. Разложением многочлена из $\mathbb{C}[x]$ на неприводимые над $\mathbb{C}$ множители является разложение на линейные множители (4.16). Разложением многочлена из $\mathbb{R}[x]$ на неприводимые над $\mathbb{R}$ множители является разложение (4.16).

Следствие 4.60. *Пусть* (4.29) — *разложение многочлена* $f(x) \in F[x]$ *на неприводимые множители. Тогда для любого делителя* $d(x)$ *этого многочлена найдутся такие целые неотрицательные* $l_1, l_2, \ldots, l_s$ *и такое* $c \in F$, *что*

$$d(x) = c\, p_1^{l_1}(x) p_2^{l_2}(x) \ldots p_s^{l_s}(x).$$

Доказательство. Если $d(x)$ — произвольный делитель многочлена $f(x)$, то для некоторого $q(x)$ имеем $f(x) = q(x)d(x)$, поэтому в разложении многочлена $d(x)$ на неприводимые могут встретиться только многочлены $c_1 p_1(x), c_2 p_2(x), \ldots, c_s p_s(x)$, где c_j — константы ($j = 1, 2, \ldots, s$), так как противное противоречило бы единственности разложения $f(x)$. □

Теорема 4.61. *Пусть*

$$\begin{aligned} f(x) &= p_1^{k_1}(x) p_2^{k_2}(x) \ldots p_t^{k_t}(x)\ \ p_{t+1}^{k_{t+1}}(x) p_{t+2}^{k_{t+2}}(x) \ldots p_s^{k_s}(x), \\ g(x) &= p_1^{l_1}(x) p_2^{l_2}(x) \ldots p_t^{l_t}(x)\ \ p_{s+1}^{l_{s+1}}(x) p_{s+2}^{l_{s+2}}(x) \ldots p_r^{l_r}(x), \end{aligned}$$

и многочлены $p_j(x)$ *неприводимы,* $p_i(x) \not\vdots\, p_j(x)$ $(i \neq j;\ i, j = 1, 2, \ldots, r)$, $m_j = \min\left\{k_j,\ l_j\right\}$ $(j = 1, 2, \ldots, t)$. *Тогда*

$$d(x) = p_1^{m_1}(x) p_2^{m_2}(x) \ldots p_t^{m_t}(x)$$

является наибольшим общим делителем многочленов $f(x)$ *и* $g(x)$.

Доказательство. Очевидно, $d(x)$ является делителем как многочлена $f(x)$, так и многочлена $g(x)$ и по следствию 4.63 любой их общий делитель является делителем многочлена $d(x)$. □

Теорема 4.62. *Неприводимый множитель кратности* k *многочлена* $f(x)$ *является неприводимым множителем кратности* $k-1$ *многочлена* $f'(x)$. *Простой неприводимый множитель многочлена* $f(x)$ *не входит в разложение многочлена* $f'(x)$ *на неприводимые.*

Доказательство. Аналогично доказательству теоремы 4.28. □

Следствие 4.63. *Пусть* (4.29) — *разложение многочлена* $f(x) \in F[x]$ *на неприводимые множители. Тогда многочлен*

$$d(x) = p_1^{k_1-1}(x) p_2^{k_2-1}(x) \ldots p_s^{k_s-1}(x) \tag{4.31}$$

является наибольшим общим делителем многочленов $f(x)$ *и* $f'(x)$ *и, следовательно,*

$$\frac{f(x)}{d(x)} = p_1(x) p_2(x) \ldots p_s(x). \tag{4.32}$$

Доказательство. Формула (4.31) следует из теорем 4.61 и 4.62. Соотношение 4.32 непосредственно вытекает из (4.29) и (4.31). □

Замечание 4.64. Следствие 4.63 дает алгоритм *избавления от кратных множителей*, т. е. алгоритм построения по заданному многочлену $f(x)$ многочлена $g(x)$ с теми же множителями в разложении на неприводимые, но кратности 1. Заметим, что $g(x)$ имеет те же корни, что и $f(x)$, но все корни $g(x)$ — простые. В некоторых случаях это позволяет определить все корни многочлена $f(x)$.

Пример 4.65. Избавимся от кратных множителей в многочлене $f(x) = x^5 + 7x^4 + 10x^3 - 18x^2 - 27x + 27$. Имеем $f'(x) = 5x^4 + 28x^3 + 30x^2 - 36x - 27$. НОД многочленов $f(x)$ и $f'(x)$ (найденный алгоритмом Евклида) равен $d(x) = -9+x^3+5x^2+3x$, откуда $f(x)/d(x) = x^2+2x-3 = (x-1)(x+3)$. Последовательным делением $f(x)$ на $x-1$ и $x+3$ (по схеме Горнера) определим кратность корней $1, -3$ в многочлене $f(x)$. Получим $f(x) = (x-1)^2(x+3)^3$.

Рассмотрим еще один способ, позволяющий иногда найти все корни многочлена и сразу указать их кратности. Пусть $d_1(x)$ — НОД многочлена $f(x)$ и его производной $f'(x)$; $d_2(x)$ — НОД многочлена $d_1(x)$ и его производной $d_1'(x)$; $d_3(x)$ — НОД многочленов $d_2(x)$ и $d_2'(x)$ и т. д. до тех пор, пока не будет получен многочлен $d_s(x)$ нулевой степени. Положим

$$g_1(x) = \frac{f(x)}{d_1(x)}, \quad g_2(x) = \frac{d_1(x)}{d_2(x)}, \quad \ldots, \quad g_s(x) = \frac{d_{s-1}(x)}{d_s(x)}$$

и далее

$$f_1(x) = \frac{g_1(x)}{g_2(x)}, \quad f_2(x) = \frac{g_2(x)}{g_3(x)}, \quad \ldots, \quad f_s(x) = g_s(x).$$

Легко проверить, что все корни многочленов $f_1(x), f_2(x), \ldots, f_s(x)$ простые, причем корнями многочлена $f_j(x)$ являются все корни кратности j многочлена $f(x)$ и только они ($j = 1, 2, \ldots, s$). Корней кратности большей s у $f(x)$ нет.

Пример 4.66. Для многочлена $f(x) = x^7 + 2x^6 + 3x^5 + x^4 - x^3 - 3x^2 - 2x - 1$ получаем

$$d_1(x) = x^4 + 2x^3 + 3x^2 + 2x + 1, \quad d_2(x) = x^2 + x + 1, \quad d_3(x) = 1,$$

далее

$$g_1(x) = x^3 - 1, \quad g_2(x) = x^2 + x + 1 \quad , g_3(x) = x^2 + x + 1,$$

откуда

$$f_1(x) = x - 1, \quad f_2(x) = 1, \quad f_3(x) = x^2 + x + 1 = \left(x + \frac{1}{2} - \frac{\sqrt{3}}{2}i\right) \cdot \left(x + \frac{1}{2} + \frac{\sqrt{3}}{2}i\right).$$

Итак, многочлен $f(x)$ имеет один простой корень 1 и два трехкратных корня $-\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$.

4.11. Многочлены с рациональными коэффициентами

В данном разделе рассматриваются три взаимосвязанные задачи: проблема отыскания рациональных корней многочлена с рациональными коэффициентами; задача установления неприводимости над $\mathbb{Q}$ заданного многочлена и задача разложения заданного многочлена на неприводимые над $\mathbb{Q}$ многочлены. Оказывается, достаточно научиться решать эти задачи не над полем $\mathbb{Q}$, а над кольцом $\mathbb{Z}$.

Очевидно, произвольный многочлен $f(x) \in \mathbb{Q}[x]$ домножением его коэффициентов на НОК их знаменателей можно превратить в многочлен $g(x) \in \mathbb{Z}[x]$. При этом все корни многочлена $f(x)$ являются корнями многочлена $g(x)$ и наоборот. Таким образом, задача отыскания корней многочлена с рациональными коэффициентами сводится к задаче отыскания корней многочлена с целыми коэффициентами. Кроме того, легко видеть, что $f(x)$ неприводим над $\mathbb{Q}$ тогда и только тогда, когда наприводим над $\mathbb{Q}$ многочлен $g(x)$. Далее мы установим более сильный результат: многочлен $f(x)$ неприводим над $\mathbb{Q}$ тогда и только тогда, когда $g(x)$ неприводим над $\mathbb{Z}$ и, следовательно, задачи проверки неприводимости над $\mathbb{Q}$ и нахождения разложения на неприводимые над $\mathbb{Q}$ сведена к соответствующим задачам над $\mathbb{Z}$.

Приведем следующее необходимое условие на рациональные корни многочлена с целыми коэффициентами.

Теорема 4.67. *Пусть*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{Z}[x],$$

$p, q \in \mathbb{Z}$, $p \neq 0$, $q \neq 0$, $a_0 \neq 0$, НОД$(p, q) = 1$, $f(p/q) = 0$, *тогда*

1) $a_0 \,\vdots\, q$,

2) $a_n \,\vdots\, p$,

3) $f(m) \,\vdots\, (p - mq)$ *для любого* $m \in \mathbb{Z}$.

Доказательство. 1), 2) Так как $f(p/q) = 0$, то

$$f\left(\frac{p}{q}\right) = a_0 \cdot \left(\frac{p}{q}\right)^n + a_1 \cdot \left(\frac{p}{q}\right)^{n-1} + \ldots + a_{n-1} \cdot \frac{p}{q} + a_n = 0.$$

Домножая обе части полученного равенства на q^n, получаем

$$a_0p^n + a_1p^{n-1}q + \ldots + a_{n-1}pq^{n-1} + a_nq^n = 0. \tag{4.33}$$

Слагаемые в левой части равенства (4.33) со второго до последнего делятся на q, следовательно, первое слагаемое a_0p^n также делится на q, но так как НОД$(p, q) = 1$, то $a_0 \,\vdots\, q$. Аналогично, слагаемые с первого до предпоследнего в левой части равенства (4.33) делятся на p, следовательно, последнее слагаемое a_nq^n также делится на p, но так как НОД$(p, q) = 1$, то $a_n \,\vdots\, p$.

3) Разделим $f(x)$ на $x - m$. Получим в частном $q(x)$, а в остатке $f(m)$:

$$f(x) = (x - m)g(x) + f(m). \tag{4.34}$$

Заметим, что $g(x) \in \mathbb{Z}[x]$. Из (4.34) при $x = p/q$ получаем $0 = (p/q - m)g(p/q) + f(m)$. Домножаем обе части этого равенства на q^n и применяем рассуждения, аналогичные пп. 1), 2).

□

Теорема 4.67 позволяет предложить следующий метод нахождения всех рациональных корней многочлена $f(x) \in \mathbb{Z}[x]$. Для всех делителей p коэффициента a_n и всех делителей q коэффициента a_0, если НОД$(p, q) = 1$, проверим (например, с помощью схемы Горнера) равенство $f(p/q) = 0$. Проверку целесообразно начать с $q = 1$ и всех p, делящих a_n: полученные в результате них значения $f(m)$, где $m = p/q \in \mathbb{Z}$, можно использовать для отсеивания дальнейших пробных p и q с помощью условия $f(m) \,\vdots\, (p - mq)$. Как только один из корней p/q найден, многочлен $f(x)$ можно разделить на $x - p/q$ и применим ту же процедуру к частному.

Пример 4.68. Найдем все рациональные корни многочлена $f(x) = 6x^6 - 13x^5 - 4x^4 + 24x^3 - 21x^2 + 4x + 10$. Делителями свободного коэффициента являются числа ± 1, ± 2, ± 5, ± 10. Положительными делителями старшего коэффициента являются числа 1, 2, 3, 6. Таким образом, рациональными корнями многочлена $f(x)$ могут являться только следующие числа:

$$\pm 1,\ \pm 2,\ \pm 5,\ \pm 10,\ \pm\frac{1}{2},\ \pm\frac{5}{2},\ \pm\frac{1}{3},\ \pm\frac{2}{3},\ \pm\frac{5}{3},\ \pm\frac{10}{3},\ \pm\frac{1}{6},\ \pm\frac{5}{6}.$$

Так как $f(1) = 6$, то отсеиваются все числа p/q, для которых $6 \not\vdots\ p - q$, т. е. отсеиваем

$$5,\ 10,\ \frac{10}{3},\ \frac{1}{6},\ -10,\ -\frac{5}{2},\ -\frac{1}{3},\ -\frac{5}{3},\ -\frac{10}{3},\ -\frac{1}{6},\ -\frac{5}{6}.$$

Так как $f(-1) = -24$, то отсеиваются все числа p/q, для которых $-24 \not\vdots\ p + q$, т. е. из оставшихся отсеиваем

$$\frac{5}{2},\ \frac{2}{3},\ \frac{5}{6}.$$

В списке возможных кандидатов остались

$$2,\ \frac{1}{2},\ \frac{1}{3},\ \frac{5}{3},\ -2,\ -5,\ -\frac{1}{2},\ -\frac{2}{3}.$$

Схемой Горнера устанавливая значение многочлена $f(x)$ для каждого из этих чисел, получаем, что его рациональными корнями являются только $5/3$ и $-1/2$.

Рассмотрим теперь задачи определения неприводимости многочлена над $\mathbb{Q}$ и разложения заданного многочлена на неприводимые.

Ненулевой многочлен $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{Z}[x]$ называется *примитивным*, если его коэффициенты $a_0, a_1, \ldots, a_n$ взаимно просты в совокупности.

Утверждение 4.69 (Лемма Гаусса). *Произведение примитивных многочленов является примитивным многочленом.*

Доказательство. Пусть

$$\begin{aligned} f(x) = g(x)h(x) &= a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n, \\ g(x) &= b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m, \\ h(x) &= c_0x^k + c_1x^{k-1} + \ldots + c_{k-1}x + c_k \end{aligned}$$

и многочлены $g(x)$ и $h(x)$ примитивны. Поэтому для любого простого p найдутся такие s и t, что

$$b_i \vdots\ p \quad (i = 0, 1, \ldots, s-1), \quad b_s \not\vdots\ p, c_j \vdots\ p \quad (j = 0, 1, \ldots, t-1), \quad c_t \not\vdots\ p.$$

Докажем, что $a_{s+t} \not\vdots\ p$, что докажет примитивность многочлена $f(x)$. Имеем

$$a_{s+t} = \sum_{i+j=s+t} b_ic_j = \ldots + b_{s-2}c_{t+2} + b_{s-1}c_{t+1} + b_sc_t + b_{s+1}c_{t-1} + b_{s+2}c_{t-2} + \ldots$$

Каждое из слагаемых $b_{s-1}c_{t+1}$, $b_{s-2}c_{t+2}$ и т. д. делится на p, так как $b_i \vdots\ p$ при $i < s$. Каждое из слагаемых $b_{s+1}c_{t-1}$, $b_{s+2}c_{t-2}$ и т. д. делится на p, так как $c_j \vdots\ p$ при $j < t$. Но $b_sc_t \not\vdots\ p$. Следовательно, $a_{s+t} \not\vdots\ p$. □

Теорема 4.70. *Если многочлен $f(x) \in \mathbb{Z}[x]$ допускает разложение на многочлены $g(x) \in \mathbb{Q}[x]$ и $h(x) \in \mathbb{Q}[x]$, то $f(x)$ допускает также разложение на многочлены $cg(x) \in \mathbb{Z}[x]$ и $c_1h(x) \in \mathbb{Z}[x]$, где c, c_1 — некоторые константы.*

Доказательство. Очевидно, найдутся такие целые p, q, p_1, q_1 что

$$g(x) = \frac{p}{q}\widehat{g}(x), \quad h(x) = \frac{p_1}{q_1}\widehat{h}(x), \quad \widehat{g}(x) \in \mathbb{Z}[x], \quad \widehat{h}(x) \in \mathbb{Z}[x],$$

НОД$(p,q) = 1$, НОД$(p_1,q_1) = 1$, причем $\widehat{g}(x)$, $\widehat{h}(x)$ — примитивные. Таким образом,

$$f(x) = g(x)h(x) = \frac{pp_1}{qq_1}\widehat{g}(x)\widehat{h}(x).$$

Так как $f(x) \in \mathbb{Z}[x]$, то qq_1 должно сократиться с $pp_1\widehat{g}(x)\widehat{h}(x)$. Но по утверждению 4.69 многочлен $\widehat{g}(x)\widehat{h}(x)$ примитивен, поэтому знаменатель qq_1 должен сократиться с pp_1, т. е. $(pp_1)/(qq_1) \in \mathbb{Z}$. Полагая, например, $c = p_1/q_1$, $c_1 = q_1/p_1$, получаем $f(x) = \big(cg(x)\big)\big(c_1h(x)\big)$, причем

$$cg(x) = \frac{pp_1}{qq_1}\widehat{g}(x) \in \mathbb{Z}[x], \quad c_1h(x) = \widehat{h}(x) \in \mathbb{Z}[x].$$

$\square$

Следствие 4.71. *Многочлен $f(x) \in \mathbb{Z}[x]$ неприводим над $\mathbb{Q}$ тогда и только тогда, когда он неприводим над $\mathbb{Z}$, т. е. не допускает разложения на многочлены ненулевой степени из $\mathbb{Z}[x]$.*

Приведенное следствие существенно облегчает доказательство неприводимости многочленов над полем $\mathbb{Q}$.

Приведем одно из известных достаточных условий неприводимости многочлена над $\mathbb{Q}$.

Теорема 4.72 (Признак Эйзенштейна)**.** *Пусть*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \qquad (a_0 \neq 0)$$

— примитивный многочлен, причем существует простое p, такое, что

$$a_0 \not\vdots\, p, \qquad a_j \,\vdots\, p \quad (j = 1, 2, \ldots, n), \qquad a_n \not\vdots\, p^2,$$

тогда $f(x)$ неприводим над $\mathbb{Z}$ и, следовательно, над $\mathbb{Q}$.

Доказательство. Предположим противное. Пусть $f(x) = g(x)h(x)$, где

$$g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m \in \mathbb{Z}[x],$$

$$h(x) = c_0x^k + c_1x^{k-1} + \ldots + c_{k-1}x + c_k \in \mathbb{Z}[x].$$

Для определенности будем считать, что $k \leq m$. Имеем

$$a_n = b_mc_k, \tag{γ_0}$$

$$a_{n-1} = b_{m-1}c_k + b_mc_{k-1}, \tag{γ_1}$$

$$a_{n-2} = b_{m-2}c_k + b_{m-1}c_{k-1} + b_mc_{k-2}, \tag{γ_2}$$

$$a_k = b_0c_k + b_1c_{k-1} + \ldots + b_kc_0, \qquad (\gamma_{n-k})$$

$$a_0 = b_0c_0. \qquad (\gamma_n)$$

Так как $a_n \vdots p$ и $a_n \not\vdots p^2$, то из (γ_0) следует, что один из коэффициентов b_m или c_k делится на p, а другой нет. Рассмотрим случай $b_m \vdots p$, $c_k \not\vdots p$. Противоположный случай рассматривается аналогично.

Так как $a_{n-1} \vdots p$, $b_m \vdots p$ и $c_k \not\vdots p$, то из (γ_1) следует, что $b_{m-1} \vdots p$. Так как $a_{n-2} \vdots p$, $b_{m-1} \vdots p$, $b_m \vdots p$ и $c_k \not\vdots p$, то из (γ_2) следует, что $b_{m-2} \vdots p$. Продолжая аналогичные рассуждения далее, в частности, из (γ_{n-k}) получим, что $b_0 \vdots p$. Теперь из (γ_n) вытекает $a_0 \vdots p$. Противоречие. □

Следствие 4.73. *Над $\mathbb{Q}$ существует неприводимый многочлен любой, сколь угодно высокой, степени.*

Доказательство. Таковым, например, является многочлен $x^n - 2$, неприводимость которого легко проверяется с помощью теоремы 4.72, если положить $p = 2$. □

Пример 4.74. Докажем, что многочлен

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \ldots + x + 1 = \frac{x^p - 1}{x - 1},$$

где p — простое, неприводим. Легко видеть, что $\Phi_p(x)$ неприводим тогда и только тогда, когда неприводим $\Phi_p(x+1) = (x+1)^{p-1} + (x+1)^{p-2} + \ldots + (x+1) + 1$. Воспользовавшись формулой бинома Ньютона, после очевидных преобразований получаем многочлен

$$\Phi_p(x+1) = \frac{(x+1)^p - 1}{x} = x^{p-1} + \sum_{k=1}^{n-1} \frac{p(p-1)\ldots(p-k+1)}{k!} x^{p-k-1},$$

к которому применим признак Эйзенштейна. Действительно, при $1 \le k < p$, так как p — простое и $p(p-1)\ldots(p-k+1)/k! \in \mathbb{Z}$, то $(p-1)\ldots(p-k+1)/k! \in \mathbb{Z}$, поэтому

$$\frac{p(p-1)\ldots(p-k+1)}{k!} \vdots p.$$

Старший коэффициент многочлена $\Phi_p(x+1)$ равен 1 и, конечно, $1 \not\vdots p$. Свободный член равен p и, разумеется, $p \not\vdots p^2$. Следовательно, по критерию Эйзенштейна многочлен $\Phi_p(x+1)$ неприводим, поэтому неприводим и $\Phi_p(x)$.

4.12. Метод Шуберта–Кронекера

Метод Шуберта–Кронекера — это классический метод нахождения разложения многочлена с целыми коэффициентами на неприводимые множители. Он основан на двух простых наблюдениях. Пусть $f(x) \in \mathbb{Z}[x]$, $\deg f(x) = n$. Если $f(x) = g(x)h(x)$, где $g(x) \in \mathbb{Z}[x]$ и $h(x) \in \mathbb{Z}[x]$, то $\deg g(x) \le s$ или $\deg h(x) \le s$, где $s = \lfloor n/2 \rfloor$. Далее, если $x_0 \in \mathbb{Z}$, то $f(x_0) \vdots g(x_0)$ и $f(x_0) \vdots h(x_0)$.

Метод заключается в следующем. Выбираем $x_0, x_1, \ldots, x_s$ — некоторые попарно различные целые числа. Составляем всевозможные наборы $c_0, c_1, \ldots, c_s$, такие, что c_j — произвольный делитель числа $f(x_j)$ ($j = 0, 1, \ldots, s$). Для каждого такого набора восстанавливаем многочлен $g(x)$, такой, что $g(x_j) = c_j$ ($j = 0, 1, \ldots, s$) (например, по формуле интерполяционного многочлена Лагранжа). Делением проверяем $f(x) \vdots g(x)$. Если $f(x) = g(x)h(x)$, то применяем к многочленам $g(x)$ и $h(x)$ ту же процедуру. Если $f(x)$ не делится ни на один из построенных многочленов $g(x)$, то $f(x)$ неприводим над $\mathbb{Q}$.

Пример 4.75. С помощью метода Шуберта–Кронекера разложим многочлен $f(x) = x^5 - 5x^4 + 6x^3 + 2x^2 - 4x + 1$ на неприводимые над $\mathbb{Q}$. Многочлен имеет степень $n = 5$, поэтому, если он приводим, то степень одного из его сомножителей $g(x)$ не превосходит $s = \lfloor n/2 \rfloor = 2$. Пусть $x_0 = -1$, $x_1 = 0$, $x_2 = 1$. Имеем $f(-1) = -5$, $f(0) = 1$, $f(1) = 1$. Делителями $f(-1)$ являются числа ± 1, ± 5. Делителями $f(0)$ и $f(1)$ являются числа ± 1. Таким образом, $g(x)$ определяется одной из следующих интерполяционных таблиц:

x	$g(x)$							
-1	1	1	1	1	5	5	5	5
0	1	1	-1	-1	1	1	-1	-1
1	1	-1	1	-1	1	-1	1	-1

Остальные 8 наборов не рассматриваются, так как определяют многочлены, отличающиеся лишь знаком от многочленов, соответствующих приведенным наборам. Первый набор задает константу 1. Для оставшихся 7 наборов последовательно строим соответствующие интерполяционные многочлены.

По значениям $g(-1) = 1$, $g(0) = 1$, $g(1) = -1$ восстанавливаем интерполяционный многочлен $g(x) = -x^2 - x + 1$. Непосредственным делением проверяем, что $f(x) \not\vdots\, g(x)$.

По значениям $g(-1) = 1$, $g(0) = -1$, $g(1) = 1$ восстанавливаем интерполяционный многочлен $g(x) = 2x^2 - 1$. Получаем, что $f(x) \not\vdots\, g(x)$.

По значениям $g(-1) = 1$, $g(0) = -1$, $g(1) = -1$ восстанавливаем интерполяционный многочлен $g(x) = x^2 - x - 1$. Имеем $f(x) \not\vdots\, g(x)$.

По значениям $g(-1) = 5$, $g(0) = 1$, $g(1) = 1$ восстанавливаем интерполяционный многочлен $g(x) = 2x^2 - 2x + 1$. Имеем $f(x) \not\vdots\, g(x)$.

По значениям $g(-1) = 5$, $g(0) = 1$, $g(1) = -1$ восстанавливаем интерполяционный многочлен $g(x) = x^2 - 3x + 1$. Получаем $f(x) \vdots g(x)$, а именно,

$$f(x) = (x^2 - 3x + 1)(x^3 - 2x^2 - x + 1). \tag{4.35}$$

Легко проверить, что полученные множители рациональных корней не имеют и поэтому, так как их степени не превосходят 3, неприводимы. Таким образом, (4.35) является разложением $f(x)$ на неприводимые над $\mathbb{Q}$ множители и рассматривать дальнейшие наборы не нужно.

4.13. Алгебраические расширения полей

Пусть $f(x) \in F[x]$ — неприводимый над полем F многочлен и α — некоторый его корень, не принадлежащий F. Минимальное по включению поле, включающее F и содержащее α назовем *алгебраическим расширением* поля F и обозначим $F(\alpha)$. Будем также говорить, что $F(\alpha)$ получено в результате *присоединения к F элемента α*.

Пусть в $F[x]$ нашелся многочлен $f(x)$, для которого $f(\alpha) = 0$. Мы можем считать, что $f(x)$ неприводим над F, так как в противном случае $f(x)$ можно заменить на тот неприводимый множитель в его разложении, для которого α является корнем.

Теорема 4.76. *Все элементы поля $F(\alpha)$, где α — иррациональный корень неприводимого над F многочлена $f(x)$ степени n, имеют вид*

$$\beta = \gamma_0 + \gamma_1\alpha + \gamma_2\alpha^2 + \ldots + \gamma_{n-1}\alpha^{n-1} \tag{4.36}$$

для произвольных $\gamma_0, \gamma_1, \ldots, \gamma_{n-1} \in F$. *По любому* $\beta \in F(\alpha)$ *величины* $\gamma_0, \gamma_1, \ldots, \gamma_{n-1}$ *определяются единственным образом.*

Иными словами, для любого $\beta \in F(\alpha)$ *найдется единственный* $g(x) \in F[x]$, *либо равный* 0, *либо степени меньшей* n, *такой, что* $\beta = g(\alpha)$. *Для любого* $g(x) \in F[x]$ *величина* $f(\alpha)$ *принадлежит* $F(\alpha)$.

Доказательство. Так как $F \subseteq F(\alpha)$ и $\alpha \in F(\alpha)$, то в силу замкнутости поля $F(\alpha)$ относительно операций сложения, вычитания и умножения каждое число вида (4.36) принадлежит $F(\alpha)$.

Прежде чем показывать, что других чисел в $F(\alpha)$ нет, докажем что многочлен $g(x)$, такой, что $\beta = g(\alpha)$ и $\deg g(x) < n$ или $g(x) = 0$, по величине $\beta \in F(\alpha)$ определяется единственным образом. Пусть $g(\alpha) = \widetilde{g}(\alpha)$, где $\widetilde{g}(x) \in F[x]$ и $\deg \widetilde{g}(x) < n$ или $\widetilde{g}(x) = 0$. Тогда α является корнем многочлена $h(x) = g(x) - \widetilde{g}(x)$. Так как α является также корнем многочлена $f(x)$, то $f(x)$ и $h(x)$ имеют нетривиальный делитель в $\mathbb{C}[x]$ и, по следствию 4.17, в $F[x]$. Но $\deg h(x) < \deg f(x)$ или $h(x) = 0$. Первое по утверждению 4.54 невозможно, так как $f(x)$ неприводим над F, следовательно, $h(x) = 0$, т. е. $g(x) = \widetilde{g}(x)$.

Покажем, что других элементов, кроме чисел вида (4.36), в $F(\alpha)$ нет. Для этого проверим замкнутость множества всех таких чисел относительно операций сложения, вычитания и умножения и замкнутость множества ненулевых чисел такого вида относительно деления. Замкнутость относительно сложения и вычитания очевидна. Для доказательства замкнутости относительно умножения рассмотрим произведение

$$g(\alpha) \cdot \widetilde{g}(\alpha) \tag{4.37}$$

Разделив $h(x) = g(x)\widetilde{g}(x)$ на $f(x)$ получим в остатке многочлен $r(x)$, такой, что $\deg r(x) < n$ или $r(x) = 0$. Легко видеть, что $h(\alpha) = r(\alpha)$. Замкнутость относительно умножения доказана.

Для доказательства замкнутости множества ненулевых элементов вида (4.36) относительно деления достаточно показать, что для любого ненулевого $g(x) \in F[x]$, степени меньше n, величина $1/g(\alpha)$ может быть представлена в виде (4.36) (такое представление называется *освобождением от иррациональности в знаменателе*). Так как $f(x)$ неприводим над $\mathbb{Q}$, то $f(x)$ и $g(x)$ взаимно просты и, следовательно, существуют коэффициенты Безу $u(x) \in F[x]$ и $v(x) \in F[x]$, такие, что $u(x)f(x) + v(x)g(x) = 1$ и $\deg v(x) < \deg f(x) = n$. Подставляя в последнее равенство вместо x иррациональность α, получаем $v(\alpha)g(\alpha) = 1$, откуда $1/g(\alpha) = v(\alpha)$. Число $v(\alpha)$ имеет вид (4.36). □

Пример 4.77. Поле $\mathbb{C}$ является алгебраическим расширением поля $\mathbb{R}$ и получено в результате присоединения к $\mathbb{R}$ корня i (или $-i$) неприводимого над $\mathbb{R}$ многочлена $f(x) = x^2 + 1$. Многочлен $f(x)$ имеет вторую степень и поэтому каждый элемент поля $\mathbb{C} = \mathbb{R}(i)$ согласно (4.36) можно записать в виде $a + bi$, где $a, b \in \mathbb{R}$.

Пример 4.78. Рассмотрим поле $\mathbb{Q}(\sqrt{2})$. Иррациональность $\sqrt{2}$ является корнем неприводимого над $\mathbb{Q}$ многочлена $x^2 - 2$. По теореме 4.76 поле $\mathbb{Q}(\sqrt{2})$ состоит из элементов вида $a + b\sqrt{2}$, где $a, b \in \mathbb{Q}$, в чем легко убедиться и непосредственно.

Пример 4.79. Рассмотрим поле $\mathbb{Q}(\sqrt[4]{3})$. Иррациональность $\alpha = \sqrt[4]{3}$ является корнем неприводимого над $\mathbb{Q}$ многочлена $f(x) = x^4 - 3$. Многочлен $f(x)$ имеет степень 4, поэтому по теореме 4.76 поле $\mathbb{Q}(\sqrt[4]{3})$ состоит из элементов вида $\gamma_0 + \gamma_1\sqrt[4]{3} + \gamma_2\sqrt[4]{9} + \gamma_3\sqrt[4]{27}$, где $\gamma_j \in \mathbb{Q}$ $(j = 0, 1, 2, 3)$.

Для примера представим в таком виде число, обратное $\sqrt[4]{27}+2\sqrt[4]{9}+\sqrt[4]{3}+1$, т. е. освободимся от иррациональности в знаменателе дроби

$$w=\frac{1}{\sqrt[4]{27}+2\sqrt[4]{9}+\sqrt[4]{3}+1}.$$

Пусть $g(x)=x^3+2x^2+x+1$. Имеем $f(\sqrt[4]{3})=0$ и $w=1/g(\sqrt[4]{3})$. Многочлены $f(x)$ и $g(x)$ рассматривались в примере 4.13, в котором были найдены коэффициенты Безу

$$u(x)=\frac{1}{7}x^2-\frac{2}{7},\qquad v(x)=-\frac{1}{7}x^3+\frac{2}{7}x^2-\frac{1}{7}x+\frac{1}{7},$$

такие, что $u(x)f(x)+v(x)g(x)=1$. Подставляя в последнее равенство вместо x иррациональности $\sqrt[4]{3}$, получаем

$$w=\frac{1}{g(\sqrt[4]{3})}=v(\sqrt[4]{3})=-\frac{\sqrt[4]{27}}{7}+\frac{2\sqrt[4]{9}}{7}-\frac{\sqrt[4]{3}}{7}+\frac{1}{7}.$$

Число $\alpha\in\mathbb{C}$ называется *алгебраическим*, если оно является корнем многочлена из $\mathbb{Q}[x]$ или, что эквивалентно, из $\mathbb{Z}[x]$. Множество всех алгебраических чисел обозначим $\mathbf{A}$. Очевидно, что $\mathbb{Q}\subset\mathbf{A}$, так как рациональное число p/q является корнем многочлена первой степени $qx-p$. Можно показать, что число, которое можно представить в виде выражения, содержащего целые числа, знаки арифметических операций и знаки извлечения корня произвольной натуральной степени также является алгебраическим, однако в силу теоремы Абеля не каждое алгебраическое число можно представить таким выражением.

Можно показать, что $\mathbf{A}$ образует алгебраически замкнутое поле.

Комплексные числа, не являющиеся алгебраическими, называются *трансцендентными*. Примерами трансцендентных чисел служат π, e.

Упражнение 4.80. Докажите, что множество алгебраических чисел $\mathbf{A}$ счетно и, следовательно, (так как $\mathbb{C}$ континуально) множество трансцендентных чисел континуально. Указание: Высотой многочлена $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n$ называется величина $\|f(x)\|=\max\{|a_0|,|a_1|,\ldots,|a_n|\}$. Докажите, что множество многочленов $f(x)$ из $\mathbb{Z}[x]$ заданной высоты конечно. Представьте $\mathbb{Z}[x]$ в виде счетного объединения конечных множеств.

4.14. Границы для комплексных и вещественных корней многочленов

Теорема 4.81. *Пусть*

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n\in\mathbb{C}[x],\quad a_0\neq 0,\quad A=\max_{j=1,2,\ldots,n}|a_j|$$

и $f(\alpha)=0$, $\alpha\in\mathbb{C}$, *тогда*

$$|\alpha|<\frac{A}{|a_0|}+1.$$

Доказательство. Покажем, что если

$$|\alpha|\geq\frac{A}{|a_0|}+1, \tag{4.38}$$

то α не может быть корнем многочлена $f(x)$. Действительно,

$$\begin{aligned} f(\alpha) &\geq |a_0|\cdot|\alpha|^n - |a_1\alpha^{n-1}+\ldots+a_n| \geq \\ &\geq |a_0|\cdot|\alpha|^n - A|\alpha|^{n-1}+\ldots+1| = \\ &= |a_0|\cdot|\alpha|^n - A\cdot\frac{|\alpha|^n-1}{|\alpha|-1} > \\ &> |a_0|\cdot|\alpha|^n - A\cdot\frac{|\alpha|^n}{A}\cdot|a_0| = \\ &= 0 \end{aligned}$$

согласно (4.38). Итак, $f(\alpha) > 0$, т. е. α не является корнем. □

Теорема 4.82. *Пусть*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{R}[x], \quad a_0 > 0$$

и k — минимальное, такое, что $a_k < 0$,

$$B = \max_{a_j<0} |a_j|.$$

Тогда для любого положительного вещественного корня α справедливо нервенство

$$\alpha \leq \sqrt[k]{\frac{B}{a_0}} + 1.$$

Доказательство. Пусть $\alpha > \sqrt[k]{\frac{B}{a_0}} + 1$, откуда

$$(x-1)^k > \frac{B}{a_0}. \tag{4.39}$$

Покажем тогда, что α не может быть корнем многочлена $f(x)$. Имеем

$$\begin{aligned} f(\alpha) &\geq a_0\alpha^n - B(\alpha^{n-k} + \alpha^{n-k-1} + \ldots + \alpha + 1) = \\ &= a_0\alpha^n - B\cdot\frac{\alpha^{n-k+1}-1}{\alpha-1} > \\ &> a_0\alpha^n - B\cdot\frac{\alpha^{n-k+1}}{\alpha-1} = \\ &= \frac{\alpha^{n-k+1}}{\alpha-1}\big(a_0\alpha^{k-1}(\alpha-1)-B\big) > \\ &> \frac{\alpha^{n-k+1}}{\alpha-1}\big(a_0(\alpha-1)^k-B\big) > \\ &> 0, \end{aligned}$$

где последнее неравенство следует из (4.39). Итак, $f(\alpha) > 0$, т. е. α не является корнем. □

Замечание 4.83. Если $a_j \geq 0$ ($j = 0, 1, \ldots, n$), то положительных корней многочлен $f(x)$, легко видеть, не имеет.

Замечание 4.84. Для нахождения нижней оценки отрицательных корней многочлена $f(x)$ достаточно применить теорему 4.82 к многочлену

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

Для нахождения нижней оценки положительных корней достаточно применить теорему 4.82 к многочлену

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

Для нахождения верхней оценки отрицательных корней достаточно применить теорему 4.82 к многочлену

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

4.15. Распределение корней многочленов на вещественной оси

Системой Штурма многочлена $f(x)$ называется конечная последовательность многочленов

$$f_0(x) = f(x), f_1(x), f_2(x), \ldots, f_s(x),$$

такая, что

1) многочлены $f_j(x)$, $f_{j+1}(x)$ не имеют общих вещественных корней ($j = 0, 1, \ldots, s-1$);
2) если $f_0(\alpha) = 0$, то функция $f_0(x)f_1(x)$ возрастает в окрестности α, меняя знак с $-$ на $+$;
3) если $f_j(\alpha) = 0$, то $f_{j-1}(\alpha)f_{j+1}(\alpha) < 0$ ($j = 1, 2, \ldots, s-1$);
4) $f_s(x)$ не имеет вещественных корней.

Пусть $f_0(x) = f(x)$, $f_1(x) = f'(x)$ и $f_j(x)$ — взятый с противоположным знаком остаток при делении $f_{j-2}(x)$ на $f_{j-1}(x)$ ($j = 2, 3, \ldots, s$). Вычисления заканчиваются, когда при делении $f_{s-1}(x)$ на $f_s(x)$ в остатке не будет получен нулевой многочлен (так как степени многочленов $f_j(x)$ убывают, то процесс завершится). Итак,

$$f(x) = q_1(x)f'(x) - f_2(x), \tag{δ_1}$$

$$f'(x) = q_2(x)f_2(x) - f_3(x), \tag{δ_2}$$

$$f_{j-1}(x) = q_j(x)f_j(x) - f_{j+1}(x), \tag{δ_j}$$

$$f_{s-2}(x) = q_{s-1}(x)f_{s-1}(x) - f_s(x), \qquad (\delta_{s-1})$$

$$f_{s-1}(x) = q_s(x)f_s(x). \qquad (\delta_s)$$

Таким образом, многочлены $f_j(x)$ лишь множителями отличаются от многочленов, получаемых в алгоритме Евклида, следовательно, $f_s(x)$ есть НОД $f(x)$ и $f'(x)$ и, в частности, по теореме 4.63 $f_s(x)$ — константа тогда и только тогда, когда $f(x)$ не имеет кратных корней.

Теорема 4.85. *Для произвольного многочлена $f(x) \in \mathbb{R}[x]$, не имеющего кратных корней, система Штурма существует. В частности, системой Штурма многочлена $f(x)$ является последовательность $f_0(x) = f(x)$, $f_1(x) = f'(x)$, $f_2(x), \dots, f_s(x)$, построенная согласно (δ_0)–(δ_s).*

Доказательство. Для системы, построенной согласно (δ_1)–(δ_s), докажем выполнение свойств 1)–4) в определении системы Штурма.

1) Если $f_j(\alpha) = f_{j+1}(\alpha) = 0$, то согласно (γ_j) $f_{j-1}(\alpha) = 0$. Из рассмотрения (γ_{j-1}) получаем $f_{j-2}(\alpha) = 0$ и т. д. Из рассмотрения (γ_2) и (γ_1) получаем $f(\alpha) = f'(\alpha) = 0$, что не возможно, так как $f(x)$ не имеет кратных корней.

2) Если $f(\alpha) = 0$ и $f'(\alpha) > 0$, то $f(x)$ возрастает в окрестности α, следовательно, $f(x)f'(x)$ возрастает. Если $f(\alpha) = 0$ и $f'(\alpha) < 0$, то $f(x)$ убывает в окрестности α, и следовательно, $f(x)f'(x)$ возрастает.

3) Пусть $f_j(\alpha) = 0$. Подставляя α в левую и правую части равенства (δ_j), получаем $f_{j-1}(\alpha) = -f_{j+1}(\alpha)$, поэтому $f_{j-1}(\alpha) = -f_{j+1}(\alpha) < 0$.

4) Как уже отмечалось, $f_s(x)$ есть НОД многочленов $f(x)$ и $f_0(x)$. Так как $f(x)$ не имеет кратных корней, то по теореме 4.63 $f_s(x)$ — ненулевая константа, следовательно, $f_s(x)$ корней не имеет.

□

Замечание 4.86. При построении системы Штурма по формулам (δ_0)–(δ_s) многочлены $f_j(x)$ $(j = 0, 1, \dots, s)$ можно умножать на произвольные положительные константы. Легко видеть, что доказательство теоремы 4.85 распространяется и на такую систему.

Если в конечной последовательности вещественных чисел $\alpha_0, \alpha_1, \dots, \alpha_s$ имеется k переходов от одного знака к другому (нулевые числа пропускаем), то говорят, что последовательность содержит *k перемен знака*.

Теорема 4.87 (Штурм)**.** *Пусть многочлен $f(x) \in \mathbb{R}[x]$ не содержит кратных корней и $f(a) \neq 0$, $f(b) \neq 0$, где $a < b$, $a, b \in \mathbb{R}$. Тогда число действительных корней многочлена $f(x)$, принадлежащих отрезку $[a, b]$, равно $W(a) - W(b)$, где через $W(\alpha)$ обозначено число перемен знака в последовательности значений многочленов системы Штурма $f_0(\alpha), f_1(\alpha), \dots, f_s(\alpha)$.*

Доказательство. Пусть x движется от a к b. Проследим, как при этом меняется $W(x)$. Очевидно, $W(x)$ не меняется, когда x проходит через интервал, на котором нет корней ни одного из многочленов системы Штурма. Покажем, во-первых, что $W(x)$ уменьшается на 1, т. е. теряется одна перемена знака, если x проходит через корень многочлена $f(x)$. Во-вторых, покажем, что $W(x)$ не меняется, когда x проходит через корень одного из многочленов $f_j(x)$ $(j = 1, 2, \dots, s)$. Это докажет, что число корней на отрезке $[a, b]$ равно $W(a) - W(b)$.

Пусть x проходит через корень α многочлена $f(x)$. Согласно свойству 2) в определении системы Штурма $f(x)f_1(x)$ в окрестности α возрастает и, следовательно, при прохождении x через α меняет знак с $-$ на $+$. Это означает, что если в окрестности α многочлен $f(x)$ возрастает, то $f_1(x) > 0$:

	$x < \alpha$	$x > \alpha$
$f(x)$	$-$	$+$
$f_1(x)$	$+$	$+$
$f_2(x)$	$*$	$*$
$\vdots$	$\vdots$	$\vdots$
$f_s(x)$	$*$	$*$

Таким образом, в приведенной таблице число перемен знака в столбце, соответствующем $x < \alpha$, на 1 больше, чем в столбце, соответствующем $x > \alpha$, т.е. теряется одна перемена знака. Если же в окрестности α многочлен $f(x)$ убывает, то $f_1(x) < 0$:

	$x < \alpha$	$x > \alpha$
$f(x)$	$+$	$-$
$f_1(x)$	$-$	$-$
$f_2(x)$	$*$	$*$
$\vdots$	$\vdots$	$\vdots$
$f_s(x)$	$*$	$*$

Аналогично приходим к выводу, что теряется одна перемена знака.

Пусть теперь x проходит через корень α одного из многочленов $f_j(x)$ $(j = 1, 2, \dots, s)$. Согласно свойству 3) в определении системы Штурма $f_{j-1}(x)f_{j+1}(x) < 0$. Это возможно в следующих 4 случаях (два из них соответствуют возрастающей в окрестности α функции $f_j(x)$, а два — убывающей):

	I.		II.		III.		IV.	
	$x < \alpha$	$x > \alpha$	$x < \alpha$	$x > \alpha$	$x < \alpha$	$x > \alpha$	$x < \alpha$	$x > \alpha$
$f_{j-1}(x)$	$-$	$-$	$+$	$+$	$-$	$-$	$+$	$+$
$f_j(x)$	$-$	$+$	$-$	$+$	$+$	$-$	$+$	$-$
$f_{j+1}(x)$	$+$	$+$	$-$	$-$	$+$	$+$	$-$	$-$

В каждом случае число перемен знака в столбце, соответствующем $x < \alpha$, совпадает с числом перемен знака в столбце, соответствующем $x > \alpha$. Итак, если x проходит через корень многочлена $f_j(x)$ $(j = 1, 2, \ldots, s)$, то $W(x)$ не изменяется. $\square$

Локализовать вещественные корни многочлена $f(x)$ — значит найти интервалы на вещественной оси, на каждом из которых содержится ровно один корень и других вещественных корней нет. После того, как корни локализованы, для их уточнения используют различные численные методы (деления пополам, секущих, касательных Ньютона и др.).

Пример 4.88. С помощью теоремы Штурма локализуем вещественные корни многочлена $f(x) = x^5 - 4x - 2$. Теорема 4.82 дает следующую верхнюю оценку на положительные корни: $\alpha \le \sqrt[4]{4/1} = \sqrt{2} < 2$. Применяя ту же теорему к многочлену $-f(-x) = x^5 - 4x + 2$, получаем нижнюю оценку на величину отрицательных корней: $\alpha \ge -\sqrt[4]{4/1} = -\sqrt{2} > -2$. Итак, все вещественные корни многочлена $f(x)$ лежат на интервале $(-2, 2)$.

Построим систему Штурма.

$$f_0(x) = f(x), \qquad f_1(x) = f'(x) = 5x^4 - 4.$$

При делении $f_0(x)$ на $f_1(x)$ получаем в остатке $r_2(x) = -(16/5)x - 2$. Меняем знак у $r_2(x)$ и домножаем на $5/2$, получаем

$$f_2(x) = 8x + 5.$$

При делении $f_1(x)$ на $f_2(x)$ получаем в остатке $r_3(x) = -13259/4096$, поэтому

$$f_3(x) = 1.$$

Полученный многочлен $f_3(x)$ является наибольшим общим делителем $f(x)$ и $f'(x)$. Таким образом, $f(x)$ и $f'(x)$ взаимно просты, поэтому $f(x)$ кратных корней не имеет.

Вычислим значения многочленов системы Штурма на концах интервала $(-2, 2)$. Число перемен знаков составит $W(-2) = 3$, $W(2) = 0$ (см. таблицу ниже). Таким образом, многочлен $f(x)$ имеет $W(-2) - W(2) = 3$ вещественных корня. Вычислим значения многочленов системы Штурма в середине интервала $(-2, 2)$. Число перемен знака в точке 0 составит $W(0) = 1$. Следовательно, имеется $W(0) - W(2) = 1$ положительный корень и $W(-2) - W(0) = 2$ отрицательных корня. Для локализации отрицательных корней вычислим значения многочленов системы Штурма в точке -1. Приходим к выводу, что корни локализованы на интервалах $(-2, -1)$, $(-1, 0)$ и $(0, 2)$.

Знаки значений многочленов системы Штурма в рассматриваемых точках сведены в таблицу.

	-2	-1	0	2
$f_0(x) = f(x) = x^5 - 4x - 2$	$-$	$+$	$-$	$+$
$f_1(x) = f'(x) = 5x^4 - 4$	$+$	$+$	$-$	$+$
$f_2(x) = 8x + 5$	$-$	$-$	$+$	$+$
$f_3(x) = 1$	$+$	$+$	$+$	$+$
$W(x)$	3	2	1	0

Теорема 4.87 была сформулирована и доказана для случая многочлена $f(x)$ без кратных корней. Если $f(x)$ имеет кратные корни, то НОД многочленов $f(x)$ и $f'(x)$ имеет положительную степень, поэтому $f_s(x)$ в системе, построенной согласно (γ_1)–(γ_s), может не удовлетворять свойству 4) в определении системы Штурма. Тем не менее, справедлив результат, обобщающий теорему 4.87. на случай многочлена с корнями произвольной кратности.

Теорема 4.89. *Пусть $f(x)$ — многочлен из $\mathbb{R}[x]$, такой, что $f(a) \ne 0$, $f(b) \ne 0$, где $a < b$, $a, b \in \mathbb{R}$. Пусть также $f_0(x), f_1(x), \ldots, f_s(x)$ — система, построенная согласно (γ_1)–(γ_s). Тогда число действительных корней (без учета их кратности) многочлена $f(x)$, принадлежащих отрезку $[a, b]$, равно $W(a) - W(b)$, где через $W(\alpha)$ обозначено число перемен знака в последовательности $f_0(\alpha), f_1(\alpha), \ldots, f_s(\alpha)$.*

Доказательство. Многочлен $f_s(x)$ является наибольшим общим делителем многочленов $f(x)$ и $f_s(x)$. Из (γ_1)–(γ_s) получаем, что каждый из многочленов $f_0(x), f_1(x), \ldots, f_s(x)$ делится на $f_s(x)$. Пусть $g_j(x) = f_j(x)/f_s(x)$ $(j = 0, 1, \ldots, s)$. Легко видеть, что система $g_0(x), g_1(x), \ldots, g_s(x)$ является системой Штурма для многочлена $g(x) = f(x)/f_s(x)$, все корни которого по следствию 4.63 (см. также замечание 4.64) совпадают с корнями многочлена $f(x)$, но имеют кратность 1. Поэтому число корней (без учета кратности) многочлена $f(x)$ совпадает с разностью в числе перемен знака в последовательностях значений $g_0(a), g_1(a), \ldots, g_s(a)$ и $g_0(b), g_1(b), \ldots, g_s(b)$. Но при заданном α последовательность $f_0(\alpha), f_1(\alpha), \ldots, f_s(\alpha)$ получается из последовательности $g_0(\alpha), g_1(\alpha), \ldots, g_s(\alpha)$ умножением на константу $f_s(\alpha)$. Так как $f_s(a) \neq 0$, $f_s(b) \neq 0$ (иначе было бы $f_s(b) = 0$ или $f_s(b) = 0$), то число перемен знака в этих последовательностях одно и то же. □

Некоторую информацию о вещественных корнях (иногда достаточно полную) дает теорема Бюдана–Фурье.

Теорема 4.90 (Бюдан–Фурье). *Число вещественных корней ненулевого многочлена* $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{R}[x]$, *расположенных на интервале* (a, b) $(f(a) \neq 0,\ f(b) \neq 0)$ *с учетом их кратностей равно или на четное число меньше разности между числом перемен знака в последовательности* $f(a), f'(a), f''(a), \ldots, f^{(n)}(a)$ *и числом перемен знака в последовательности* $f(b), f'(b), f''(b), \ldots, f^{(n)}(b)$.

Доказательство. Обозначим $N(f(x))$ — число корней многочлена $f(x)$ на интервале (a, b), а $L(f(x))$ — разность между числом перемен знака в последовательности $f(a), f'(a), f''(a), \ldots, f^{(n)}(a)$ и числом перемен знака в последовательности $f(b), f'(b), f''(b), \ldots, f^{(n)}(b)$.

Заметим, что $f^{(n)}(x)$ — константа, равная $n!a_0$.

Сначала докажем, что $N(f(x))$ и $L(f(x))$ могут отличаться лишь на четное число (т. е. имеют одинаковую четность). Когда x увеличивается и проходит простой корень, $f(x)$ меняет знак. Когда x увеличивается и проходит k-кратный корень, $f(x)$ меняет знак, если k нечетно, и не меняет знака, если k четно. Поэтому $N(f(x))$ четно, если $f(a)$ и $f(b)$ имеют одинаковые знаки, и $N(f(x))$ нечетно, если $f(a)$ и $f(b)$ имеют разные знаки. Легко проверить, что аналогичное утверждение справедливо и для $L(f(x))$, поэтому $N(f(x))$ и $L(f(x))$ могут отличаться лишь на четное число.

Теперь докажем, что

$$N(f(x)) \le N(f'(x)) + 1. \tag{4.40}$$

Действительно, по теореме Ролля между любыми двумя корнями многочлена $f(x)$ лежит корень его производной $f'(x)$. Кроме того, по теореме 4.28 каждый k-кратный корень многочлена $f(x)$ является $(k-1)$-кратным корнем многочлена $f'(x)$. Отсюда $N(F'(x)) \ge N(f(x)) - 1$.

Очевидно,

$$L(f'(x)) \le L(f(x)). \tag{4.41}$$

Теперь индукцией по $\deg f(x)$ докажем, что $N(f(x)) \le L(f(x))$. Это завершит доказательство теоремы. Если $\deg f(x) = 0$, то $L(f(x)) = N(f(x)) = 0$. Если $\deg f(x) > 0$, то, используя (4.40), (4.41) и применяя предположение индукции к $f'(x)$, получаем

$$N(f(x)) \le N(f'(x)) + 1 \le L(f'(x)) + 1 \le L(f(x)) + 1.$$

Но так как $L(f(x))$ и $N(f(x))$ могут отличаться лишь на четное число, то $N(f(x)) \leq L(f(x))$. $\square$

Упражнение 4.91. Докажите следующее усиление теоремы 4.90.

Число вещественных корней ненулевого многочлена $f(x) \in \mathbb{R}[x]$, расположенных на интервале (a,b) ($f(a) \neq 0$, $f(b) \neq 0$) с учетом их кратностей равно или на четное число меньше разности между числом перемен знака в последовательности $f(a), f'(a), f''(a), \ldots, f^{(n)}(a)$ и числом перемен знака в последовательности, полученной из $f(b), f'(b), f''(b), \ldots, f^{(n)}(b)$ заменой встречающихся нулей такими ненулевыми числами, что если $f^{(k-1)}(b) \neq 0$, $f^{(k)}(b) = \ldots = f^{(k+l-1)}(b) = 0$, $f^{(k+l)}(b) \neq 0$, то число c_{k+i}, заменяющее $f^{(k+i)}(b)$ ($i = 0, 1, \ldots, l-1$), имеет тот же знак, что и $f^{(k+l)}(b)$, если $l-i$ четно, и противоположный знак, если $l-i$ нечетно.

Указание: вместо интервала (a,b) рассмотрите интервал $(a, b-\varepsilon)$, где ε — достаточно маленькое число.

Следствие 4.92 (Правило знаков Декарта)**.** *Число положительных корней ненулевого многочлена* $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in \mathbb{R}[x]$ *с учетом их кратностей равно или на четное число меньше числа перемен знака в последовательности коэффициентов* $a_0, a_1, \ldots, a_n$*. В частности, если число перемен знака равно* 0 *или* 1*, то* $f(x)$ *соответственно либо не имеет положительных корней, либо имеет ровно один положительный корень.*

Доказательство. Не нарушая общности, можно считать, что $f(0) \neq 0$, так как в противном случае $f(x) = a_0x^n + \ldots + a_{n-r}x^r$, $a_{n-r} \neq 0$, и мы можем разделить $f(x)$ на x^r. При этом ни число положительных корней, ни число перемен знака в последовательности коэффициентов не изменятся.

Так как $f^{(k)}(0) = k!a_{n-k}$, то число перемен знака в последовательности $f(0), f'(0), \ldots, f^{(n)}(0)$ совпадает с числом перемен знака в последовательности коэффициентов многочлена $f(x)$. Пусть M — достаточно большое число, такое, что знак в последовательности $f(M), f'(M), \ldots, f^{(n)}(M)$ не меняется и $f(\alpha) \neq 0$ при $\alpha \geq M$. Теперь достаточно применить теорему 4.90 к интервалу $[0, M]$. $\square$

Пример 4.93. Правило знаков Декарта, примененное к многочлену $f(x) = x^5 - 4x - 2$ (1 перемена знака) из примера 4.88, сразу приводит к выводу, что $f(x)$ имеем 1 положительный корень. Это же правило, примененное к многочлену $f(-x) = -x^5 + 4x - 2$ (2 перемены знака), позволяет утверждать лишь то, что $f(x)$ имеет либо 2, либо 0 отрицательных корней.

Теперь применим теорему Бюдана–Фурье.

	-2	-1	$-\varepsilon$	0	2
$f(x) = x^5 - 4x - 2$	$-$	$+$	$-$	$-$	$+$
$f'(x) = 5x^4 - 4$	$+$	$+$	$-$	$-$	$+$
$f''(x) = 20x^3$	$-$	$-$	$-$	0	$+$
$f'''(x) = 60x^2$	$+$	$+$	$+$	0	$+$
$f^{IV}(x) = 120x$	$-$	$-$	$-$	0	$+$
$f^{V}(x) = 120$	$+$	$+$	$+$	$+$	$+$
$W(x)$	5	4	3	1	0

Число перемен знака в последовательности значений производной в точках -2, -1, 0 и 2 составит соответственно $W(-2) = 5$, $W(-1) = 4$, $W(0) = 1$, $W(2) = 0$, т. е. имеем один корень на интервале $(-2, -1)$ и 1 корень на интервале $(0, 2)$. Получить информацию о возможных корнях на интервале $(-1, 0)$ в данном случае удается, если вместо 0 рассмотреть точку $-\varepsilon$ для достаточно малого положительного ε, такого, что на интервале $(-\varepsilon, 0)$ многочлен $f(x)$ корней не имеет. Получаем $W(\varepsilon) = 3$. Следовательно, на интервале $(-1, 0)$ находится $W(-1) - W(\varepsilon) = 1$ вещественный корень.

Глава 5

Линейные пространства

5.1. Аксиоматическое определение линейного пространства

Пусть V — непустое множество, F — поле. V называется *линейным* (или *векторным*) *пространством* над полем F, если

I. задано *правило сложения*, ставящее в соответствие любым двум элементам a, b из V единственный элемент c из V, называемый *суммой* и обозначаемый $c = a + b$;

II. задано *правило умножения на число*, ставящее в соответствие каждому a из V и каждому α из F единственный элемент d из V, обозначаемый $d = \alpha a$ (или $d = \alpha \cdot a$);

III. выполняются следующие свойства (*аксиомы линейного пространства*):

(1) $\forall a, b, c \in V \ \ a + (b + c) = (a + b) + c$ (*ассоциативность*[1]),

(2) $\forall a, b \in V \ \ a + b = b + a$ (*коммутативность*),

(3) $\exists o \ \forall a \in V \ \ a + o = a$ (o называется *нулевым* элементом[2]),

(4) $\forall a \in V \ \exists b \in V \ \ a + b = o$ (b называется элементом, *противоположным* к a и обозначается $-a$),

(5) $\forall a \in V \ \ 1 \cdot a = a$,

(6) $\forall a \in V \ \forall \alpha, \beta \in F \ \ \alpha(\beta a) = (\alpha\beta a)$,

(7) $\forall a, b \in V \ \forall \alpha \in F \ \ \alpha(a + b) = \alpha a + \alpha b$ (дистрибутивность I),

(8) $\forall a \in V \ \forall \alpha, \beta \in F \ \ (\alpha + \beta)a = \alpha a + \beta a$ (дистрибутивность II).

Элементы множества V называются *векторами*, элементы множества F называются *скалярами* (или, просто, числами). Далее, как правило, векторы мы будем обозначать латинскими буквами, скаляры — греческими.

Линейное пространство над полем $\mathbb{R}$ называется *вещественным*. Линейное пространство над полем $\mathbb{C}$ называется *комплексным*.

[1]Скобки в выражениях $a + (b + c)$, $(a + b) + c$ указывают на порядок выполнения операции сложения. Как следует из аксиомы (3) в выражении вида $a + b + c$ их можно опустить.

[2]Нулевой вектор $o \in V$ следует отличать от числа $0 \in F$.

Если указана природа элементов множества V и определены операции над этими элементами, то пространство будем называть *конкретным*. Рассмотрим важные типы конкретных линейных пространств.

5.2. Примеры линейных пространств

Нулевое пространство $\{o\}$. В этом случае $V = \{o\}$. В качестве F можно рассматривать любое поле. Операции определяются тривиальным образом: $o + o = o$, $\alpha o = o$. Легко проверить, что указанное множество с такими операциями является линейным пространством над полем F.

Пространство геометрических векторов $\mathbf{V}_3$. Элементами этого пространства являются геометрические векторы, т. е. направленные отрезки, в пространстве. Геометрический вектор, начало которого находится в точке A, а конец — в точке B обозначается $\overrightarrow{AB}$. Два вектора считаются равными, если они сонаправлены и имеют одинаковую длину. Ввиду этого удобно считать, что все векторы закреплены в одной точке $\mathcal{O}$, называемой *полюсом* или *началом координат*. Такое рассмотрение также удобно тем, что с каждым вектором ассоциируется некоторая точка пространства — его конец, и, наоборот, с каждой точкой пространства связан единственный вектор, называемый *радиус-вектором точки*, начало которого закреплено в полюсе, а конец указывает на эту точку. Векторы складываются по правилу параллелограмма: суммой двух радиус-векторов называется диагональ параллелограмма, построенного на этих векторах. Векторы можно умножать на вещественные числа. Под произведением радиус-вектора на число α понимается вектор, длина которого равна длине исходного вектора, умноженной на $|\alpha|$, а направление совпадает с направлением исходного вектора, если $\alpha > 0$, и заменяется на противоположное, если $\alpha < 0$. Легко проверяются аксиомы (1)–(8). Таким образом, $\mathbf{V}_3$ является линейным пространством над полем $\mathbb{R}$. Аналогичные совокупности векторов на плоскости и на прямой обозначим $\mathbf{V}_2$, $\mathbf{V}_1$ соответственно. Эти совокупности также являются линейными пространствами над полем $\mathbb{R}$.

Арифметическое пространство F^n. Элементами этого пространства являются столбцы высоты n, составленные из чисел. Два столбца

$$a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, \qquad b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}.$$

назовем равными, если $\alpha_j = \beta_j$ $(j = 1, 2, \ldots, n)$.

Операции сложения столбцов и умножения их на числа из F определены по следующим правилам:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1+\beta_1 \\ \alpha_2+\beta_2 \\ \vdots \\ \alpha_n+\beta_n \end{pmatrix}, \quad \alpha\cdot\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \\ \vdots \\ \alpha\alpha_n \end{pmatrix}.$$

Легко проверить, что аксиомы (1)—(8) выполнены и поэтому F^n образует линейное пространство над полем F, называемое *арифметическим n-мерным пространством*. В частности, нулевым вектором в пространстве F^n является

$$o = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Пространство многочленов $F[x]$. Элементами этого пространства являются многочлены с коэффициентами из поля F. Легко видеть, что относительно обычных операций сложения и умножения многочленов на числа из F множество $F[x]$ образует линейное пространство над полем F. В частности, нулевым вектором в этом пространстве является нулевой многочлен.

5.3. Простейшие следствия из аксиом

Утверждение 5.1. *В любом линейном пространстве нулевой элемент единственен.*

Доказательство. Предположим, что существует два нулевых элемента $o_1, o_2 \in V$, тогда положив в аксиоме (3) $a = o_1$, $o = o_2$, получаем $o_1 + o_2 = o_1$. С другой стороны, положив $a = o_2$, $o = o_1$ и воспользовавшись аксиомой (1), получаем $o_1+o_2 = o$. Приравнивая правые части полученных равенств, получаем, $o_1 = o_2$. □

Нулевой вектор o часто называется нулем, однако его следует отличать от числа 0 из поля F.

Утверждение 5.2. *Для любого вектора $a \in V$ существует единственный противоположный элемент.*

Доказательство. Пусть b_1, b_2 — элементы пространства V, противоположные вектору a. По аксиоме (4)

$$b_1 + (a + b_2) = b_1 + o = b_1.$$

Аналогично,

$$(b_1 + a) + b_2 = o + b_2 = b_2.$$

По аксиоме (4) правые части приведенных равенств совпадают, поэтому $b_1 = b_2$. □

Утверждение 5.3. *Для любых a, b из V уравнение $a + x = b$ имеет и единственное решение.*

Доказательство. Подстановкой легко убеждаемся, что $x = (-a) + b$ является решением уравнения. Для доказательства единственности предположим, что имеется два решения x_1, x_2. Тогда $a + x_1 = a + x_2 = b$. Прибавляя ко всем частям равенства $(-a)$, получаем $(-a) + a + x_1 = (-a) + a + x_2$, откуда $x_1 = x_2$. □

Назовем *разностью* векторов b и a решение уравнения $a + x = b$. Разность обозначим $b - a$. Мы установили, что $b - a = (-a) + b = b + (-a)$.

Утверждение 5.4. *$0a = o$ для любого $a \in V$.*

Доказательство. Имеем $a + 0a = 1a + 0a = (1 + 0)a = 1a = a$. Таким образом, $a + 0a = a$, откуда $0a = a - a = o$. □

Утверждение 5.5. *$\alpha o = o$ для любого $\alpha \in F$.*

Доказательство. Пусть $a \in V$. Имеем $\alpha a + \alpha o = \alpha(a + o) = \alpha a$. Таким образом, $\alpha a + \alpha o = \alpha a$, откуда $\alpha o = \alpha a - \alpha a = o$. □

Утверждение 5.6. *Для любых векторов a, b из V и любых чисел α, β из F справедливы равенства*:

1) $(-\alpha)a = -(\alpha a)$;
2) $(-1)a = -a$;
3) $(\alpha - \beta)a = \alpha a - \beta a$;
4) $\alpha(a - b) = \alpha a - \alpha b$.

Доказательство. 1) Покажем, что вектор $(-\alpha)a$ является противоположным к αa. Действительно, $\alpha a + (-\alpha)a = (\alpha - \alpha)a = 0a = o$. Пункты 2)–4) вытекают из 1). □

Утверждение 5.7. *Пусть $a \in V$ и $\alpha \in F$. Если $\alpha a = 0$, то $\alpha = 0$ или $a = 0$.*

Доказательство. Если $\alpha = 0$, то утверждение справедливо. Пусть $\alpha \neq 0$, тогда $a = \dfrac{1}{\alpha}(\alpha a) = \dfrac{1}{\alpha} o = o$. □

Сумма $a_1 + a_2 + \ldots + a_n$ определяется по индукции следующим образом:

$$a_1 + a_2 + \ldots + a_n = \begin{cases} a_1, & \text{при } n = 1, \\ (a_1 + a_2 + \ldots + a_{n-1}) + a_n, & \text{при } n > 1. \end{cases}$$

Т. е. в выражении $a_1 + a_2 + \ldots + a_n$ операция сложения выполняется слева направо.

Рассмотрим обобщения свойств ассоциативности, коммутативности и дистрибутивности.

Утверждение 5.8 (Обобщенная ассоциативность). *Пусть $a_1, a_2, \ldots, a_n$ — векторы из V, тогда для любого $k \in \{1, 2, \ldots, n\}$*

$$(a_1 + \ldots + a_k) + (a_{k+1} + \ldots + a_n) = a_1 + a_2 + \ldots + a_n.$$

Доказательство. Докажем свойство индукцией по n. При $n \leq 3$ доказываемое равенство либо тривиально, либо представляет собой запись свойства ассоциативности. При $n > 3$ имеем

$$(a_1 + \ldots + a_k) + (a_{k+1} + \ldots + a_n) = (a_1 + \ldots + a_k) + (a_{k+1} + \ldots + a_{n-1}) + a_n.$$

По предположению индукции

$$(a_1 + a_2 + \ldots + a_k) + (a_{k+1} + a_2 + \ldots + a_{n-1}) = a_1 + a_2 + \ldots + a_{n-1},$$

откуда получаем требуемое. □

Утверждение 5.9 (Обобщенная коммутативность). *Пусть $j_1, j_2, \ldots, j_n$ — перестановка чисел $1, 2, \ldots, n$, и пусть $a_1, a_2, \ldots, a_n$ — векторы из V, тогда*

$$a_{j_1} + a_{j_2} + \ldots + a_{j_n} = a_1 + a_2 + \ldots + a_n.$$

Доказательство. Свойство докажем индукцией по n. При $n \leq 2$ доказываемое равенство либо тривиально, либо представляет собой запись свойства коммутативности. Пусть k такое, что $j_k = n$, тогда

$$a_{j_1} + a_{j_2} + \ldots + a_{j_n} = (a_{j_1} + \ldots + a_{j_{k-1}}) + a_n + (a_{j_{k+1}} + \ldots + a_{j_n}).$$

Два раза используя предположение индукции и применяя ассоциативность, выводим

$$\begin{aligned}(a_{j_1} + \ldots + a_{j_{k-1}}) + a_n + (a_{j_{k+1}} + \ldots + a_{j_n}) &= \\ = (a_{j_1} + \ldots + a_{j_{k-1}}) + (a_{j_{k+1}} + \ldots + a_{j_n}) + a_n &= \\ = (a_1 + \ldots + a_{n-1}) + a_n&,\end{aligned}$$

откуда получаем требуемое. □

Утверждение 5.10 (Обобщенная дистрибутивность). *Пусть $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_n$ — числа из F и пусть $a, a_1, a_2, \ldots, a_n$ — векторы из V. Тогда*

1) $\alpha(a_1 + a_2 + \ldots + a_n) = \alpha a_1 + \alpha a_2 + \ldots + \alpha a_n$;

2) $(\alpha_1 + \alpha_2 + \ldots + \alpha_n)a = \alpha_1 a + \alpha_2 a + \ldots + \alpha_n a$.

Доказательство. Обе формулы доказываются индукцией по n. □

Для сокращенной записи выражений вида $a_1 + a_2 + \ldots + a_n$ используется знак суммирования $\sum$. Полагают

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n.$$

Формулы из утверждений (5.8)–(5.10) теперь можно записать, используя знак суммирования, следующим образом:

$$\sum_{i=1}^{k} a_i + \sum_{i=k+1}^{n} a_i = \sum_{i=1}^{n} a_i, \qquad \sum_{i=1}^{n} a_{i_j} = \sum_{i=1}^{n} a_i$$

$$\alpha \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \alpha a_i, \qquad \left(\sum_{i=1}^{n} \alpha_i\right) \cdot a = \sum_{i=1}^{n} \alpha_i a.$$

5.4. Линейные подпространства

Пусть V — линейное пространство над полем F. Множество $W \subseteq V$ называется (линейным) *подпространством*, если оно является пространством над F относительно операций сложения векторов и умножения их на числа, определенных в V.

Теорема 5.11 (Критерий линейного подпространства)**.** *Для того, чтобы непустое подмножество W линейного пространства V являлось подпространством необходимо и достаточно, чтобы*

1. *$a+b \in W$ для любых a, b из W (замкнутость относительно сложения векторов),*
2. *$\alpha a \in W$ для любого α из F и любого a из W (замкнутость относительно умножения векторов на числа).*

Доказательство. Необходимость. Условия 1–2 включены в определение линейного пространства.
Достаточность. Проверим, что все 8 аксиом линейного пространства в условиях теоремы выполнены. Проверка аксиом (1)–(2), (5)–(8) тривиальна.

Проверим аксиому (3). Покажем, что $o \in W$. Пусть $\alpha = 0$, $a \in W$, тогда $o = 0a = \alpha a \in W$ по условию 2.

Проверим аксиому (4). Покажем, что $-a \in W$ для любого $a \in W$. Пусть $\alpha = -1$, тогда $-a = (-1)a = \alpha a \in W$ по условию 2. □

Легко проверить, что согласно приведенному критерию подпространствами являются, например, множество радиус-векторов, концы которых лежат на прямой, проходящей через полюс, в пространстве $\mathbf{V}_3$; множество арифметических векторов с нулевой первой (и, вообще, любой фиксированной) компонентой в пространстве F^n; множество четных (нечетных) многочленов в пространстве $F[x]$; множество многочленов степени не выше n в пространстве $F[x]$.

5.5. Линейные комбинации и линейные оболочки

Пусть V — линейное пространство над полем F. *Системой векторов* назовем произвольную конечную последовательность векторов $a_1, \dots, a_n$ из V (возможно с повторениями). Иногда нам будет нужна пустая система, т. е. система, не содержащая ни одного вектора.

Линейной комбинацией векторов $a_1, \dots, a_n$ *с коэффициентами* $\alpha_1, \dots, \alpha_n$ называется выражение $\alpha_1 a_1 + \ldots + \alpha_n a_n$.

Множество всех линейных комбинаций векторов $a_1, \dots, a_n$ называется *линейной оболочкой* этих векторов и обозначается $L(a_1, \dots, a_n)$. Формально:

$$L(a_1, \dots, a_n) = \{\alpha_1 a_1 + \ldots + \alpha_n a_n : \alpha_1, \dots, \alpha_n \in F\}.$$

Теорема 5.12. *Пусть* $a_1, \dots, a_n$ *— произвольные векторы из* V. *Тогда* $L(a_1, \dots, a_n)$ *— подпространство в* V *и для любого другого подпространства* W, *содержащего векторы* $a_1, \dots, a_n$, *справедливо* $L(a_1, \dots, a_n) \subseteq W$.

Доказательство. Докажем сперва, что $L(a_1, \dots, a_n)$ — подпространство в V. Действительно, если $a = \alpha_1 a_1 + \ldots + \alpha_n a_n$ и $b = \beta_1 a_1 + \ldots + \beta_n a_n$, то $a + b = (\alpha_1 + \beta_1) a_1 + \ldots + (\alpha_n + \beta_n) a_n$. Таким образом, $a + b \in L(a_1, \dots, a_n)$. Если $\alpha \in F$, то $\alpha a = (\alpha\alpha_1) a_1 + \ldots + (\alpha\alpha_n) a_n$. Таким образом, $\alpha a \in F$. Достаточные условия в теореме 5.11 выполнены.

Пусть теперь W — произвольное подпространство, содержащее векторы $a_1, \dots, a_n$. Тогда для любых чисел $\alpha_1, \dots, \alpha_n$ из F справедливо $\alpha_j a_j \in W$ $(j = 1, \dots, n)$ (замкнутость относительно операции умножения на число) и $\alpha_1 a_1 + \ldots + \alpha_n a_n \in W$ (замкнутость относительно операции сложения), т. е. $b \in W$ для произвольного $b \in L(a_1, \dots, a_n)$. □

Говорят, что вектор b *линейно выражается* через систему $a_1, \dots, a_n$, если $b = \alpha_1 a_1 + \ldots + \alpha_n a_n$ для некоторых $\alpha_1, \dots, \alpha_n$, иными словами, $b \in L(a_1, \dots, a_n)$. Говорят, что система $b_1, \dots, b_m$ *линейно выражается* через систему $a_1, \dots, a_n$, если для всех $i \in \{1, \dots, m\}$ вектор b_i линейно выражается через систему $a_1, \dots, a_n$.

Утверждение 5.13. *Система* $b_1, \dots, b_m$ *линейно выражается через систему* $a_1, \dots, a_n$ *тогда и только тогда, когда* $L(b_1, \dots, b_m) \subseteq L(a_1, \dots, a_n)$.

Доказательство. Необходимость. Если система $b_1, \dots, b_m$ линейно выражается через систему $a_1, \dots, a_n$, т. е. $b_i \in L(a_1, \dots, a_n)$ для всех $i \in \{1, \dots, m\}$, то по замкнутости $\beta_1 b_1 + \ldots + \beta_m b_m \in L(a_1, \dots, a_n)$ для любых $\beta_1, \dots, \beta_m$ из F.
Достаточность. Пусть $L(b_1, \dots, b_m) \subseteq L(a_1, \dots, a_n)$. Так как $b_i \in L(b_1, \dots, b_m)$, то $b_i \in L(a_1, \dots, a_n)$ $(i = 1, \dots, m)$. Т. е. система $b_1, \dots, b_m$ линейно выражается через систему $a_1, \dots, a_n$. □

Очевидно следующее

Утверждение 5.14. *Отношение линейной выразимости является рефлексивным и транзитивным, но, в общем случае, не симметричным.*

Системы $a_1, \dots, a_n$ и $b_1, \dots, b_m$ называются *эквивалентными*, если $a_1, \dots, a_n$ линейно выражается через $b_1, \dots, b_m$, а $b_1, \dots, b_m$ линейно выражается через $a_1, \dots, a_n$.

Следствие 5.15. *Системы* $a_1, \dots, a_n$ *и* $b_1, \dots, b_m$ *эквивалентны тогда и только тогда, когда* $L(a_1, \dots, a_n) = L(b_1, \dots, b_m)$.

Очевидно следующее

Утверждение 5.16. *Отношение эквивалентности линейных систем является рефлексивным, симметричным и транзитивным.*

5.6. Линейная зависимость векторов

Линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_n a_n$ векторов $a_1, \ldots, a_n \in V$ называется *тривиальной*, если $\alpha_1 = \ldots = \alpha_n = 0$. Очевидно, тривиальная комбинация равна нулевому вектору. Говорят, что (непустая) система векторов $a_1, \ldots, a_n$ *линейно зависима*, если существует нетривиальная комбинация этих векторов, равная нулевому вектору, иными словами, если найдутся числа $\alpha_1, \ldots, \alpha_n \in F$, такие, что $\alpha_j \neq 0$ для некоторого $j \in \{1, \ldots, n\}$ и

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = o. \tag{5.1}$$

В противном случае (непустая) система называется *линейно независимой*. В силу важности вводимых терминов переформулируем определение линейной независимости. Система $a_1, \ldots, a_n$ называется линейно независимой, если равенство (5.1) возможно лишь в случае $\alpha_1 = \ldots = \alpha_n = 0$.

Утверждение 5.17 (Система из одного вектора)**.** *Система, состоящая из одного вектора линейно зависима тогда и только тогда, когда вектор нулевой.*

Доказательство. Для системы, состоящей лишь из одного вектора a, равенство (5.1) примет вид $\alpha a = 0$, откуда $\alpha = 0$, но тогда комбинация тривиальная, или $a = 0$, что и утверждается. ☐

Утверждение 5.18 (Система из двух векторов)**.** *Система, состоящая из двух векторов линейно зависима тогда и только тогда, когда векторы пропорциональны.*

Доказательство. Равенство (5.1) принимает вид $\alpha_1 a_1 + \alpha_2 a_2 = 0$, причем $\alpha_1 \neq 0$ или(и) $\alpha_2 \neq 0$. В первом случае имеем $a_1 = (-\alpha_2/\alpha_1)a_2$, во втором имеем $a_2 = (-\alpha_1/\alpha_2)a_1$. ☐

Следствие 5.19. *Два вектора геометрического пространства* $\mathbf{V}_2$ ($\mathbf{V}_3$) *линейно зависимы тогда и только тогда, когда они коллинеарны.*

Утверждение 5.20. *Если подсистема некоторой системы линейно зависима, то и вся система линейно зависима.*

Доказательство. Пусть $a_1, \ldots, \alpha_n$ — исходная система, $a_{j_1}, \ldots, a_{j_m}$ — линейно зависимая подсистема ($1 \leq j_1 < j_2 < \ldots < j_m \leq n$). По определению линейной зависимости найдутся такие $\alpha_{j_1}, \ldots, \alpha_{j_m}$, не все равные нулю, что

$$\alpha_{j_1} a_{j_1} + \ldots + \alpha_{j_m} a_{j_m} = o.$$

В последнем равенстве добавим к левой части тривиальную комбинацию векторов, не вошедших в подсистему:

$$0a_{i_1} + \ldots + 0a_{i_{n-m}},$$

где

$$\{i_1, \ldots, i_{n-m}\} = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_m\}.$$

Полученная комбинация будет нулевой, но не тривиальной. ☐

Следствие 5.21. *Любая подсистема линейно независимой системы линейно независима.*

Теорема 5.22 (Критерий линейной зависимости)**.** *Система $a_1, \ldots, a_n$, где $n \geq 2$, линейно зависима тогда и только тогда, когда*

$$a_j = L(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$$

для некоторого $j \in \{1, \ldots, n\}$.

Доказательство. Необходимость. Пусть система $a_1, \ldots, a_n$ линейно зависима, тогда $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$, причем $\alpha_j \neq 0$ для некоторого j, откуда

$$a_j = \left(-\frac{\alpha_1}{\alpha_j}\right) a_1 + \ldots + \left(-\frac{\alpha_{j-1}}{\alpha_j}\right) a_{j-1} + \left(-\frac{\alpha_{j+1}}{\alpha_j}\right) a_{j+1} + \ldots + \left(-\frac{\alpha_n}{\alpha_j}\right) a_n.$$

Достаточность. Пусть

$$a_j = \alpha_1 a_1 + \ldots \alpha_{j-1} a_{j-1} + \alpha_{j+1} a_{j+1} + \ldots + \alpha_n a_n$$

для некоторых $\alpha_1, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_n$. Тогда

$$\alpha_1 a_1 + \ldots \alpha_{j-1} a_{j-1} + \alpha_j a_j + \alpha_{j+1} a_{j+1} + \ldots + \alpha_n a_n = 0,$$

где $\alpha_j = -1$. Получили нулевую нетривиальную (так как $\alpha_j = -1 \neq 0$) комбинацию. □

Следствие 5.23. *Три вектора геометрического пространства* $\mathbf{V}_3$ *образуют линейно зависимую систему тогда и только тогда, когда векторы компланарны.*

Теорема 5.24 (Усиленный критерий линейной зависимости)**.** *Система $a_1, \ldots, a_n$ линейно зависима тогда и только тогда, когда $a_1 = o$ или $a_j = L(a_1, \ldots, a_{j-1})$ для некоторого $j \in \{1, \ldots, j-1, j+1, \ldots n\}$.*

Доказательство. Необходимость Пусть система $a_1, \ldots, a_n$ линейно зависимая, тогда $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$. Пусть j — максимальный индекс, такой, что $\alpha_j \neq 0$, тогда $\alpha_1 a_1 + \ldots + \alpha_j a_j = 0$, откуда

$$a_j = \left(-\frac{\alpha_1}{\alpha_j}\right) a_1 + \ldots + \left(-\frac{\alpha_{j-1}}{\alpha_j}\right) a_{j-1}.$$

Достаточность Если $n \geq 2$, то воспользуемся предыдущим критерием. Если $n = 1$, то $a_1 = 0$. □

Следствие 5.25. *Если система $a_1, \ldots, a_n$ линейно независима, а система $a_1, \ldots, a_n, b_1, \ldots, b_m$ — линейно зависима, то найдется такое $j \in 1, \ldots, m$, что*

$$b_j = L(a_1, \ldots, a_n, b_1, \ldots, b_{j-1}).$$

Лемма 5.26 (Теорема о замене). *Если система $a_1, \dots, a_n$ линейно выражается через систему $b_1, \dots, b_m$, причем $n \geq m$, то первая система линейно зависимая.*

Доказательство. Предположим противное. Пусть в условиях теоремы первая система линейно независима.

Так как a_1 линейно выражается через $b_1, \dots, b_m$, то объединенная система $a_1, b_1, \dots, b_m$ линейно зависима и эквивалентна $b_1, \dots, b_m$. По следствию 5.25 в новой системе найдется

$$b_{i_1} \in L(a_1, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_m),$$

поэтому система $b_1, \dots, b_m$ эквивалентна $a_1, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_m$.

Так как a_2 линейно выражается через систему $b_1, \dots, b_m$, эквивалентную $a_1, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_m$, то объединенная система $a_1, a_2, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_m$ линейно зависима и эквивалентна $b_1, \dots, b_m$. По следствию 5.25 в этой новой системе найдется

$$b_{i_2} \in L(a_1, a_2, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_{i_2-1}, b_{i_2+1}, \dots, b_m),$$

поэтому $b_1, \dots, b_m$ эквивалентна $a_1, a_2, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_{i_2-1}, b_{i_2+1}, \dots, b_m$.

Продолжая эти рассуждения далее (можно провести индукцию), мы получим цепочку эквивалентных систем

$$\begin{aligned} & b_1, \dots, b_m \sim \\ & \quad \sim a_1, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_m \sim \\ & \quad\quad \sim a_1, a_2, b_1, \dots, b_{i_1-1}, b_{i_1+1}, \dots, b_{i_2-1}, b_{i_2+1}, \dots, b_m \sim \dots \\ & \quad\quad\quad \dots \sim a_1, \dots, a_m. \end{aligned}$$

Поэтому система $b_1, \dots, b_m$ эквивалентна системе $a_1, \dots, a_m$. Следовательно, система $a_1, \dots, a_n$, выражающаяся через $b_1, \dots, b_m$, выражается через свою подсистему $a_1, \dots, a_m$, следовательно, она линейно зависима, что противоречит предположению. □

Следствие 5.27. *Эквивалентные линейно независимые системы содержат одинаковое число векторов.*

Базой системы векторов называется ее любая линейно независимая подсистема, эквивалентная исходной системе.

Следствие 5.28. *Любая ненулевая система векторов $a_1, \dots, a_m$ имеет базу.*

Доказательство. Пусть a_{i_1} — некоторый ненулевой вектор из V. Система, состоящая из одного вектора a_{i_1}, линейно независима. Поэтому, если она эквивалентна исходной системе, то является ее базой. В противном случае найдется вектор $a_{i_2} \notin L(a_{i_1})$. По усиленному критерию линейной зависимости система a_{i_1}, a_{i_2} — независимая. Если она эквивалентна исходной системе, то является ее базой. В противном случае найдется вектор $a_{i_3} \notin L(a_{i_1}, a_{i_2})$ и т. д. Описанный процесс оборвется, так как исходная система конечна. □

Следствие 5.29. *Для любой системы число векторов в произвольной базе одинаково.*

Число векторов в базе называется *рангом* системы.

Упражнение 5.30. Докажите, что база — это наибольшая (по числу векторов) линейно независимая подсистема данной системы.

Упражнение 5.31. Докажите, что база — это наименьшая (по числу векторов) подсистема данной системы, эквивалентная всей системе.

5.7. Базис линейного пространства

Система векторов $a_1, \ldots, a_n$ линейного пространства V называется *полной*, если $V = L(a_1, \ldots, a_n)$. Линейное пространство, в котором существует полная система, называется *конечномерным*. В противном случае пространство называется *бесконечномерным*.

Базисом пространства V называется полная линейно независимая система.

Утверждение 5.32. *В любом ненулевом конечномерном пространстве V существует базис.*

Доказательство. Пространство V конечномерное, поэтому $V = L(a_1, \ldots, a_n)$ для некоторых $a_1, \ldots, a_n$. В качестве базиса пространства возьмем базу системы $a_1, \ldots, a_n$. □

Утверждение 5.33. *Для любого пространства число векторов в любом базисе одинаково.*

Доказательство. Пусть $a_1, \ldots, a_n$, $b_1, \ldots, b_m$ — два базиса пространства V. Системы $a_1, \ldots, a_n$, $b_1, \ldots, b_m$ эквивалентны и линейно независимы, поэтому $n = m$. □

Число векторов в базисе конечномерного пространства V называется *размерностью пространства* V и обозначается $\dim V$. Размерность нулевого пространства считается равной нулю.

Утверждение 5.34. *Для того, чтобы линейно независимая система была базисом, необходимо и достаточно, чтобы любая система из большего числа векторов этого пространства была линейно зависимой.*

Доказательство. Необходимость. Пусть система $a_1, \ldots, a_n$ является базисом. Система векторов $b_1, \ldots, b_m$ выражается через линейно независимую систему $a_1, \ldots, a_n$ и если $m > n$, то по теореме о замене $b_1, \ldots, b_m$ линейно зависима. *Достаточность.* Пусть b — произвольный вектор из V. Рассмотрим систему $a_1, \ldots, a_n, b$. По условию она линейно зависима. Теперь из усиленного критерия линейной зависимости следует, что $b \in L(a_1, \ldots, a_n)$, т. е. система $a_1, \ldots, a_n$ — полная. □

Замечание 5.35. Утверждение 5.34 позволяет дать следующее определение базиса, эквивалентное исходному: базис — это наибольшая (по мощности) линейно независимая система. Заметим, что такое определение не означает, что базис пространства определен единственным образом.

Утверждение 5.36. *Для того, чтобы полная система $a_1, \ldots, a_n$ была базисом, необходимо и достаточно, чтобы либо $n = 1$ и $a_1 \neq o$, либо любая система из меньшего числа векторов не была бы полной.*

Доказательство. Необходимость. Пусть система $b_1, \ldots, b_m$ полная, тогда линейно независимая система $a_1, \ldots, a_n$ выражается через $b_1, \ldots, b_m$, откуда из теоремы о замене $m \geq n$.

Достаточность. Ни для какого $j = \{1, \ldots, n\}$ система $a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n$ не является базисом, поэтому $a_j \notin L(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$. Линейная независимость системы векторов $a_1, \ldots, a_n$ следует теперь из критерия линейной независимости. □

Замечание 5.37. Утверждение 5.36 позволяет дать следующее определение базиса, эквивалентное исходному: базис — это наименьшая (по мощности) полная система. Заметим, что такое определение не означает, что базис пространства определен единственным образом.

Утверждение 5.38. *Пусть* $\dim V = n$ *и система* $a_1, \ldots, a_n$ *— линейно независима, тогда она полна и, следовательно, является базисом.*

Доказательство. Так как $\dim V = n$, то любая система из большего числа векторов вектора линейно независима. Поэтому $a_1, \ldots, a_n$ полная по утверждению 5.34. □

Утверждение 5.39. *Пусть* $\dim V = n$ *и система* $a_1, \ldots, a_n$ *— полная, тогда она линейно независима и, следовательно, является базисом.*

Доказательство. Так как $\dim V = n$, то любая система из меньшего числа векторов полна. Поэтому система $a_1, \ldots, a_n$ линейно независима по утверждению 5.36. □

Размерность и базис арифметического пространства. Рассмотрим систему векторов арифметического пространства F^n:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{5.2}$$

где j-ая компонента вектора e_j равна 1, а все остальные компоненты вектора e_j равны 0 ($j = 1, 2, \ldots, n$). Покажем, что система $e_1, e_2, \ldots, e_n$ образует базис пространства F^n и, следовательно, $\dim F^n = n$.

Действительно, система (5.2) линейно независима, так как, приравнивая все компоненты в левой и правой части равенства

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = o, \tag{5.3}$$

получаем $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, т. е. нулевая комбинация (5.3) — тривиальная.

Система (5.2) — полная, так как для произвольного вектора

$$a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in F^n$$

имеем $a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$.

Систему (5.2) будем называть *стандартным базисом арифметического n-мерного пространства.*

Размерность и базис пространства геометрических векторов.

Утверждение 5.40.

1. $\dim \mathbf{V}_1 = 1$. *Система из одного вектора пространства* $\mathbf{V}_1$ *образует базис этого пространства тогда и только тогда, когда вектор ненулевой.*

2. $\dim \mathbf{V}_2 = 2$. *Система из двух векторов пространства* $\mathbf{V}_2$ *образует базис этого пространства тогда и только тогда, когда векторы неколлинеарны.*

3. $\dim \mathbf{V}_3 = 3$. *Система из трех векторов пространства* $\mathbf{V}_3$ *образует базис этого пространства тогда и только тогда, когда векторы некомпланарны.*

Доказательство.

1. Утверждение очевидно.

2. Уже было показано, что любые два вектора пространства $\mathbf{V}_2$ линейно зависимы тогда и только тогда, когда они коллинеарны. Покажем, что любые два неколлинеарных вектора a, b образуют полную систему в пространстве $\mathbf{V}_2$ и, следовательно, образуют базис этого пространства. Пусть c — произвольный вектор пространства $\mathbf{V}_2$. Как обычно, все векторы отложены из полюса O. Проведем через конец вектора c прямую, параллельную вектору b. Эта прямая пересечет прямую, на которой лежит вектор a в некоторой точке A. Затем проведем через конец вектора c прямую, параллельную вектору a. Эта прямая пересечет прямую, на которой лежит вектор b в некоторой точке B. Будем иметь $c = \overrightarrow{OA} + \overrightarrow{OB}$, $\overrightarrow{OA} = \alpha a$, $\overrightarrow{OB} = \beta b$ для некоторых α и β, откуда откуда $c = \alpha a + \beta b$. Итак, любой вектор пространства $\mathbf{V}_2$ может быть выражен в виде линейной комбинации векторов a, b, и, следовательно, a, b — полная система.

3. Доказывается аналогично. □

Размерность пространства многочленов. Покажем, что пространство $F[x]$ не является конечномерным. Действительно, предположим, что $f_1(x), f_2(x), \ldots, f_n(x)$ — некоторая полная система в пространстве $F[x]$, а m — максимум степеней многочленов из этой системы. Очевидно, что никакая их линейная комбинация не дает ни одного многочлена степени выше m. Таким образом, система $f_1(x), f_2(x), \ldots, f_n(x)$ полной не является.

Рассмотрим пространство многочленов степени не выше n. Очевидно, что система $1, x, x^2, \ldots, x^n$ образует базис этого пространства, называемый стандартным. Таким образом, пространство многочленов степени не выше n имеет конечную размерность, равную $n+1$.

5.8. Координаты векторов

Пусть $e_1, \ldots, e_n$ — базис пространства V над полем F. Тогда для произвольного вектора a из V найдутся такие числа $\alpha_1, \ldots, \alpha_n$ из F, что $a = \alpha_1 e_1 + \ldots + \alpha_n e_n$. Коэффициенты $\alpha_1, \ldots, \alpha_n$ разложения вектора по базису называются *координатами вектора*.

Утверждение 5.41. *В произвольном базисе координаты вектора определены единственным образом.*

Доказательство. Пусть $e_1, \ldots, e_n$ — базис пространства V. Предположим, что для вектора a нашлось два разложения по базису

$$a = \alpha_1 e_1 + \ldots + \alpha_n e_n = \beta_1 e_1 + \ldots + \beta_n e_n.$$

Откуда получаем

$$(\alpha_1 - \beta_1)e_1 + \ldots + (\alpha_n - \beta_n)e_n = o.$$

Так как система $e_1, \ldots, e_n$ — линейно независимая, то все коэффициенты в полученнной линейной комбинации равны 0, откуда $\alpha_j = \beta_j$ $(j = 1, \ldots, n)$. $\square$

Координаты удобно записывать в столбец, который мы будем называть *координатным столбцом* и обозначать $[a]_e$. Таким образом, если $a = \alpha_1 e_1 + \ldots + \alpha_n e_n$, то

$$[a]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Координатный столбец можно рассматривать как вектор арифметического пространства F^n. Имея в виду эту интерпретацию, получаем

Утверждение 5.42. *Для произвольных векторов a, b из V и произвольного α из F справедливо*

$$[a+b]_e = [a]_e + [b]_e,$$
$$[\alpha a]_e = \alpha [a]_e.$$

Таким образом, при сложении векторов их координаты складываются, при умножении на число умножаются на это число.

Доказательство. Пусть

$$a = \alpha_1 e_1 + \ldots + \alpha_n e_n,$$
$$b = \beta_1 e_1 + \ldots + \beta_n e_n.$$

Тогда

$$a + b = (\alpha_1 + \beta_1)e_1 + \ldots + (\alpha_n + \beta_n)e_n,$$

т. е. $[a+b]_e = [a]_e + [b]_e$. Второе равенство доказывается аналогично. □

5.9. Изоморфизм линейных пространств

Пусть пространства V и V' заданы над одним и тем же полем F. Биекция $\varphi : V \to V'$ называется *изоморфизмом*, если для любых векторов a, b из V и любого α из F справедливо

1. $\varphi(a + b) = \varphi a + \varphi b$,
2. $\varphi(\alpha a) = \alpha(\varphi a)$.

Замечание 5.43. Пусть $\dim V = n$. Из утверждения 5.42 следует, что отображение $\varphi : V \to F^n$, ставящее в соответствие вектору из V столбец его координат в некотором фиксированном базисе, является изоморфизмом.

Свойство 5.44. *Отображение, обратное к изоморфизму, является изоморфизмом.*

Доказательство. Отображение $\varphi^{-1} : V' \to V$, обратное к изоморфизму $\varphi : V \to V'$, является биекцией. Осталось показать выполнение свойств 1, 2 для отображения φ^{-1}. Пусть a', b' — произвольные векторы из V', тогда $\varphi^{-1}a' \in V$ и $\varphi^{-1}b' \in V$. Из определения изоморфизма φ имеем

$$\varphi(\varphi^{-1}a' + \varphi^{-1}b') = \varphi\varphi^{-1}a' + \varphi\varphi^{-1}b',$$
$$\varphi(\varphi^{-1}a' + \varphi^{-1}b') = a' + b'.$$

Применяя к обеим частям последнего равенства отображение φ^{-1}, получаем

$$\varphi^{-1}a' + \varphi^{-1}b' = \varphi^{-1}a' + \varphi^{-1}b'.$$

Равенство $\varphi^{-1}(\alpha a) = \alpha(\varphi^{-1}a)$ доказывается аналогично. □

Свойство 5.45. *Образом нулевого вектора пространства V при изоморфизме φ является нулевой вектор пространства V', т. е.*[3] $\varphi o = o$.

[3] Заметим, что в следующем равенстве o в левой части — вектор пространства V, тогда как o в правой части — вектор пространства V'.

Доказательство. Из определения изоморфизма

$$\varphi o = \varphi(0a) = 0\varphi a = o.$$

□

Свойство 5.46. *При изоморфизме линейно зависимая система отображается в линейно зависимую систему.*

Доказательство. Пусть система $a_1, \ldots, a_n$ линейно зависимая, тогда найдутся такие числа $\alpha_1, \ldots, \alpha_n$, что

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = o, \tag{5.4}$$

причем $\alpha_j \neq 0$ для некоторого $j \in \{1, \ldots, n\}$. Применим отображение φ к обеим частям равенства (5.4). Из определения изоморфизма получаем

$$\alpha_1(\varphi a_1) + \ldots + \alpha_n(\varphi a_n) = o.$$

Нами получена нетривиальная нулевая комбинация векторов $\varphi a_1, \ldots, \varphi a_n$, значит векторы линейно зависимые. □

Свойство 5.47. *При изоморфизме линейно независимая система отображается в линейно независимую систему.*

Доказательство. Утверждение является следствием свойств 5.44, 5.46. □

Свойство 5.48. *При изоморфизме $\varphi : V \to V'$ базис пространства V отображается в базис пространства V'.*

Доказательство. По свойству 5.46 базис отображается в линейно независимую систему. При этом по свойству 5.47 любая линейно зависимая система отображается в линейно зависимую. Таким образом, базис отображается в наибольшую линейно независимую систему, т. е. в базис. □

Пространства V и V' называются *изоморфными*, если между ними существует изоморфизм $\varphi : V \to V'$. Если V и V' изоморфны, то пишут $V \cong V'$.

Пример 5.49. Пусть $\dim V = n$, тогда, как следует из предыдущего примера, $V \cong F^n$.

Утверждение 5.50. *Отношение изоморфности пространств является отношением эквивалентности.*

Доказательство. Рефлексивность. Очевидно, что тождественное преобразование $\varepsilon : V \to V'$, определяемое формулой $\varepsilon a = a$ является изоморфизмом. Таким образом, $V \cong V$.
Симметричность. Если $V \cong V'$, то $V' \cong V$ по свойству 1.
Транзитивность. Докажем, что если $V \cong V'$ и $V' \cong V''$, то $V \cong V''$. Действительно, если φ и φ' — изоморфизмы из V в V' и из V' в V'', то произведение изоморфизмов $\psi = \varphi\varphi'$, определяемой формулой $\psi a = \varphi(\varphi' a)$ есть изоморфизм V в V''. □

Теорема 5.51 (Критерий изоморфности пространств). *Для того, чтобы пространства V и V', заданные над одним и тем же полем F, были изоморфны, необходимо и достаточно, чтобы* $\dim V = \dim V'$.

Доказательство. Необходимость Необходимость следует из свойства 5.
Достаточность Так как $V \cong F^n$ и $V' \cong F^n$, то, в силу симметричности и транзитивности отношения эквивалентности, имеем $V \cong V'$. □

5.10. Размерность подпространства

Утверждение 5.52. *Любое подпространство V_1 конечномерного линейного пространства V является конечномерным.*

Доказательство. Если $V_1 = \{o\}$, то утверждение доказано. Пусть a_1 — некоторый ненулевой вектор из V_1. Система, состоящая из одного вектора a_1, линейно независима. Поэтому, если $V_1 = L(a_1)$, то система образует базис пространства V_1. В противном случае в V_1 найдется вектор $a_2 \notin L(a_1)$. По усиленному критерию линейной зависимости система a_1, a_2 — независимая. Поэтому, если $V_1 = L(a_1, a_2)$, то система образует базис пространства V_1. В противном случае в V_1 найдется вектор $a_3 \notin L(a_1, a_2)$ и т. д. Описанный процесс оборвется по крайней мере через n шагов, где $n = \dim V$, так как любая система из $n + 1$ вектора линейно зависима. □

Утверждение 5.53. *Пусть V_1, V_2 — конечномерные подпространства, причем $V_1 \subseteq V_2$, тогда* $\dim V_1 \leq \dim V_2$.

Доказательство. Пусть $a_1, \ldots, a_l$, $b_1, \ldots, b_m$ — базисы подпространств V_1 и V_2 соответственно. Так как $V_1 \subseteq V_2$, то первая система линейно выражается через вторую и поэтому по лемме о замене $l \leq m$. □

Утверждение 5.54. *Пусть V_1, V_2 — конечномерные подпространства, причем $V_1 \subseteq V_2$ и $\dim V_1 = \dim V_2$, тогда $V_1 = V_2$.*

Доказательство. Базис $a_1, \ldots, a_l$ подпространства V_1 является линейно независимой системой векторов из V_2. Так как $l = \dim V_1 = \dim V_2$, то $a_1, \ldots, a_l$ является базисом подпространства V_2 и, следовательно, $V_1 = L(a_1, \ldots, a_l) = V_2$. □

5.11. Сумма и пересечение подпространств

Суммой подпространств V_1 и V_2 пространства V называется множество

$$V_1 + V_2 = \{x = y + z : y \in V_1, z \in V_2\}.$$

Пересечение подпространств понимается в обычном смысле:

$$V_1 \cap V_2 = \{x : x \in V_1, x \in V_2\}.$$

Утверждение 5.55. *Сумма подпространств одного и того же пространства является подпространством.*

Доказательство. Так как $o \in V_1$, $o \in V_2$, то $o + o \in V_1 + V_2$ и $V_1 + V_2 \neq \emptyset$. Докажем замкнутость относительно сложения векторов и умножения их на числа. Пусть $x \in V_1 + V_2$ и $x' \in V_1 + V_2$, тогда найдутся такие векторы y, y' из V_1 и z, z' из V_2, что $x = y + z$, $x' = y' + z'$. Так как $y + y' \in V_1$ и $z + z' \in V_2$, то $x + x' = (y + y') + (z + z') \in V_1 + V_2$. Так как $\alpha y \in V_1$ и $\alpha z \in V_2$, то $\alpha x = \alpha y + \alpha z \in V_1 + V_2$. □

Утверждение 5.56. *Пересечение подпространств одного и того же пространства является подпространством.*

Доказательство. Так как $o \in V_1$, $o \in V_2$, то $o \in V_1 \cap V_2$ и $V_1 \cap V_2 \neq \emptyset$. Докажем замкнутость относительно сложения векторов и умножения их на числа. Пусть $x \in V_1 \cap V_2$ и $x' \in V_1 \cap V_2$. Так как $x + x' \in V_1$ и $x + x' \in V_2$, то $x + x' \in V_1 \cap V_2$. Так как $\alpha x \in V_1$ и $\alpha x \in V_2$, то $\alpha x \in V_1 \cap V_2$. □

Лемма 5.57. *Любую линейно независимую систему векторов можно дополнить до базиса конечномерного пространства.*

Доказательство. Утверждение доказывается аналогично доказательству утверждения 5.52. □

Теорема 5.58 (Размерность суммы подпространств)**.** *Если V_1, V_2 — конечномерные подпространства пространства V, то*

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$$

(*формула Грассмана*).

Доказательство. Пусть $a_1, \ldots, a_k$ — базис подпространства $V_1 \cap V_2$. Векторами $b_1, \ldots, b_l$ дополним систему $a_1, \ldots, a_k$ до базиса пространства V_1. Векторами $c_1, \ldots, c_m$ дополним систему $a_1, \ldots, a_m$ до базиса пространства V_2. Для доказательства теоремы покажем, что система

$$a_1, \ldots, a_k, b_1, \ldots, b_l, c_1, \ldots, c_m$$

образует базис пространства $V_1 + V_2$.

Легко проверить полноту системы. Для доказательства ее линейной независимости рассмотрим нулевую линейную комбинацию:

$$\underbrace{\alpha_1 a_1 + \ldots + \alpha_k a_k + \beta_1 b_1 + \ldots + \beta_l b_l}_{-c \in V_1} + \underbrace{\gamma_1 c_1 + \ldots + \gamma_m c_m}_{c \in V_2} = 0. \tag{5.5}$$

Пусть

$$c = \gamma_1 c_1 + \ldots + \gamma_m c_m. \tag{5.6}$$

Очевидно, что $c \in V_2$ и тогда по (5.5) имеем $c \in V_1$. Следовательно, $c \in V_1 \cap V_2$ и поэтому найдутся такие $\delta_1, \ldots, \delta_k$, что

$$c = \delta_1 a_1 + \ldots + \delta_k a_k. \tag{5.7}$$

Из (5.6, 5.7) получаем два разложения вектора c по базису пространства V_2:

$$c = 0a_1 + \ldots + 0a_k + \gamma_1 c_1 + \ldots + \gamma_m c_m,$$
$$c = \delta_1 a_1 + \ldots + \delta_k a_k + 0c_1 + \ldots + 0c_m.$$

Так как разложение по базису единственно, то все коэффициенты в этих разложениях нулевые, поэтому $c = 0$ и, следовательно, нулевая линейная комбинация (5.5) — тривиальная. □

Понятие суммы и пересечения подпространств распространяется на произвольное конечное их число. А именно, под *суммой* $V_1 + V_2 + \ldots + V_s$ подпространств $V_1, V_2, \ldots, V_s$ понимается

$$V_1 + V_2 + \ldots + V_s = \{a_1 + a_2 + \ldots + a_s : \ a_i \in V_i \ (i = 1, 2, \ldots, s)\}.$$

Под *пересечением* $V_1 \cap V_2 \cap \ldots \cap V_s$ понимается

$$V_1 \cap V_2 \cap \ldots \cap V_s = \{a : \ a \in V_i \ (i = 1, 2, \ldots, s)\}.$$

Легко проверить справедливость утверждений 5.55 и 5.56 для таких сумм и пересечений соответственно.

5.12. Прямая сумма подпространств

Сумма подпространств $V_1 + \ldots + V_s$ называется *прямой*, если для произвольного вектора $x \in V_1 + \ldots + V_s$ векторы $x_1, \ldots, x_s$, такие, что

$$x = x_1 + \ldots + x_s, \tag{5.8}$$

где $x_i \in V_i$ $(i = 1, \ldots, s)$, определены единственным образом. Для прямой суммы будем использовать обозначение $V_1 \dot{+} \ldots \dot{+} V_s$. Представление вектора x в виде (5.8) называют его *разложением по подпространствам*. Таким образом, в случае прямой суммы это разложение определено единственным образом.

Лемма 5.59 (Единственность разложения o)**.** *Для того, чтобы сумма* $V_1 + \ldots + V_s$ *являлась прямой, необходимо и достаточно, чтобы из условий* $o = x_1 + \ldots + x_s$, $x_i \in V_i$ $(i = 1, \ldots, s)$ *следовало* $x_i = o$ $(i = 1, \ldots, s)$.

Доказательство. Необходимость следует из определения прямой суммы.
Достаточность. Пусть для некоторого x нашлось два разложения:

$$x = x_1 + \ldots + x_s = y_1 + \ldots + y_s,$$

где $x_i \in V_i$, $y_i \in V_i$ $(i = 1, \ldots, s)$. Тогда

$$o = \underbrace{(x_1 - y_1)}_{\in V_1} + \ldots + \underbrace{(x_s - y_s)}_{\in V_s}$$

Из единственности разложения вектора o теперь следует, что $x_i = y_i$ $(i = 1, \ldots, s)$. □

Теорема 5.60 (Критерий прямой суммы). *Для того, чтобы сумма $V_1 + \ldots + V_s$ конечномерных подпространств была прямой, необходимо и достаточно, чтобы*

$$\dim(V_1 + \ldots + V_s) = \dim V_1 + \ldots + \dim V_s.$$

Доказательство. Рассмотрим систему, состоящую из базисов пространств V_i:

$$\left.\begin{array}{ll} \text{базис } V_1: & a_{11}, \ldots, a_{1n_1}; \\ \text{базис } V_2: & a_{21}, \ldots, a_{2n_2}; \\ \multicolumn{2}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ \text{базис } V_s: & a_{s1}, \ldots, a_{sn_2} \end{array}\right\} \text{система } A.$$

Для доказательства теоремы покажем, что для того, чтобы сумма $V_1 + \ldots + V_s$ была прямой, необходимо и достаточно, чтобы совокупная система A образует базис этой суммы.

Необходимость. Легко видеть, что A — полная система. Докажем ее линейную независимость. Пусть

$$\sum_{j=1}^{n_1} \alpha_{1j} a_{1j} + \ldots + \sum_{j=1}^{n_s} \alpha_{sj} a_{sj} = o. \tag{5.9}$$

Проинтерпретируем приведенное равенство как разложение o по подпространствам V_i. Так как сумма подпространств прямая, то данное разложение единственно, и поэтому

$$\sum_{j=1}^{n_1} \alpha_{1j} a_{1j} = \ldots = \sum_{j=1}^{n_s} \alpha_{sj} a_{sj} = o.$$

Векторы в каждой из приведенных линейных комбинаций образуют базис, поэтому все коэффициенты равны нулю. Итак, линейная комбинация (5.9) тривиальная.

Достаточность. Для доказательства достаточности воспользуемся леммой. Для этого покажем, что если система A линейно независима, то разложение o по подпространствам V_i определяется единственным образом. Действительно, в противном случае найдутся такие векторы a_i, что

$$o = a_1 + \ldots + a_s,$$

где $a_i \in V_i$ $(i = 1, \ldots, s)$ и $a_{i'} \neq 0$ для некоторого i'. Каждый вектор a_i разложим по базису $a_{i1}, \ldots, a_{in_i}$:

$$o = \sum_{j=1}^{n_1} \alpha_{1j} a_{1j} + \ldots + \sum_{j=1}^{n_s} \alpha_{sj} a_{sj}.$$

В приведенной линейной комбинации найдутся ненулевые коэффициенты (среди чисел $\alpha_{i'1}, \ldots, \alpha_{i'n_{i'}}$), поэтому система A — линейно зависима, что противоречит условию. $\square$

Следствие 5.61. *Сумма* $V_1 + V_2$ *прямая тогда и только тогда, когда* $V_1 \cap V_2 = \emptyset$.

Доказательство. Утверждение следует из доказанного критерия и теоремы о размерности суммы подпространств. □

Пусть линейное пространство V раскладывается в прямую сумму подпространств: $V = V_1 \dot{+} V_2$. Тогда для произвольного $x \in V$ определяются единственным образом векторы $y \in V_1$, $z \in V_2$, такие, что $x = y + z$. Вектор y называется *проекцией* вектора x на подпространство V_1 параллельно V_2.

5.13. Изменение координат вектора при замене базиса

Пусть $e_1, \ldots, e_n$ и $e'_1, \ldots, e'_n$ — два базиса пространства V. Каждый вектор второго базиса, который также будем называть «новым», разложим по первому базису, называемому также «старым»:

$$e'_j = \alpha_{1j}e_1 + \ldots + \alpha_{nj}e_n \quad \text{т. е.} \quad [e'_j]_e = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{nj} \end{pmatrix} \quad (j = 1, \ldots, n). \tag{5.10}$$

Найдем связь координат произвольного вектора x в этих базисах. Пусть

$$x = x_1e_1 + \ldots + x_ne_n = x'_1e'_1 + \ldots + x'_ne'_n,$$

т. е.

$$[x]_e = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad [x]_{e'} = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

Имеем

$$x = \sum_{j=1}^{n} x'_je'_j = \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} \alpha_{ij}e_i = \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{n} \alpha_{ij}x'_j\right)}_{x_i} e_i.$$

Так как координаты вектора в любом фиксированном базисе определяются единственным образом, то

$$x_i = \sum_{j=1}^{n} \alpha_{ij}x'_j \qquad (i = 1, 2, \ldots, n)$$

(формула, связывающая координаты одного и того же вектора в разных базисах).

5.14. Системы координат

В этом разделе векторы любых линейных пространств иногда будут называться точками. Такое изменение в названии аппелирует к геометрическим представлениям: как уже отмечалось, вектор пространства $\mathbf{V}_3$, закрепленный в полюсе, можно ассоциировать с точкой, являющейся его концом.

Системой координат линейного пространства V назовем совокупность точки O из V и базиса $e_1, e_2, \ldots, e_n$ этого пространства. *Координатами точки* $A \in V$ в этой системе координат назовем координаты вектора $A - O$ в базисе $e_1, e_2, \ldots, e_n$, т. е. $[A]_{O,e} = [A - O]_e$. Исследуем как меняются координаты точки при изменении системы координат.

Пусть «старая» система координат $O, e_1, \ldots, e_n$ заменена на «новую» $O', e'_1, \ldots, e'_n$, причем справедливы формулы (5.10) и

$$O' - O = \sum_{j=1}^{n} \gamma_j e_j, \quad \text{т. е.} \quad [O']_{O,e} = [O' - O]_e = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Пусть A — произвольная точка пространства V, причем

$$A - O' = x_1 e_1 + \ldots + x_n e_n = x'_1 e'_1 + \ldots + x'_n e'_n,$$

т. е.

$$[A - O']_e = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad [A - O']_{e'} = [A]_{O',e'} = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

Так как $A - O = (O' - O) + (A - O')$, то

$$\sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} \gamma_i e_i + \sum_{j=1}^{n} x'_j e'_j = \sum_{i=1}^{n} \gamma_i e_i + \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} \alpha_{ij} e_i = \\ = \sum_{i=1}^{n} \gamma_i e_i + \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_{ij} x'_j \right) e'_i.$$

Так как координаты вектора в любом фиксированном базисе определяются единственным образом, то

$$x_i = \gamma_i + \sum_{j=1}^{n} \alpha_{ij} x'_j \qquad (i = 1, 2, \ldots, n)$$

(формула, связывающая координаты одной и той же точки в разных системах координат).

5.15. Линейное многообразие

Рассмотрим линейное пространство V над полем F. Пусть a_0 — произвольный вектор, L — подпространство в V. Под записью $a_0 + L$ будем понимать множество всевозможных векторов вида $a_0 + x$, где x — произвольный вектор из L. Множество $a_0 + L$ называется *линейным многообразием*, порожденным вектором a_0 и подпространством L. Подпространство L называется также *несущим* подпространством для многообразия $a_0 + L$.

Утверждение 5.62. *Следующие три утверждения эквивалентны*:

1. $a_0 + L = b_0 + L$,
2. $b_0 \in a_0 + L$,
3. $b_0 - a_0 \in L$.

Доказательство. Импликации $1 \to 2$, $2 \to 3$ очевидны. Докажем импликацию $3 \to 1$ Пусть $b_0 - a_0 \in L$, тогда $b_0 = a_0 + l$ для некоторого $l \in L$. Рассмотрим произвольный вектор $a = a_0 + x$ из $a_0 + L$, $x \in L$. Имеем $a = b_0 + (x - l)$, поэтому $a \in b_0 + L$, откуда $a_0 + L \subseteq b_0 + L$. Аналогично можно показать, что $b_0 + L \subseteq a_0 + L$. □

Замечание 5.63. Импликация $2 \to 1$ означает, что многообразие порождается любым своим представителем. Эквивалентность $1 \sim 3$ дает критерий совпадения двух линейных многообразий с одинаковым несущим подпространством.

Утверждение 5.64. *Пусть a_0, b_0 — векторы, L, L' — подпространства. Тогда, если $a_0 + L = b_0 + L'$, то $L = L'$. Таким образом, несущее подпространство линейного многообразия определяется единственным образом.*

Доказательство. Из предыдущего утверждения следует, что, если $a_0 + L = b_0 + L'$, то $a_0 + L = a_0 + L'$, поэтому для любого x из L найдется такой y из L', что $a_0 + x = a_0 + y$, откуда $x = y$, т. е. $x \in L'$, а, следовательно, $L \subseteq L'$. Аналогично можно показать, что $L' \subseteq L$. □

Из доказанного утверждения следует корректность следующего определения. *Размерностью* линейного многообразия $a_0 + L$ называется размерность его несущего подпространства L: $\dim(a_0 + L) = \dim L$.

Линейное многообразие размерности 1 называется *прямой*, размерности 2 — *плоскостью*[4], размерности $n - 1$ — *гиперплоскостью*.

[4] Иногда плоскостью называют линейное многообразие произвольной размерности

Глава 6

Матрицы и системы линейных уравнений

6.1. Метод Гаусса решения систем линейных уравнений

Рассмотрим произвольную *систему m линейных уравнений* с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2; \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{6.1}$$

где $a_{ij} \in F$, $b_i \in F$ ($i = 1, \ldots, m$; $j = 1, \ldots, n$). *Решением* (или *частным решением*) системы (6.1) называется набор $x = (x_1, \ldots, x_n)^\top \in F^n$ значений неизвестных x_j, при подстановке которых в систему получаем m верных тождеств. Множество всех частных решений системы будем называть *общим решением*. Если множество решений пусто, то система называется *несовместной*. Две системы назовем *эквивалентными*, если они имеют одинаковое число неизвестных и их множества решений совпадают.

Коэффициенты левой части системы (6.1) запишем в матрицу A, коэффициенты правой части системы — в столбец b:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Матрицу A будем называть *основной матрицей* системы. Блочную матрицу (A, b) будем называть *расширенной матрицей системы*.

Основная идея метода Гаусса решения системы линейных уравнений — ее упрощение с помощью серии преобразований строк расширенной матрицы, переводящих систему уравнений в эквивалентную.

Пример 6.1. Решим систему линейных уравнений

$$\begin{cases} x_1 + 2x_2 - x_3 + x_5 = 1; \\ 3x_1 + 4x_2 - x_3 + 2x_4 - x_5 = 1; \\ 4x_1 + 6x_2 - 2x_3 + 2x_4 = 2; \\ x_1 + 2x_2 - x_3 + x_4 + x_5 = 0. \end{cases} \tag{6.2}$$

Запишем расширенную матрицу:

$$\left(\begin{array}{rrrrr|r} 1 & 2 & -1 & 0 & 1 & 1 \\ 3 & 4 & -1 & 2 & -1 & 1 \\ 4 & 6 & -2 & 2 & 0 & 2 \\ 1 & 2 & -1 & 1 & 1 & 0 \end{array}\right).$$

Из второй строки вычтем первую, умноженную на 3. Из третьей строки вычтем первую, умноженную на 4. Из четвертой строки вычтем первую. Легко видеть, что данные преобразования соответствуют процедуре исключения неизвестной x_1: из первого уравнения мы выражаем x_1 и полученное выражение подставляем в остальные уравнения. Матрица примет вид:

$$\left(\begin{array}{rrrrr|r} 1 & 2 & -1 & 0 & 1 & 1 \\ 0 & -2 & 2 & 2 & -4 & -2 \\ 0 & -2 & 2 & 2 & -4 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

Вторую строку прибавим к первой м вычтем из третьей. Затем поделим вторую строку на -2. Матрица примет вид:

$$\left(\begin{array}{rrrrr|r} 1 & 0 & 1 & 2 & -3 & -1 \\ 0 & 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

Из первой строки вычтем четвертую, умноженную на 2. Ко второй строке прибавим четвертую. Переставим третью и четвертую строки. Матрица примет вид:

$$\left(\begin{array}{rrrrr|r} 1 & 0 & 1 & 0 & -3 & 1 \\ 0 & 1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right). \tag{6.3}$$

Таким образом, исходная система уравнений эквивалентна системе:

$$\begin{cases} x_1 + x_3 - 3x_5 = 1; \\ x_2 - x_3 + 2x_5 = 0; \\ x_4 = -1; \\ 0 = 0. \end{cases} \tag{6.4}$$

Выразим x_1, x_2, x_4 через x_3, x_5:

$$\begin{aligned} x_1 &= 1 - x_3 + 3x_5; \\ x_2 &= 0 + x_3 - 2x_5; \\ x_4 &= -1 \end{aligned} \tag{6.5}$$

— легко видеть, что какие бы значения не принимали *свободные неизвестные* x_3, x_5 всегда найдутся значения *связанных неизвестных*, при которых все уравнения системы (6.2), и, следовательно, системы (6.4), превращаются в верные равенства. С другой стороны, любое решение исходной системы (6.2) удовлетворяет также и системе (6.5). Таким образом, общее решение системы (6.2) можно записать в виде:

$$\begin{aligned} x_1 &= 1 - t_1 + 3t_2;\\ x_2 &= t_1 - 2t_2;\\ x_3 &= t_1;\\ x_4 &= -1\\ x_5 &= t_2, \end{aligned} \tag{6.6}$$

где t_1, t_2 — произвольные числа[1]. Решение (6.6) можно записать в «столбцовой» форме:

$$\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5 \end{pmatrix} = \begin{pmatrix} 1\\ 0\\ 0\\ -1\\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -1\\ 1\\ 1\\ 0\\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 3\\ -2\\ 0\\ 0\\ 1 \end{pmatrix}.$$

Следующие преобразования строк матрицы будем называть *элементарными преобразованиями*:

1) перестановка (транспозиция) i-й и j-й строк ($i \neq j$),

2) домножение i-й строки на число $\alpha \in F$ ($\alpha \neq 0$),

3) прибавление к i-й строке j-й строки, умноженной на α.

Лемма 6.2. *Элементарные преобразования строк расширенной матрицы системы линейных уравнений не меняют множества решений.*

Доказательство. Докажем, что преобразование 3-го типа не меняет множества решений. Запишем систему, получающуюся из (6.1) преобразованием 3-го типа: Система, полученная из (6.1) преобразованием 3-го типа, будет отличаться от (6.1) только i-м уравнением, которое примет вид:

$$(a_{i1} + \alpha a_{j1})x_1 + (a_{i2} + \alpha a_{j2})x_2 + \ldots + (a_{i1} + \alpha a_{j1})x_n = b_i + \alpha b_j. \tag{6.7}$$

Пусть $x' = (x'_1, \ldots, x'_n)^\top$ — частное решение системы (6.1). Оно удовлетворяет всем уравнениям системы (6.1) и, следовательно, всем уравнениям новой системы, в том числе уравнению (6.7). Последнее справедливо, так как тождество

$$(a_{i1} + \alpha a_{j1})x'_1 + (a_{i2} + \alpha a_{j2})x'_2 + \ldots + (a_{i1} + \alpha a_{j1})x'_n = b_i + \alpha b_j. \tag{6.8}$$

получается прибавлением к равенству

$$a_{i1}x'_1 + a_{i2}x'_2 + \ldots + a_{i1}x'_n = b_i \tag{6.9}$$

равенства

$$a_{j1}x'_1 + a_{j2}x'_2 + \ldots + a_{j1}x'_n = b_j, \tag{6.10}$$

умноженного на α. Итак, любое частное решение системы (6.1) является частным решением новой системы. Пусть теперь $x' = (x'_1, \ldots, x'_n)^\top$ — частное решение новой системы. Оно удовлетворяет всем уравнениям системы (6.1), в том числе i-му уравнению. Последнее справедливо, так как тождество (6.9) получается вычитанием из равенства (6.8) равенства (6.10), умноженного на α. Таким образом, любое решение новой системы является решением системы (6.1). □

[1] Форма записи ответа, в общем случае не единственна

Обозначим через e_j столбец, у которого все, кроме j-ой компоненты равны нулю, а j-ая компонента равна 1. Будем говорить, что матрица $A \in F^{m\times n}$ имеет *простейший*, или *упрощенный*, *вид*, если ее столбцы с номерами $j_1, \dots, j_r$ суть столбцы $e_1, \dots, e_r$ соответственно и (при $m > r$) последние $m - r$ строк — нулевые:

$$A = \begin{pmatrix} * & \dots & * & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * \\ * & \dots & * & 0 & * & \dots & * & 1 & * & \dots & * & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * & 1 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

(знаком «$*$» обозначены произвольные элементы; столбцы $e_1, e_2, \dots, e_r$ могут занимать произвольные места и не обязательно стоят в таком порядке, что $j_1 < j_2 < \dots < j_r$). Например, матрица (6.3) имеет простейший вид. Матрицу простейшего вида, которую можно получить из A элементарными преобразованиями ее строк назовем *простейшим видом* матрицы A. Заметим, что в общем случае простейший вид определяется неоднозначно.

Лемма 6.3. *С помощью элементарных преобразований строк произвольную матрицу можно привести к простейшему виду.*

Доказательство. На вход следующей процедуры (*алгоритм Гаусса—Жордана*) подается матрица $A \in F^{m\times n}$.

procedure GAUSS—JORDAN(var A)
    $i := 1$
    for $j := 1, \dots, n$ **do**
        найти $k \in \{i, i+1, \dots, m\}$, такое, что $a_{kj} \neq 0$
        если такого k не существует, то **continue**
        переставить k-ю и i-ю строки
        разделить i-ю строку на a_{ij} (*ведущий*, или *главный*, элемент)
        for $l := 1, \dots, i-1, i+1, \dots, m$ **do**
            из l-й строки вычесть i-ю, умноженную на a_{lj}
        end
        $i := i + 1$
    end
end

Легко видеть, что описанный алгоритм приводит матрицу A к простейшему виду. $\square$

Для решения системы линейных уравнений приведем *расширенную* матрицу (A, b) к простейшему виду (A', b'). Пусть столбцы $j_1, \dots, j_r$ матрицы A суть

$e_1, \dots e_r$. По матрице (A', b') построим систему уравнений, эквивалентную исходной. Если полученная система содержит уравнения вида

$$0x_1 + \ldots + 0x_n = a'_j,$$

где $a'_j \neq 0$ (т. е. в (A', b') есть строка $(0, 0, \ldots, 0, \alpha)$ с $\alpha \neq 0$), то система, очевидно, несовместна. В противном случае выразим неизвестные $x_{j_1}, \ldots, x_{j_r}$ через оставшиеся неизвестные:

$$\begin{aligned}
x_{j_1} &= c_{10} + c_{11}x_{k_1} + \ldots + c_{1n-r}x_{k_{n-r}},\\
x_{j_2} &= c_{20} + c_{21}x_{k_1} + \ldots + c_{2n-r}x_{k_{n-r}},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
x_{j_r} &= c_{r0} + c_{r1}x_{k_1} + \ldots + c_{1n-r}x_{k_{n-r}},
\end{aligned} \tag{6.11}$$

где $\{j_1, \ldots, j_r\} \cup \{k_1, \ldots, k_{n-r}\} = \{1, 2, \ldots, n\}$.

Неизвестные $x_{k_1}, \ldots, x_{k_{n-r}}$ назовем *свободными*, а $x_{j_1}, \ldots, x_{j_r}$ — *связанными*. Легко видеть, что какие бы значения не принимали свободные неизвестные всегда по формулам (6.11) найдутся значения связанных неизвестных, при которых все уравнения системы превращаются в верные равенства. С другой стороны, любое решение исходной системы линейных уравнений удовлетворяет также и системе (6.11). Таким образом, общее решение системы можно записать в виде:

$$\begin{aligned}
x_{j_1} &= c_{10} + c_{11}t_1 + \ldots + c_{1n-r}t_{n-r},\\
x_{j_2} &= c_{20} + c_{21}t_1 + \ldots + c_{2n-r}t_{n-r},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
x_{j_r} &= c_{r0} + c_{r1}t_1 + \ldots + c_{1n-r}t_{n-r},\\
x_{k_1} &= \qquad\quad t_1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
x_{k_{n-r}} &= \qquad\qquad\qquad\qquad t_{n-r},
\end{aligned} \tag{6.12}$$

где $t_1, \ldots, t_{n-r}$ — произвольные числа из F. В частности, если $n = r$, то решение системы единственно. Формулы (6.12) можно записать в «столбцовой» форме. Для простоты обозначений предположим, что $j_1 = 1, \ldots, j_r = r$ (этого всегда можно добиться перенумерацией неизвестных), тогда имеем:

$$\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_r\\ x_{r+1}\\ x_{r+2}\\ \vdots\\ x_n \end{pmatrix} = \begin{pmatrix} c_{10}\\ c_{20}\\ \vdots\\ c_{r0}\\ 0\\ 0\\ \vdots\\ 0 \end{pmatrix} + t_1 \cdot \begin{pmatrix} c_{11}\\ c_{21}\\ \vdots\\ c_{r1}\\ 1\\ 0\\ \vdots\\ 0 \end{pmatrix} + \ldots + t_{n-r} \cdot \begin{pmatrix} c_{1n-r}\\ c_{2n-r}\\ \vdots\\ c_{rn-r}\\ 0\\ 0\\ \vdots\\ 1 \end{pmatrix}. \tag{6.13}$$

Мы доказали следующие утверждения:

Теорема 6.4. *Система с расширенной матрицей* (A, b) *несовместна тогда и только тогда, когда в некотором (любом) простейшем виде этой матрицы есть строка* $(0, 0, \ldots, 0, \alpha)$, *где* $\alpha \neq 0$.

Теорема 6.5. *Если система линейных уравнений с раширенной матрицей* (A, b) *совместна, то найдется такой вектор* c_0 *и такие линейно независимые векторы* $c_1, \ldots, c_{n-r}$, *что*

$$M(A, b) = \{x = c_0 + t_1 c_1 + \ldots + t_{n-r} c_{n-r} : t_1, \ldots, t_{n-r} \in F\},$$

где r *— число строк в некотором простейшем виде матрицы* A. *В частности, если* $n = r$, *то решение* c_0 *единственное.*

Говорят, что матрица $A = (a_{ij}) \in F^{m \times n}$ имеет *ступенчатый* вид, если найдутся $j_1, j_2, \ldots, j_r$, такие, что $j_1 < j_2 < \ldots < j_r$, $a_{ik} = 0$, $a_{ij_i} \neq 0$ $(i = 1, 2, \ldots, r,\ k = 1, 2, \ldots, j_i - 1)$ и $a_{ij} = 0$ $(i = m - r + 1, \ldots, m;\ j = 1, 2, \ldots, n)$:

$$A = \begin{pmatrix} 0 & \ldots & 0 & a_{1j_1} & * & \ldots & * & * & * & \ldots & * & * & * & \ldots & * \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & a_{2j_2} & * & \ldots & * & * & * & \ldots & * \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & a_{rj_r} & * & \ldots & * \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}. \tag{6.14}$$

Элементы $a_{1j_1}, a_{1j_2}, \ldots, a_{1j_r}$ называются *ведущими*, или *главными*. Если $r = m$ и $j_i = i$ $(i = 1, 2, \ldots, r)$, то ступенчатая матрица называется *верхней треугольной*. Легко видеть, что следующая процедура (*прямой ход метода Гаусса*) приводит матрицу A к ступенчатому виду.

```
procedure GAUSS(var A)
  i := 1
  for j := 1, ..., n do
    найти k ∈ {i, i + 1, ..., m}, такое, что a_kj ≠ 0
    если такого k не существует, то continue
    переставить k-ю и i-ю строки
    for l := i + 1, ..., m do
      из l-й строки вычесть i-ю, умноженную на a_lj/a_ij
    end
    i := i + 1
  end
end
```

Один из методов решения системы линейных уравнений основан на приведении расширенной матрицы (A, b) элементарными преобразованиями строк к ступенчатому виду (A', b'). Пусть $j_1, \ldots, j_r$ — номера столбцов ведущих элементов. Как и раньше, неизвестные $x_{j_1}, \ldots, x_{j_r}$ назовем *связанными*, а остальные переменные свободными. По матрице (A', b') запишем соответствующую ей систему линейных уравнений. Если полученная система содержит уравнения вида

$$0x_1 + \ldots + 0x_n = b'_j,$$

где $b'_j \neq 0$, то она, очевидно, несовместна. В противном случае выполняем следующую процедуру, называемую *обратным ходом* метода Гаусса. Из r-го уравнения выразим неизвестную x_{j_r} через свободные неизвестные $x_{j_r+1}, \ldots, x_n$. Полученное выражение подставим в $(r - 1)$-е уравнение и выразим оттуда неизвестную $x_{j_{r-1}}$. Полученное выражение подставим в $(r - 2)$-е уравнение и т. д. до тех пор, пока не выразим все связанные неизвестные через свободные и не получим формулы вида (6.11). Описанная процедура особенно удобна в том случае, когда решение единственно, т. е. свободные переменные отсутствуют.

В заключение заметим, что ступенчатую матрицу элементарными преобразованиями строк легко привести к виду

$$A' = \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

т. е. к матрице, которая является и простейшей, и ступенчатой. Впрочем, алгоритм из доказательства леммы 6.3 сразу приводит произвольную матрицу именно к такому виду.

6.2. Ранг матрицы

Пусть $A \in F^{m\times n}$. Рассмотрим столбцы $a_1, \dots, a_n$ матрицы $A = (a_1, \dots, a_n)$ как векторы пространства F^m, а строки $\widetilde{a}_1, \dots, \widetilde{a}_m$ как векторы пространства F^n. Базу (соответственно ранг) системы $a_1, \dots, a_n$ назовем *столбцовой базой* (соответственно *столбцовым рангом*) матрицы A. Базу (соответственно ранг) системы $\widetilde{a}_1, \dots, \widetilde{a}_n$ назовем *строчечной базой* (соответственно *строчечным рангом*) матрицы A.

Утверждение 6.6. *Строчечный и столбцовый ранги матрицы простейшего вида совпадают.*

Доказательство. Непосредственно проверяется, что столбцы $j_1, \dots, j_r$ матрицы простейшего вида (т. е. векторы $e_1, \dots, e_r$) образуют столбцовую базу. С другой стороны, легко видеть, что строки $1, 2, \dots, r$ образуют строчечную базу. Таким образом, строчечный и столбцовый ранги матрицы простейшего вида равны r. □

Замечание 6.7. Также легко видеть, что строчечный и столбцовый ранги ступенчатой матрицы (6.14) раны r.

Утверждение 6.8. *Пусть матрица $A' = (a'_1, \dots, a'_n)$ получена из матрицы $A = (a_1, \dots, a_n)$ серией элементарных преобразований ее строк. Если $a_{j_1}, \dots, a_{j_r}$ — столбцовая база матрицы A, то $a'_{j_1}, \dots, a'_{j_r}$ — столбцовая база матрицы A'.*

Доказательство. Докажем, что система $a'_{j_1}, \dots, a'_{j_r}$ линейно независима. Так как система $a_{j_1}, \dots, a_{j_r}$ линейно независима, то система линейных уравнений (рассмотренная относительно неизвестных $\alpha_1, \dots, \alpha_r$)

$$\alpha_1 a_{j_1} + \ldots + \alpha_r a_{j_r} = 0 \tag{6.15}$$

имеет единственное (нулевое) решение. С матрицей этой системы осуществим те же преобразования, что и с матрицей A. Очевидно, что получим систему

$$\alpha'_1 a_{j_1} + \ldots + \alpha'_r a_{j_r} = 0.$$

Эта система эквивалентна (6.15), поэтому она тоже имеет единственное (нулевое) решение, что означает линейную независимость векторов $a'_{j_1}, \ldots, a'_{j_r}$.

Теперь покажем, что для произвольного $j \in \{1, \ldots, n\}$ вектор a'_j линейно выражается через $a'_{j_1}, \ldots, a'_{j_r}$. Система

$$\alpha_1 a_{j_1} + \ldots + \alpha_r a_{j_r} = a_j$$

совместна, поэтому совместной будет система

$$\alpha_1 a'_{j_1} + \ldots + \alpha_r a'_{j_r} = a'_j,$$

полученная из исходной той же серией элементарных преобразований, с помощью которых мы из матрицы A получили A'. □

Следствие 6.9. *При элементарных преобразованиях строк столбцовый ранг матрицы не изменяется.*

Аналогично можно доказать

Следствие 6.10. *При элементарных преобразованиях столбцов строчечный ранг матрицы не изменяется.*

Пример 6.11. Найдем столбцовую базу матрицы и выразим остальные столбцы через базу:

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ -3 & -6 & -9 & 3 & -3 \\ 2 & 4 & 6 & 1 & 2 \\ 2 & 5 & 7 & -1 & -1 \end{pmatrix}.$$

Элементарными преобразованиями со строками приведем матрицу к простейшему виду:

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ -3 & -6 & -9 & 3 & -3 \\ 2 & 4 & 6 & 1 & 2 \\ 2 & 5 & 7 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 1 & 1 & 1 & -3 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 0 & 1 & -3 & 7 \\ 0 & 1 & 1 & 1 & -3 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & 0 & 7 \\ 0 & 1 & 1 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

В полученной матрице базу образуют 1-й, 2-й и 4-й столбцы, следовательно, в матрице A также 1-й, 2-й и 4-й столбцы образуют базу. При этом 3-й столбец равен сумме 1-го и 2-го, а 5-й равен линейной комбинации 1-го и 2-го с коэффициентами 7, -3.

Утверждение 6.12. *Пусть матрица A' получена из матрицы A серией элементарных преобразований ее строк. Если $\widetilde{a}_1, \ldots, \widetilde{a}_m$ — строки матрицы A, $\widetilde{a}'_1, \ldots, \widetilde{a}'_m$ — строки матрицы A', то $L(\widetilde{a}_1, \ldots, \widetilde{a}_m) = L(\widetilde{a}'_1, \ldots, \widetilde{a}'_m)$.*

Доказательство. Рассмотрим, например, преобразование третьего типа. Легко видеть, что системы

$$\widetilde{a}_1, \ldots, \widetilde{a}_i, \ldots, \widetilde{a}_j, \ldots, \widetilde{a}_m;$$
$$\widetilde{a}_1, \ldots, \widetilde{a}_i + \alpha\widetilde{a}_j, \ldots, \widetilde{a}_j, \ldots, \widetilde{a}_m$$

эквивалентны. □

Следствие 6.13. *При элементарных преобразованиях строк строчечный ранг матрицы не изменяется.*

Аналогично можно доказать

Следствие 6.14. *При элементарных преобразованиях столбцов столбцовый ранг не изменяется.*

Теорема 6.15. *Строчечный и столбцовый ранги произвольной матрицы совпадают.*

Доказательство. Любую матрицу серией элементарных преобразований, сохраняющих столбцовый и строчечный ранги, можно привести к простейшему виду. Но для матрицы простейшего вида эти ранги равны, следовательно, они будут равны и для исходной матрицы. □

Столбцовый≡строчечный ранг назовем просто *рангом* матрицы и обозначим $\operatorname{rank} A$.

Замечание 6.16. Как мы уже видели, для нахождения столбцовой базы удобно применять элементарные преобразования строк матрицы, приводящие ее к простейшему (или ступенчатому) виду. Это позволяет также сразу найти коэффициенты линейных комбинаций, выражающих столбцы матрицы через базу. Аналогично, для нахождения строчечной базы удобно применять элементарные преобразования столбцов матрицы. Если же нужно определить только ранг матрицы, то можно применять любые элементарные преобразования строк и столбцов.

Пример 6.17. Найдем ранг матрицы

$$A = \begin{pmatrix} 25 & -75 & 50 & 25 \\ -39 & 119 & -80 & -39 \\ 18 & -57 & 39 & 18 \\ 103 & -305 & 203 & 99 \\ 97 & -293 & 191 & 117 \end{pmatrix}.$$

Ко второму столбцу прибавим первый, умноженный на 3; из третьего столбца вычтем первый, умноженный на 2; из четвертого столбца вычтем первый. Получим:

$$\begin{pmatrix} 25 & 0 & 0 & 0 \\ -39 & 2 & -2 & 0 \\ 18 & -3 & 3 & 0 \\ 103 & 4 & -3 & -4 \\ 97 & -2 & -3 & 20 \end{pmatrix}.$$

К третьему столбцу прибавим второй:

$$\begin{pmatrix} 25 & 0 & 0 & 0 \\ -39 & 2 & 0 & 0 \\ 18 & -3 & 0 & 0 \\ 103 & 4 & 1 & -4 \\ 97 & -2 & -5 & 20 \end{pmatrix}.$$

К четвертому столбцу прибавим третий, умноженный на 4:

$$\begin{pmatrix} 25 & 0 & 0 & 0 \\ -39 & 2 & 0 & 0 \\ 18 & -3 & 0 & 0 \\ 103 & 4 & 1 & 0 \\ 97 & -2 & -5 & 0 \end{pmatrix}$$

Ранг матрицы равен 3.

Упражнение 6.18. Показать, как с помощью элементарных преобразований строк найти строчечную базу матрицы A.

Теорема 6.19. *Пусть столбцы $a_{j_1}, \ldots, a_{j_r}$ составляют столбцовую базу матрицы A. Тогда элементарными преобразованиями строк матрицу A можно привести к простейшему виду, так, что на месте этих столбцов будут стоять векторы $e_1, e_2, \ldots, e_r$.*

Доказательство. Переставим столбцы $a_{j_1}, \ldots, a_{j_r}$ в начало матрицы и применим к полученной матрице алгоритм Гаусса–Жордана. Этот алгоритм приводит матрицу к виду, который является и простейшим, и ступенчатым. Если на месте первых r столбцов не получены $e_1, e_2, \ldots, e_r$, то эти столбцы не составляют базы полученной матрицы, а, следовательно, столбцы $a_{j_1}, \ldots, a_{j_r}$ не составляют базы для исходной матрицы. □

Теорема 6.20 (Кронекер–Капелли). *Система линейных уравнений с расширенной матрицей (A, b) совместна тогда и только тогда, когда* $\operatorname{rank} A = \operatorname{rank}(A, b)$.

Доказательство. Рассмотрим два доказательства теоремы.

1-*й способ*. Элементарными преобразованиями строк матрицы (A, b) приведем ее к простейшему виду (A', b'). Из теоремы 6.4 следует, что система совместна тогда и только тогда, когда $\operatorname{rank}(A', b') = \operatorname{rank} A'$, но $\operatorname{rank}(A', b') = \operatorname{rank}(A, b)$, $\operatorname{rank} A' = \operatorname{rank} A$, откуда следует доказываемое.

2-*й способ*. Систему уравнений можно записать в виде $a_1x_1 + \ldots + a_nx_n = b$, где $a_1, \ldots, a_n$ — столбцы матрицы A. Очевидно, что совместность системы эквивалентна условию $b \in L(a_1, \ldots, a_n)$. Докажем, что это эквивалентно условию $\operatorname{rank}\{a_1, \ldots, a_n\} = \operatorname{rank}\{a_1, \ldots, a_n, b\}$.

Если $b \in L(a_1, \ldots, a_n)$, то $\operatorname{rank}\{a_1, \ldots, a_n\} = \operatorname{rank}\{a_1, \ldots, a_n, b\}$.

Пусть теперь $\operatorname{rank}\{a_1, \ldots, a_n\} = \operatorname{rank}\{a_1, \ldots, a_n, b\}$. Если $a_{j_1}, \ldots a_{j_r}$ — база системы $a_1, \ldots, a_n$, то $\operatorname{rank}\left\{a_{j_1}, \ldots, a_{j_r}\right\} = \operatorname{rank}\left\{a_{j_1}, \ldots, a_{j_r}, b\right\}$, поэтому система $a_{j_1}, \ldots, a_{j_r}, b$ линейно зависима, и, следовательно, найдется вектор в этой системе, линейно выражающийся через остальные. Так как система $a_{j_1}, \ldots, a_{j_r}$ линейно независима, то этим вектором может быть только b, т. е. $b \in L(a_{j_1}, \ldots, a_{j_r}) = L(a_1, \ldots, a_n)$. □

Следствие 6.21. *Число связанных переменных совместной системы линейных уравнений равно рангу матрицы и, следовательно, не зависит от способа приведения расширенной матрицы системы к простейшему виду элементарными преобразованиями строк.*

6.3. Пространство решений системы линейных однородных уравнений

Система линейных уравнений с нулевой правой частью ($b = 0$) называется *однородной*. Исследуем множество $M(A, 0)$.

Теорема 6.22. *Пусть* $A \in F^{m\times n}$, $\operatorname{rank} A = r$. *Множество* $M(A, 0)$ *решений системы однородных уравнений с матрицей* A *есть линейное подпространство в* F^n *размерности* $n - r$.

Доказательство. Утверждение сразу следует из формул (6.13), описывающих множество $M(A, 0)$. □

Базис пространства $M(A, 0)$ называется *фундаментальной системой решений* системы линейных однородных уравнений.

Следствие 6.23. *Любая линейно независимая система из* $n - r$ *частных решений системы линейных однородных уравнений составляет ее фундаментальную систему решений.*

6.4. Множество решений системы линейных уравнений общего вида

Теорема 6.24. *Пусть* $A \in F^{m\times n}$, $b \in F^m$, $\operatorname{rank} A = r$. *Множество* $M(A, b)$ *решений совместной системы уравнений с расширенной матрицей* (A, b) *есть линейное многообразие в* F^n *размерности* $n - r$. *В частности, если система совместна, то она имеет единственное решение тогда и только тогда, когда* $n = r$.

Доказательство. Утверждение сразу следует из формул (6.13), описывающих множество $M(A, b)$. □

Теорема 6.25. *Пусть* $A \in F^{m\times n}$, $b \in F^m$, $a_0 \in M(A, 0)$, *тогда*

$$M(A, b) = x_0 + M(A, 0). \tag{6.16}$$

Таким образом, несущим подпространством множества решений системы линейных уравнений является линейное пространство решений соответствующей однородной системы.

Доказательство. Утверждение сразу следует из формул (6.13). □

Замечание 6.26. Можно дать следующую словесную формулировку теоремы: общее решение $M(A, b)$ неоднородной системы уравнений есть ее частное решение x_0 плюс общее решение $M(A, 0)$ соответствующей однородной системы.

Утверждение 6.27. *Для того, чтобы линейное уравнение являлось следствием совместной линейной системы необходимо и достаточно, чтобы оно являлось линейной комбинацией уравнений этой системы.*

Доказательство. Достаточность условий очевидна. Для доказательства необходимости заметим, что добавление уравнения-следствия к совместной линейной системе не должно менять ранга этой системы, откуда получаем требуемое. □

6.5. Описание подпространства и линейного многообразия арифметического пространства в виде множества решений системы линейных уравнений

Ранее мы видели, что множество решений однородной (соответственно неоднородной) системы линейных уравнений есть линейное подпространство (соответственно многообразие) арифметического пространства. В этом разделе мы покажем обратное: любое линейное подпространство (соответственно многообразие) арифметического пространства является множеством решений некоторой однородной (соответственно неоднородной) системы линейных уравнений.

Теорема 6.28. *Для любого подпространства L арифметического пространства F^n существует матрица $A \in F^{m\times n}$, такая, что $L = M(A, 0)$.*

Доказательство. Пусть $L = L(b_1, \dots, b_s)$, где $b_i = (b_{i1}, \dots, b_{in})^\top$ $(i = 1, \dots, s)$. Рассмотрим систему линейных однородных уравнений

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n = 0; \\ b_{21}x_1 + b_{22}x_2 + \ldots + b_{2n}x_n = 0; \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ b_{s1}x_1 + b_{s2}x_2 + \ldots + b_{sn}x_n = 0. \end{cases} \tag{6.17}$$

Пусть векторы

$$\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}, \begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{pmatrix}, \dots, \begin{pmatrix} a_{m1} \\ a_{m2} \\ \vdots \\ a_{mn} \end{pmatrix}$$

образуют фундаментальную систему решений для (6.17). Покажем, что матрица $A = (a_{ij})$ — искомая. Иными словами, покажем, что множество решений системы (6.1) есть заданное линейное подпространство. Действительно, этой системе удовлетворяют координаты каждого вектора $b_1, \dots, b_s$, следовательно, этой системе удовлетворяют координаты произвольного вектора из L. Таким образом, $L \subseteq M(A, 0)$. С другой стороны, ранг системы (6.17) равен размерности пространства L, поэтому размерность m пространства решений системы (6.17) равна $m = n - \dim L$. Откуда

$$\dim M(A, 0) = n - m = n - (n - \dim L) = \dim L.$$

Итак, $L \subseteq M(A, 0)$ и $\dim M(A, 0) = \dim L$. Следовательно, $L = M(A, 0)$. $\square$

Пример 6.29. Пусть $L = L(b_1, b_2)$, где

$$b_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Найдем систему линейных уравнений, множество решений которой совпадает с L.
1-*й способ*. Воспользуемся методом, описанным в доказательстве теоремы 6.28. Рассмотрим систему

$$\begin{cases} x_1 + x_2 = 0, \\ x_1 + x_2 + x_3 + x_4 = 0. \end{cases}$$

Фундаментальную систему ее решений составляют, например, векторы

$$b_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix},$$

поэтому в качестве искомой системы можно взять

$$\begin{cases} x_1 - x_2 = 0, \\ x_3 - x_4 = 0. \end{cases} \tag{6.18}$$

2-*й способ*. Рассмотрим матрицу B, составленную из столбцов b_1, b_2, и матрицу $\widetilde{B}$, составленную из столбцов b_1, b_2, x. Для того, чтобы $x \in L(b_1, b_2)$, необходимо и достаточно, чтобы $\operatorname{rank} B = \operatorname{rank} \widetilde{B}$. Элементарными преобразованиями строк матрицы B приведем ее к ступенчатому виду. Такие же преобразования проделаем с матрицей $\widetilde{B}$:

$$\widetilde{B} = \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 1 & 1 & x_2 \\ 0 & 1 & x_3 \\ 0 & 1 & x_4 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 0 & x_2 - x_1 \\ 0 & 1 & x_3 \\ 0 & 1 & x_4 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 1 & x_3 \\ 0 & 0 & x_2 - x_1 \\ 0 & 0 & x_4 - x_3 \end{array}\right).$$

Ранг матрицы без последнего столбца равен 2. Для того, чтобы ранг всей матрицы был также равен 2 необходимо и достаточно, чтобы $x_2 - x_1 = 0$ и $x_4 - x_3$. Эти равенства и составляют искомую систему.

Теорема 6.30. *Для любого линейного многообразия $b_0 + L$ арифметического пространства F^n существуют матрица $A \in F^{m\times n}$ и вектор $b \in F^m$, такие, что $b_0 + L = M(A, 0)$.*

Доказательство. Для подпространства L построим однородную систему линейных уравнений с матрицей A, такую, что $L = M(A, 0)$. Положим $b = Ab_0$. Из теоремы 6.24 следует, что $b_0 = M(A, b)$. □

Пример 6.31. Пусть $L = L(b_1, b_2)$, где L — такое же, как в примере 6.29, и

$$b_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Найдем систему линейных уравнений, множество решений которой совпадает с $b_0 + L$.

Подставляя компоненты вектора b_0 в левую часть системы (6.18) получим числа 1, 0, следовательно, множество решений системы

$$\begin{cases} x_1 - x_2 = 1, \\ x_3 - x_4 = 0 \end{cases}$$

совпадает с $b_0 + L$.

Описание линейного многообразия с помощью системы линейных уравнений удобно для нахождении пересечения линейных многообразий.

Пример 6.32. Найдем пересечение линейных многообразий $a_0 + L(a_1, a_2)$ и $b_0 + L(b_1, b_2)$, где

$$a_0 = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$b_0 = \begin{pmatrix} -4 \\ -3 \\ 1 \\ -2 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Находим системы линейных уравнений, определяющих каждое из многообразий:

$$\begin{cases} x_1 - x_2 - x_3 = 1, \\ 2x_1 - 2x_2 - x_4 = 0, \end{cases} \qquad \begin{cases} x_1 - 2x_2 = 2, \\ -2x_2 + x_4 = 4 \end{cases}$$

Пересечение многообразий определяет система, составленная из этих систем:

$$\begin{cases} x_1 - x_2 - x_3 = 1, \\ 2x_1 - 2x_2 - x_4 = 0, \\ x_1 - 2x_2 = 2, \\ -2x_2 + x_4 = 4. \end{cases}$$

Множество решений этой системы есть одномерное линейное многообразие $c_0 + L(c_1)$, где

$$c_0 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 4 \end{pmatrix}, \quad c_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

6.6. Матричные операции

Пусть F — произвольное поле. Таблица

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

в которой $a_{ij} \in F$ ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$), называется *матрицей* из m строк и n столбцов, или, кратко, матрицей размера $m \times n$ (а также $(m \times n)$-матрицей). Числа a_{ij} называются *элементами* матрицы, первый индекс элемента a_{ij} указывает номер строки, второй индекс указывает номер столбца. Сокращенное обозначение матрицы, указывающее общий вид ее элементов: $A = (a_{ij})$, где, разумеется, вместо i, j можно использовать другие переменные. Если $m = n$, матрица называется *квадратной* матрицей порядка n. Диагональ квадратной матрицы, идущая от левого верхнего к правому нижнему углу (т. е. составленная из элементов $a_{11}, a_{22}, \dots, a_{nn}$) называется *главной диагональю*. Диагональ, идущая от правого верхнего к левому нижнему углу называется *побочной диагональю*.

Множество всех матриц с элементами из X обозначим $X^{m\times n}$.

Матрицы из $F^{m\times n}$ можно складывать и умножать на числа из F. При этом получаются матрицы из $F^{m\times n}$. Под *суммой* $A+B$ матриц $A=(a_{ij})\in F^{m\times n}$, $B=(b_{ij})\in F^{m\times n}$ понимается матрица $C=(c_{ij})\in F^{m\times n}$, в которой $c_{ij}=a_{ij}+b_{ij}$ ($i=1,2,\ldots,m$; $j=1,2,\ldots,n$). Под *произведением* αA матрицы $A=(a_{ij})\in F^{m\times n}$ на число $\alpha\in F$ понимается матрица $C=(c_{ij})\in F^{m\times n}$, в которой $c_{ij}=\alpha a_{ij}$.

Легко доказываются следующие свойства этих операций.

Утверждение 6.33. *Пусть A, B, C — матрицы из $F^{m\times n}$, а α, β — числа из F. Тогда*

1) $A+(B+C)=(A+B)+C$;
2) $A+B=B+C$;
3) $A+O=A$, *где O —* нулевая *матрица размера $m\times n$, т. е. матрица, в которой все элементы равны* 0;
4) *для любой матрицы A найдется матрица B, называемая* противоположной *и обозначаемая $-A$, такая, что $A+B=O$; элементы противоположной матрицы равны $b_{ij}=-a_{ij}$.*
5) $(\alpha\beta)A=\alpha(\beta A)$;
6) $1\cdot A=A$;
7) $(\alpha+\beta)A=\alpha A+\beta A$;
8) $\alpha(A+B)=\alpha A+\alpha B$.

Следствие 6.34. *Относительно операций сложения матриц и умножения матрицы на число $F^{m\times n}$ является линейным пространством над полем F размерности mn. В частности, $F^{m\times n}$ изоморфно арифметическому пространству F^{mn}.*

Так как $F^{m\times n}$ — линейное пространство, то все результаты, полученные для линейных пространств относятся и к $F^{m\times n}$. В частности, в $F^{m\times n}$ вводится операция вычитания и т. д.

Пусть $A=(a_{ij})\in F^{m\times n}$, $B=(b_{jl})\in F^{n\times k}$. Под *произведением* матрицы A на матрицу B называется матрица $C=(c_{il})\in F^{m\times k}$, в которой

$$c_{il}=\sum_{j=1}^{n}a_{ij}b_{jl}=a_{i1}b_{1l}+a_{i2}b_{2l}+\ldots+a_{in}b_{nl}.$$

Пример 6.35.

$$(a_1,a_2,\ldots,a_n)\cdot\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=\sum_{j=1}^{n}a_jb_j=a_1b_1+a_2b_2+\ldots+a_nb_n;$$

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \cdot (b_1, b_2, \ldots, b_n) = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_n \\ \hdotsfor{4} \\ a_m b_1 & a_m b_2 & \ldots & a_m b_n \end{pmatrix};$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Пример 6.36. Пусть $A = (a_{ij}) \in F^{m\times n}$, $b = (b_i) \in F^m$, $x = (b_j) \in F^n$. Систему линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

можно записать в матричном виде следующим образом: $Ax = b$.

Утверждение 6.37.

1) *Если* $A \in F^{m\times n}$, $B \in F^{n\times p}$, $C \in F^{p\times q}$, *то* $(AB)C = A(BC)$ (*ассоциативность*);

2) *Если* $A \in F^{m\times n}$, $B \in F^{n\times p}$, $C \in F^{n\times p}$, *то* $A(B + C) = AB + AC$ (*дистрибутивность I*);

3) *Если* $A \in F^{m\times n}$, $B \in F^{m\times n}$, $C \in F^{n\times p}$, *то* $(A + B)C = AC + BC$ (*дистрибутивность II*);

4) *Если* $A \in F^{m\times n}$, $B \in F^{n\times p}$, *то* $\alpha(AB) = (\alpha A)B$.

5) *Если* $A \in F^{m\times n}$, *то* $AE_n = A$ *и* $E_m A = A$, *где* E_n, E_m — единичные *матрицы порядков* n *и* m *соответственно*, *т. е. матрицы*, *в которых все элементы равны* 0, *кроме элементов главной диагонали*, *равных* 1.

Доказательство. Все свойства проверяются непосредственно. Докажем, например, 1-е свойство. Во-первых, все действия, которые необходимо выполнить при вычислении левой и правой частей доказываемого равенства, осуществимы (размеры матриц согласованы для вычисления всех произведений). Далее, размеры матриц, получающиеся в левой части, совпадают с размерами матрицы, получающейся в правой части, и равны $m \times q$. Докажем теперь, что соответствующие элементы совпадают. Пусть

$$A = (a_{ij}), \qquad B = (b_{jk}), \qquad C = (c_{kl}).$$

Обозначим

$$D = AB = (d_{ik}), \qquad G = BC = (g_{jl}),$$

$$F = (AB)C = (f_{il}), \quad H = A(BC) = (h_{il}).$$

По определению матричного произведения имеем

$$d_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Подставляя полученное выражение в формулу для f_{il}, получаем:

$$f_{il} = \sum_{k=1}^{p} d_{ik} c_{kl} = \sum_{k=1}^{p} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) c_{kl} = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij} b_{jk} c_{kl}.$$

По определению матричного произведения имеем

$$g_{jl} = \sum_{k=1}^{p} b_{jk} c_{kl}.$$

Подставляя полученное выражение в формулу для h_{il}, получаем:

$$h_{il} = \sum_{j=1}^{n} a_{ij} g_{jl} = \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{p} b_{jk} c_{kl} \right) = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij} b_{jk} c_{kl}.$$

Имеем $f_{il} = h_{il}$ $(i = 1, 2, \ldots, m,\ l = 1, 2, \ldots, q)$, откуда $F = H$. □

Утверждение 6.38. *j-й столбец произведения AB есть линейная комбинация столбцов матрицы A с коэффициентами, взятыми из j-го столбца матрицы B. i-я строка произведения AB есть линейная комбинация строк матрицы B с коэффициентами, взятыми из i-й строки матрицы A.*

Следствие 6.39. *Если $A \in F^{m\times n}$, $B \in F^{n\times p}$, то*

$$\operatorname{rank}(AB) \le \operatorname{rank} A, \qquad \operatorname{rank}(AB) \le \operatorname{rank} B.$$

Множество $F^{n\times n}$ квадратных матриц замкнуто относительно умножения. Отметим, что умножение не обладает свойством коммутативности.

Действия с матрицами, разбитыми на клетки. Матрицы, составленные из других матриц называются *блочными* матрицами, или матрицами, *разбитыми на клетки*. Пусть $A \in F^{m\times n_1}$, $B \in F^{m\times n_2}$. Матрицу, полученную приписыванием к A справа матрицы B обозначим $(A\ B)$. Пусть теперь $A \in F^{m_1\times n}$, $B \in F^{m_2\times n}$. Матрицу, полученную приписыванием к A снизу матрицы B обозначим

$$\begin{pmatrix} A \\ B \end{pmatrix}.$$

Можно строить матрицы из большего числа блоков, например, из блоков $A_{ij} \in F^{m_i\times n_j}$ $(i = 1, 2, \ldots, q;\ j = 1, 2, \ldots, r)$ можно составить матрицу

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1r} \\ A_{21} & A_{22} & \ldots & A_{2r} \\ \multicolumn{4}{c}{\dotfill} \\ A_{q1} & A_{q2} & \ldots & A_{qr} \end{pmatrix} \in F^{m\times n}, \tag{6.19}$$

где $m = m_1 + m_2 + \ldots + m_q$, $n = n_1 + n_2 + \ldots + n_r$. Пусть также $B_{jk} \in F^{n_j \times p_k}$ ($j = 1, 2, \ldots, r$; $k = 1, 2, \ldots, s$) и

$$B = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1s} \\ B_{21} & B_{22} & \ldots & B_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ B_{r1} & B_{r2} & \ldots & B_{rs} \end{pmatrix} \in F^{n \times p}, \tag{6.20}$$

где $p = p_1 + p_2 + \ldots + p_s$.

Утверждение 6.40. *Пусть матрицы A и B построены из блоков согласно* (6.19) *и* (6.20), *тогда*

$$C = \begin{pmatrix} C_{11} & C_{12} & \ldots & C_{1s} \\ C_{21} & C_{22} & \ldots & C_{2s} \\ \ldots & \ldots & \ldots & \ldots \\ C_{q1} & C_{q2} & \ldots & C_{qs} \end{pmatrix} \in F^{m \times p},$$

где

$$C_{ik} = A_{i1}B_{1k} + A_{i2}B_{2k} + \ldots + A_{ir}B_{rk} \qquad (i = 1, 2, \ldots, q;\ k = 1, 2, \ldots, s).$$

Доказательство. Следует из определения матричного умножения. □

Пример 6.41.

$$A \cdot (B_1,\ B_2) = (AB_1,\ AB_2)$$

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \cdot B = \begin{pmatrix} A_1 B \\ A_2 B \end{pmatrix}$$

$$(A_1,\ A_2) \cdot \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1 B_1 + A_2 B_2$$

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \cdot \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

Элементарные преобразования как умножение матриц. Пусть E_{ij} — матрица, полученная из единичной перестановкой ее i-й и j-й строк:

$$E_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix};$$

$E_i(\alpha)$ — матрица, полученная из единичной умножением i-й строки на число α:

$$E_i(\alpha) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix};$$

$E_{ij}(\alpha)$ — матрица, полученная из единичной прибавлением к i-й строке j-й, умноженной на α:

$$E_{ij}(\alpha) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & \alpha & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Матрицы E_{ij}, $E_i(\alpha)$, $E_{ij}(\alpha)$ называются матрицами *элементарных преобразований*.

Утверждение 6.42. *Элементарные преобразования со строками матрицы $A \in F^{m\times n}$ можно осуществить домножая A* слева *на матрицы элементарных преобразований*:

- *умножение на E_{ij} осуществляет перестановку строк с номерами i и j,*
- *умножение на $E_i(\alpha)$ — умножение i-й строки на число α,*
- *умножение на $E_{ij}(\alpha)$ — прибавление к i-й строке j-й, умноженной на α.*

Элементарные преобразования со столбцами матрицы $A \in F^{m\times n}$ можно осуществить умножая A справа *на матрицы элементарных преобразований*:

- *умножение на E_{ij} осуществляет перестановку столбцов с номерами i, j,*
- *умножение на $E_i(\alpha)$ — умножение i-го столбца на число α,*
- *умножение на $E_{ij}(\alpha)$ — прибавление к j-му столбцу i-го, умноженного на α.*

Доказательство. Проверяется непосредственно. □

Транспонирование матриц. Пусть $A = (a_{ij}) \in F^{m\times n}$. Матрица $B = (b_{ji}) \in F^{n\times m}$, в которой $b_{ji} = b_{ij}$, называется *транспонированной к* A матрицей и обозначается $A^\top$.

Утверждение 6.43.

1) *Если* $A \in F^{m\times n}$, $B \in F^{m\times n}$, *то* $(A+B)^\top = A^\top + B^\top$.

2) *Если* $A \in F^{m\times n}$, $B \in F^{n\times k}$, *то* $(AB)^\top = B^\top A^\top$.

Доказательство. Первое свойство очевидно. Докажем второе свойство. Во-первых, заметим, что размеры матриц, получающихся в левой и правой частях доказываемого равенства, совпадают. Пусть $A = (a_{ij})$, $B = (b_{jk})$, $C = AB = (c_{ik})$, $D = (AB)^\top = (d_{ki})$, тогда

$$d_{ki} = c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}.$$

$F = A^\top = (f_{ji})$, $G = B^\top = (g_{kj})$, $H = B^\top A^\top = (h_{ki})$, тогда

$$f_{ji} = a_{ij}, \quad g_{kj} = b_{jk}, \quad h_{ki} = \sum_{j=1}^{n} g_{kj} f_{ji} = \sum_{j=1}^{n} b_{jk} a_{ij}.$$

Получаем, что $d_{ki} = h_{ki}$ $(k = 1, 2, \ldots, k;\ i = 1, 2, \ldots, m)$. $\square$

Глава 7

Определители

7.1. Определители второго и третьего порядков

Рассмотрим систему двух линейных уравнений с двумя неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases} \tag{7.1}$$

Вычитая из первого уравнения, умноженного на a_{22}, второе, умноженное на a_{12}, получим

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}, \tag{7.2}$$

откуда при $a_{11}a_{22} - a_{12}a_{21} \neq 0$

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}. \tag{7.3}$$

Вычитая из второго уравнения системы (7.1), умноженного на a_{11}, первое, умноженное на a_{21}, получим

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = b_1a_{11} - b_2a_{21}, \tag{7.4}$$

откуда при $a_{11}a_{22} - a_{12}a_{21} \neq 0$

$$x_2 = \frac{b_1a_{11} - b_2a_{21}}{a_{11}a_{22} - a_{12}a_{21}}. \tag{7.5}$$

Легко проверить, что условие $a_{11}a_{22} - a_{12}a_{21} \neq 0$ эквивалентно тому, что ранг основной матрицы

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{7.6}$$

системы (7.1) равен 2, откуда следует, что ранг расширенной матрицы системы тоже равен 2 — система совместна и имеет единственное решение. Так как уравнения (7.2), (7.4) являются следствиями системы (7.1), то ее единственное решение можно найти по формулам (7.3), (7.5).

Заметим, что общий знаменатель значений неизвестных просто выражается через элементы матрицы (7.6): он равен произведению элементов главной диагонали минус произведение элементов побочной диагонали. Это число называется *определителем* (второго порядка) матрицы (7.6). Определитель матрицы A обозначается либо $\det A$, либо заменой круглых скобок, заключающих элементы матрицы, прямыми чертами. Таким образом,

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

Теперь формулы (7.3), (7.5) можно записать следующим образом:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{7.7}$$

Эти формулы называются *формулами Крамера*.

Пример 7.1. Решим систему

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

Определитель из коэффициентов есть

$$\Delta = \begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} = 2 \cdot (-1) - 1 \cdot 3 = -5.$$

Он отличен от нуля, поэтому применимы формулы Крамера. Числителями для значений неизвестных будут

$$\Delta_1 = \begin{vmatrix} 2 & 1 \\ -1 & -1 \end{vmatrix} = 2 \cdot (-1) - 1 \cdot (-1) = -1, \quad \Delta_2 = \begin{vmatrix} 2 & 2 \\ 3 & -1 \end{vmatrix} = 2 \cdot (-1) - 2 \cdot 3 = -8.$$

Таким образом, решением системы является набор

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{1}{5}, \qquad x_2 = \frac{\Delta_2}{\Delta} = \frac{8}{5}.$$

Теперь рассмотрим систему трех линейных уравнений с тремя неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{7.8}$$

с матрицей коэффициентов

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{7.9}$$

Если умножить первое уравнение системы (7.8) на $a_{22}a_{33} - a_{23}a_{32}$, второе уравнение — на $a_{13}a_{32} - a_{12}a_{33}$, третье уравнение — на $a_{12}a_{23} - a_{13}a_{22}$, а затем сложить три полученных уравнения, то легко проверить, что коэффициенты при x_2, x_3 окажутся равными 0, и мы получим равенство

$$\begin{gathered}(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32})x_1 = \\ = b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - a_{13}a_{22}b_3 - a_{12}b_2a_{33} - b_1a_{23}a_{32}.\end{gathered} \tag{7.10}$$

Коэффициент при x_1 в этом равенстве называется *определителем* (третьего порядка) матрицы (7.9). Для определителей третьего порядка используется такая же символика, что и для определителей второго порядка. Таким образом,

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

Правая часть равенства (7.10) также представляет из себя определитель третьего порядка, а именно определитель матрицы, получающейся из (7.9) заменой элементов первого столбца коэффициентами из правой части:

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}.$$

Если $\Delta \neq 0$, то из (7.10) выводим

$$x_1 = \frac{\Delta_1}{\Delta}. \tag{7.11}$$

Аналогично можно получить

$$x_2 = \frac{\Delta_2}{\Delta}, \qquad x_3 = \frac{\Delta_3}{\Delta}, \tag{7.12}$$

где

$$\Delta_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \qquad \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

Условие $\Delta \neq 0$ равносильно тому, что решение системы (7.8) существует и единственно, однако этого не так проста как в случае систем двух уравнений с двумя неизвестными. Далее мы докажем более общее утверждение. Впрочем, проверить, что формулы (7.11), (7.12) дают решение системы (7.8) можно простой подстановкой этих формул в систему. Формулы (7.11), (7.12), как и (7.7), называются *формулами Крамера*.

Пример 7.2. Решим систему

$$\begin{cases} 2x_1 + x_2 + 2x_3 = 2, \\ 3x_1 - x_2 - 3x_3 = -1, \\ -x_1 + x_2 - 4x_3 = 3, \end{cases}$$

Определитель из коэффициентов есть

$$\Delta = \begin{vmatrix} 2 & 1 & 2 \\ 3 & -1 & -3 \\ -1 & 1 & -4 \end{vmatrix} = 2\cdot(-1)\cdot(-4) + 1\cdot(-3)\cdot(-1) + 2\cdot 3\cdot 1 - \\ -2\cdot(-1)\cdot(-1) - 1\cdot 3\cdot(-4) - 2\cdot(-3)\cdot 1 = 33.$$

Он отличен от нуля, поэтому применимы формулы Крамера:

$$\Delta_1 = \begin{vmatrix} 2 & 1 & 2 \\ -1 & -1 & -3 \\ 3 & 1 & -4 \end{vmatrix} = 5, \qquad \Delta_2 = \begin{vmatrix} 2 & 2 & 2 \\ 3 & -1 & -3 \\ -1 & 3 & -4 \end{vmatrix} = 72,$$

$$\Delta_3 = \begin{vmatrix} 2 & 1 & 2 \\ 3 & -1 & 3 \\ -1 & 1 & -1 \end{vmatrix} = -8,$$

откуда

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{5}{33}, \qquad x_2 = \frac{\Delta_2}{\Delta} = \frac{24}{11}, \qquad x_3 = \frac{\Delta_3}{\Delta} = -\frac{8}{33}.$$

7.2. Перестановки и подстановки

Перестановкой попарно различных элементов $k_1, \ldots, k_n$ называется упорядоченный набор $\pi = (i_1, \ldots, i_n)$, в котором $i_p \neq i_s$ при $p \neq s$ и для каждого p найдется t, такое, что $i_p = i_t$. Например, (b, a, c, f, e, d) — перестановка элементов (букв) a, b, c, d, e, f.

Далее, если не оговорено противное, рассматриваются перестановки элементов $1, \ldots, n$.

Имеется только одна перестановка из 1 элемента, две перестановки из 2 элементов: $(1, 2)$ и $(2, 1)$, 6 перестановок из 3 элементов: $(1, 2, 3)$, $(1, 3, 2)$, $(2, 1, 3)$, $(2, 3, 1)$, $(3, 1, 2)$, $(3, 2, 1)$.

Утверждение 7.3. *Число перестановок из n элементов равно*

$$n! = 1 \cdot 2 \cdot \ldots \cdot n.$$

Доказательство. В качестве i_1 в перестановке $(i_1, i_2, \ldots, i_n)$ можно выбрать любое из чисел $1, 2, \ldots, n$ — всего n возможностей. Если i_1 уже выбрано, то в качестве i_2 можно выбрать одно из $n - 1$ чисел, таким образом, число различных способов выбрать i_1 и i_2 равно $n(n - 1)$ и т. д. □

Транспозицией элементов i_p и i_s $(p \neq s)$ в перестановке $i_1, \ldots, i_p, \ldots, i_s, \ldots, i_n$ называется ее преобразование, при котором перестановка переходит в перестановку $i_1, \ldots, i_s, \ldots, i_p, \ldots, i_n$ (элементы i_p и i_s поменялись местами, остальные элементы остались на прежних местах).

Утверждение 7.4. *Все $n!$ перестановок из n элементов можно расположить в таком порядке, что каждая следующая перестановка получается из предыдущей одной транспозицией, причем начинать можно с любой перестановки.*

Доказательство. Докажем теорему методом математической индукции по параметру n. Основание индукции очевидно. Теперь предположим, что все $(n-1)!$ перестановок из n элементов можно получить с помощью транспозиций. Вначале в качестве исходной рассмотрим перестановку $1, 2, \ldots, n$. По предположению индукции с помощью только транспозиций можно получить все перестановки элементов $2, \ldots, n$. В начало каждой такой перестановки припишем 1. В результате получим все $(n-1)!$ перестановок из n элементов с первым элементом 1. В последней из выписанных перестановок совершим транспозицию элементов 1 и 2. Выпишем все перестановки элементов, стоящих на 2-ом, 3-ом, …, n-ом местах. В начало каждой такой перестановки припишем 2. В результате получим все $(n-1)!$ перестановок из n элементов с первым элементом 2. В последней из выписанных перестановок совершим транспозицию элементов 2 и 3 и т. д.

Общее количество выписанных перестановок равно $n(n-1)! = n!$, причем все эти перестановки попарно различны (на первом месте стояли по очереди все n элементов).

Нам осталось показать, что для того, чтобы получить все перестановки с помощью одних транспозиций, можно начинать с любой перестановки. Для доказательства рассмотрим исходную перестановку и перенумеруем ее элементы. $\square$

Будем говорить, что элементы i_p и i_s, где $p < s$, образуют *инверсию* в перестановке $\pi = (i_1, \ldots, i_n)$, если $i_p > i_s$. Общее число инверсий в перестановке обозначим $\sigma(\pi) = \sigma(i_1, \ldots, i_n)$. В случае $n = 1$ примем $\sigma(\pi) = 1$. Перестановка π называется *четной*, если $\sigma(\pi)$ — четно, перестановка π называется *нечетной*, если $\sigma(\pi)$ — нечетно.

Например, в перестановке $\pi = (2, 3, 1)$ имеется две инверсии: их образуют пары $2, 1$ и $3, 1$, таким образом, $\sigma(\pi) = 2$, перестановка четная.

Утверждение 7.5. *Транспозиция меняет четность перестановки.*

Доказательство. Вначале рассмотрим транспозицию соседних элементов i_p и i_{p+1}:

$$(i_1, \ldots, i_{p-1}, i_p, i_{p+1}, i_{p+2}, \ldots, i_n) \to (i_1, \ldots, i_{p-1}, i_{p+1}, i_p, i_{p+2}, \ldots, i_n).$$

Легко видеть, что если $i_p < i_{p+1}$, то число инверсий увеличивается на 1 (появляется новая инверсия элементов i_p, i_{p+1}); если $i_p > i_{p+1}$, то число инверсий уменьшается на 1 (исчезает инверсия элементов i_p, i_{p+1}). Итак, транспозиция соседних элементов меняет четность.

Теперь рассмотрим транспозицию элементов i_p и i_s, где $s > p + 1$. Покажем, что такую инверсию можно осуществить с помощью серии транспозиций соседних элементов. Сначала передвигаем i_p вправо на место элемента i_s:

$$\begin{array}{cll}
 & (i_1,\ldots,i_{p-1},i_p,i_{p+1},i_{p+2},\ldots,i_{s-2},i_{s-1},i_s,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_p,i_{p+2},\ldots,i_{s-2},i_{s-1},i_s,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_p,\ldots,i_{s-2},i_{s-1},i_s,i_{s+1},\ldots i_n) & \to \\
 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_{p+3},\ldots,i_p,i_{s-1},i_s,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_{p+3},\ldots,i_{s-1},i_p,i_s,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_{p+3},\ldots,i_{s-1},i_s,i_p,i_{s+1},\ldots i_n) &
\end{array}$$

Затем передвигаем элемент i_s влево:

$$\begin{array}{cll}
 & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_{p+3},\ldots,i_{s-1},i_s,i_p,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_{p+3},\ldots,i_s,i_{s-1},i_p,i_{s+1},\ldots i_n) & \to \\
 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_{p+2},i_s,\ldots,i_{s-2},i_{s-1},i_p,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_{p+1},i_s,i_{p+2},\ldots,i_{s-2},i_{s-1},i_p,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_s,i_{p+1},i_{p+2},\ldots,i_{s-2},i_{s-1},i_p,i_{s+1},\ldots i_n) & \to \\
\to & (i_1,\ldots,i_{p-1},i_s,i_{p+1},i_{p+2},\ldots,i_{s-2},i_{s-1},i_p,i_{s+1},\ldots i_n). &
\end{array}$$

Мы выполнили нечетное ($2(s-p)-1$) число транспозиций соседних элементов. Каждая транспозиция меняет четность перестановки. В результате всей серии четность перестановки изменилась. □

Из предыдущих двух утверждений получаем

Следствие 7.6. *При $n \geq 2$ число четных перестановок из n элементов совпадает с числом нечетных перестановок и равно $n!/2$.*

Подстановкой называется биекция φ конечного множества в себя. Чтобы однозначно определить подстановку достаточно указать образ каждого элемента. Это можно сделать, записав таблицу

$$\tau = \begin{pmatrix} k_1 & k_2 & \ldots & k_n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix}, \tag{7.13}$$

в которой в первой строке указаны в некотором фиксированном порядке элементы множества, а во втором — произвольная их перестановка. Преобразование определяется правилом $\varphi k_p = i_p$ ($p = 1,\ldots,n$). Если порядок $k_1, k_2, \ldots, k_n$ фиксирован, то любой подстановке соответствует единственная таблица вида (7.13). С другой стороны, любой таблице с таким свойством соответствует единственная подстановка.

Рассмотрим подстановки множества $1, 2, \ldots, n$. Тогда в таблице (7.13) каждая строка есть перестановка элементов $1, 2, \ldots, n$. Обозначим $\sigma(\tau) = \sigma(k_1, k_2, \ldots, k_n)+$

$\sigma(i_1, i_2, \dots, i_n)$. Удобным способом задания подстановок множества $1, 2, \dots, n$ является задание с помощью таблицы с фиксированной верхней строкой:

$$\tau = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix},$$

Утверждение 7.7. *Четность величины $\sigma(\tau)$ совпадает для всех таблиц τ, определяющих одну и ту же подстановку.*

Доказательство. Для получения всех таблиц, определяющих одну и ту же подстановку достаточно рассмотреть произвольную перестановку столбцов таблицы τ. Эти перестановки можно получить с помощью лишь одних транспозиций столбцов. Каждая такая транспозиция меняет четность обеих перестановок и, следовательно, сохраняет четность величины $\sigma(\tau)$. □

Пример 7.8. Рассмотрим подстановку

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

В первой строке имеем 5 инверсий, во второй строке — 2. Суммарное число инверсий равно 7, подстановка четная. Зададим ту же подстановку другой таблицей:

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

В первой строке нет инверсий, во второй строке имеется 5 инверсий. Таким образом, при разных записях одной и той же подстановки сохраняется четность общего чисал инверсий, но не само это число.

7.3. Определитель. Комбинаторное определение

Напомним определения определителей 2-го и 3-го порядков:

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

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

Мы видим, что в обоих случаях определитель есть алгебраическая сумма произведений некоторых элементов матрицы. Каждое слагаемое в этих суммах называется *членом* определителя. Каждый член определителя второго порядка есть произведение двух элементов матрицы, стоящих в разных строках и разных столбцах, при этом всевозможные произведения (их всего 2) использованы. Один член участвует в алгебраической сумме со знаком плюс, другой — со знаком минус. Каждый член определителя третьего порядка есть произведение трех элементов матрицы, стоящих в разных строках и разных столбцах, причем всевозможные произведения (их всего 6) использованы. Половина членов участвует в алгебраической сумме со знаком плюс, другая половина со знаком минус.

Пусть имеется квадратная матрица n-го порядка

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \tag{7.14}$$

с элементами из некоторого поля F. Рассмотрим всевозможные произведения из n элементов этой матрицы, стоящих в разных строках и разных столбцах, т. е. произведения вида

$$a_{1j_1}a_{2j_2}\dots a_{nj_n}, \tag{7.15}$$

где $(j_1, j_2, \dots, j_n)$ — перестановка чисел $1, 2, \dots, n$. Число таких произведений равно числу перестановок, т. е. $n!$. Будем считать все такие произведения членами определителя n-го порядка.

Для определения знака, с каким входит в состав определителя произведение (7.15), рассмотрим четность перестановки $(j_1, j_2, \dots, j_n)$. Заметим, что для определителей 2-го и 3-го порядков если эта перестановка четная, то соответсвующий ей член входит со знаком плюс, если перестановка нечетная, то член входит со знаком минус. Естественно сохранить это свойство и для определителей n-го порядка. Вместо четности перестановки можно взять четность подстановки

$$\begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}.$$

Мы приходим к следующему определению: *определителем*, или *детерминантом*, матрицы (7.14) называется алгебраическая сумма $n!$ всевозможных произведений элементов матрицы, взятых по одному из каждой строки и каждого столбца, причем это произведение входит со знаком плюс, если его индексы составляют четную подстановку, и со знаком минус в противоположном случае. Данное определение определителя будем называть *комбинаторным*, или *определением через полное разложение* (1-я точка зрения на определители).

Определитель матрицы A будем обозначать $\det A$. Также для обозначения определителя матрицы будем заменять круглые скобки, обрамляющие элементы матрицы, прямыми полосами, как и для определителей 2-го и 3-го порядков. Таким образом,

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{\pi = (j_1, j_2, \dots, j_n)} (-1)^{\sigma(\pi)} a_{1j_1}a_{2j_2}\dots a_{nj_n},$$

где суммирование идет по всем перестановкам $\pi = (j_1, j_2, \dots, j_n)$ элементов

$1, 2, \ldots, n$. Из определения сразу следует

$$\det A = \sum_{\tau = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix}} (-1)^{\sigma(\tau)} a_{i_1 j_1} a_{i_2 j_2} \ldots a_{i_n j_n},$$

где суммирование идет по всем подстановкам τ из n элементов, и, например,

$$\det A = \sum_{\pi = (i_1, i_2, \ldots, i_n)} (-1)^{\sigma(\pi)} a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}.$$

Приведенные три формулы называются *формулами полного разложения* определителя.

Определитель матрицы первого порядка — это (единственный) элемент этой матрицы.

На определители распространяется некоторая матричная терминология. В частности, *порядком*, *элементами*, *строкой*, *столбцом* определителя $\det A$ называются соответственно порядок, элементы, строка и столбец матрицы A.

7.4. Свойства определителя

Утверждение 7.9. *Определитель не меняется при транспонировании, т. е.* $\det A = \det A^\top$.

Доказательство. Очевидно, определители $\det A$ и $\det A^\top$ состоят из одинаковых членов. Остается доказать, что эти члены входят с одинаковыми знаками. Рассмотрим член $a_{1j_1} a_{1j_2} \ldots a_{1j_n}$ определителя $\det A$. Его знак определяется четностью подстановки

$$\tau = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix}.$$

Знак того же члена в разложении определителя $\det A^\top$ определяется, очевидно, четностью подстановки

$$\tau' = \begin{pmatrix} j_1 & j_2 & \ldots & j_n \\ 1 & 2 & \ldots & n \end{pmatrix}.$$

Перестановки τ и τ' в общем случае различны, но, легко видеть, имеют одинаковую четность. □

Из утверждения 7.9 следует, что всякое утверждение о столбцах определителя справедливо и для его строк и наоборот. Далее мы формулируем и доказываем свойства определителя, касающиеся его столбцов. Аналогичные свойства справедливы и для его строк.

Пусть $a_1, a_2, \ldots, a_n$ — столбцы матрицы $A \in F^{n \times n}$, тогда под $\det(a_1, a_2, \ldots, a_n)$ будем понимать $\det A$.

Утверждение 7.10. *Определитель матрицы, содержащей нулевой столбец, равен нулю, т.е.* $\det(a_1, \ldots, a_{j-1}, o, a_{j+1}, \ldots, a_n) = 0$.

Доказательство. В каждый член определителя входит один элемент j-го столбца. Так как любой элемент этого столбца равен нулю, то определитель равен нулю. □

Утверждение 7.11. *Транспозиция столбцов меняет знак определителя*: $\det(\ldots, a_j, \ldots, a_k, \ldots) = -\det(\ldots, a_k, \ldots, a_j, \ldots)$ *при* $j \neq k$.

Доказательство. Пусть $B = (b_{ij})$ — матрица, получающаяся из A транспозицией j-го и k-го столбцов. Очевидно, определители $\det A$ и $\det B$ состоят из одниковых членов. Рассмотрим знаки, с которыми эти члены входят в разложение определителей. Общий член определителя $\det B$ есть

$$b_{i_1 1} \ldots b_{i_j j} \ldots b_{i_k k} \ldots b_{i_n n} = a_{i_1 1} \ldots a_{i_j k} \ldots a_{i_k j} \ldots a_{i_n n}.$$

Его знак определяется четностью подстановки

$$\tau = \begin{pmatrix} i_1 & \ldots & i_j & \ldots & i_k & \ldots & i_n \\ 1 & \ldots & j & \ldots & k & \ldots & n \end{pmatrix},$$

в то время как знак того же члена в разложении определителя $\det A$ определяется четностью подстановки

$$\tau' = \begin{pmatrix} i_1 & \ldots & i_j & \ldots & i_k & \ldots & i_n \\ 1 & \ldots & k & \ldots & j & \ldots & n \end{pmatrix}.$$

Подстановка τ' получается одной транспоцизией в верхней строке из подстановки τ, следовательно, τ и τ' имеют разные четности. Таким образом, одни и те же члены входят в разложение определителей $\det A$ и $\det B$ с обратными знаками. □

Утверждение 7.12. *Определитель матрицы с двумя одинаковыми столбцами равен нулю, т.е.* $\det(a_1, \ldots, a_j, \ldots, a_k, \ldots, a_n) = 0$ *при* $a_j = a_k$ *и* $j \neq k$.

Доказательство. Вместе с членом

$$a_{i_1 1} \ldots a_{i_j j} \ldots a_{i_k k} \ldots a_{i_n n}$$

рассматриваемого определителя рассмотрим член

$$a_{i_1 1} \ldots a_{i_k j} \ldots a_{i_j k} \ldots a_{i_n n}.$$

Так как $a_j = a_k$, то эти члены равны, однако знаки, с которым они входят в состав определителя различны. Поэтому сумма этих членов равна нулю, следовательно равен нулю сам определитель. □

Утверждение 7.13. *При умножении столбца, определитель умножается на это число, т.е.* $\det(a_1, \ldots, \alpha a_j, \ldots, a_n) = \alpha \det(a_1, \ldots, a_j, \ldots, a_n)$.

Доказательство. При умножении j-го столбца на число α каждый член определителя умножается на α, следовательно, весь определитель умножается на α. $\square$

Утверждение 7.14. *Если j-й столбец определителя представлен в виде суммы $b + c$, где $b \in F^n$, $c \in F^n$, то определитель равен сумме двух определителей, первый из которых получается заменой j-го столбца исходного определителя столбцом b, а второй — заменой того же столбца столбцом c, т. е.* $\det(a_1, \ldots, b+c, \ldots, a_n) = \det(a_1, \ldots, b, \ldots, a_n) + \det(a_1, \ldots, c, \ldots, a_n)$.

Доказательство. Следует из равенства

$$a_{i_1 1} \ldots (b_{i_j} + c_{i_j}) \ldots a_{i_n n} \;=\; a_{i_1 1} \ldots a_{i_j} \ldots a_{i_n n} \;+\; a_{i_1 1} \ldots b_{i_j} \ldots a_{i_n n}.$$

$\square$

Из утверждений 7.13 и 7.14 получаем следующее обобщение:

Следствие 7.15.

$$\det\left(a_1, \ldots, \sum_{i=1}^{k} \beta_i b_i, \ldots, a_m\right) = \sum_{i=1}^{k} \beta_i \det(a_1, \ldots, b_i, \ldots, a_m),$$

т. е. определитель есть линейная функция по каждому своему столбцу (строке).

Утверждение 7.16. *Определитель с линейно зависимыми столбцами равен нулю.*

Доказательство. Так как столбцы линейно зависимы, то найдется столбец, скажем, j-й, линейно выражающийся через остальные. Используя следствие 7.15 и утверждение 7.12, получаем

$$\det A = \det\left(a_1, \ldots, a_{j-1}, \sum_{i \neq j} \alpha_i a_i, a_{j+1}, \ldots, a_m\right) =$$

$$= \sum_{i \neq j} \alpha_i \det(a_1, \ldots, a_{j-1}, a_i, a_{j+1}, \ldots, a_m) = 0.$$

$\square$

Утверждение 7.17. *Определитель не меняется, если к одному из его столбцов прибавить другой столбец, умноженный на произвольное число:*

$$\det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = \det(a_1, \ldots, a_i + \alpha a_j, \ldots, a_j, \ldots, a_n).$$

Доказательство. Из утверждений 7.13, 7.14 и 7.12 получаем:

$$\det(a_1, \ldots, a_i + \alpha a_j, \ldots, a_j, \ldots, a_n) =$$

$$= \det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) + \alpha \det(a_1, \ldots, a_j, \ldots, a_j, \ldots, a_n) =$$

$$= \det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n).$$

$\square$

Замечание 7.18. Утверждения 7.11, 7.13, 7.17 устанавливают, как меняется определитель при элементарных преобразованиях его строк и столбцов: при транспозиции строк (столбцов) определитель меняет знак, при умножении строки (столбца) на число определитель умножается на это число, при прибавлении к строке (столбцу) другой строки (столбца), умноженного на число, определитель не изменяется. В частности, если матрица A' получена из A серией элементарных преобразований ее строк (столбцов), то $\det A = 0$ тогда и только тогда, когда $\det A' = 0$.

Утверждение 7.19. *Определитель верхней треугольной матрицы равен произведению диагональных элементов*:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11}a_{22}\dots a_{nn}.$$

Доказательство. Докажем, что все члены определителя, кроме, быть может, $a_{11}a_{22}\dots a_{nn}$, равны нулю. Пусть $a_{i_1 1}a_{i_2 2}\dots a_{i_n n}$ — член определителя. Если $i_1 \neq 1$, то $a_{i_1 1} = 0$ и рассматриваемый член равен нулю. Поэтому будем считать, что $i_1 = 1$, поэтому $i_2 \neq 1$. Если, кроме того, $i_2 \neq 2$, то $a_{i_2 2} = 0$ и рассматриваемый член снова равен нулю. Поэтому будем считать, что $i_2 = 2$, поэтому $i_3 \neq 1$, $i_3 \neq 2$ и т. д. Приходим к выводу, что среди всех членов рассматриваемого определителя лишь $a_{11}a_{22}\dots a_{nn}$ может не быть равным нулю. □

Очевидно, аналогичное утверждение справедливо и для нижних треугольных матриц.

С помощью элементарных преобразований определитель можно привести к треугольному виду. Согласно замечанию 7.18 при этом определитель будет изменяться известным образом. На этом основан практичный метод вычисления определителей.

Пример 7.20. Вычислим определитель

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

Переставляем первую и вторую строки, при этом определитель сменит знак. Далее вычитаем первую строку из остальных с подходящими множителями (2, 3, 1 соответственно), так, чтобы занулить в первом столбце все элементы ниже диагонали:

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -3 & -6 \\ 0 & -5 & -11 & -9 \\ 0 & 0 & -6 & -5 \end{vmatrix}.$$

Вычитаем из третьей строки вторую:

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -3 & -6 \\ 0 & 0 & -8 & -3 \\ 0 & 0 & -6 & -5 \end{vmatrix}.$$

Вычитаем из четвертой строки третью, умноженную на $\dfrac{6}{8} = \dfrac{3}{4}$:

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -3 & -6 \\ 0 & 0 & -8 & -3 \\ 0 & 0 & 0 & -\dfrac{11}{4} \end{vmatrix} = -1 \cdot (-5) \cdot (-8) \cdot \left(-\frac{11}{4}\right) = 110$$

Теорема 7.21. *Определитель равен нулю тогда и только тогда, когда его столбцы (строки) линейно зависимы.*

Доказательство. Достаточность была доказана в утверждении 7.16. Докажем необходимость. Элементарными преобразованиями строк и/или столбцов матрицы A приведем ее к ступенчатому виду A':

$$A' = \begin{pmatrix} 0 & \dots & 0 & \alpha_1 & * & \dots & * & * & * & \dots & * & * & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & \alpha_2 & * & \dots & * & * & * & \dots & * \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \alpha_r & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

где $\alpha_j \neq 0$ $(j = 1, 2, \dots, r)$. Заметим, что $\det A = 0$ тогда и только тогда, когда $\det A' = 0$ (см. замечание 7.18). Однако из утверждения 7.19 следует, что $\det A' = 0$ тогда и только тогда $\operatorname{rank} A' = \operatorname{rank} A < n$. $\square$

Квадратная матрица $A \in F^{n \times n}$ называется *вырожденной*, а также *сингулярной*, или *особенной*, если $\det A = 0$, в противном случае матрица называется *невырожденной*. Из теоремы 7.21 следует, что A вырождена тогда и только тогда, когда $\operatorname{rank} A < n$.

Пусть $A = (a_{ij}) \in F^{n \times n}$. Говорят, что A имеет *диагональное преобладание*, если

$$|a_{ii}| > |a_{i1}| + \ldots + |a_{i,i-1}| + |a_{i,i+1}| + \ldots + |a_{in}| \qquad (i = 1, 2, \dots, n).$$

Утверждение 7.22. *Матрица с диагональным преобладанием невырождена.*

Доказательство. Предположим противное: пусть $\det A = 0$. Тогда найдутся числа $\alpha_1, \alpha_2, \dots, \alpha_n$, не все равные нулю, такие, что

$$\sum_{j=1}^{n} \alpha_j a_{ij} = 0 \qquad (i = 1, 2, \dots, n). \tag{7.16}$$

Пусть α_s — среди этих чисел максимальное по абсолютной величине. Из (7.16) при $i = s$ получаем

$$\alpha_s a_{ss} = -\alpha_1 a_{s1} - \ldots - \alpha_{s-1} a_{s,s-1} - \alpha_{s+1} a_{s,s+1} - \ldots - \alpha_n a_{sn},$$

откуда

$$\begin{gathered} |\alpha_s a_{ss}| \le |\alpha_1 a_{s1}| + \ldots + |\alpha_{s-1} a_{s,s-1}| + |\alpha_{s+1} a_{s,s+1}| + \ldots + |\alpha_n a_{sn}| \le \\ \le |\alpha_s|(|a_{s1}| + \ldots + |a_{s,s-1}| + |a_{s,s+1}| + \ldots + |a_{sn}|) < |\alpha_s| \cdot |a_{ss}|. \end{gathered}$$

Противоречие. □

7.5. Миноры. Теорема Лапласа

Пусть $A = (a_{ij}) \in F^{m\times n}$. Обозначим

$$A\begin{pmatrix} i_1, i_2, \ldots, i_k \\ j_1, j_2, \ldots, j_k \end{pmatrix} = \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_k} \\ a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_k} \end{vmatrix}.$$

Если $i_1 < i_2 < \ldots < i_k$ и $j_1 < j_2 < \ldots < j_k$, то

$$A\begin{pmatrix} i_1, i_2, \ldots, i_k \\ j_1, j_2, \ldots, j_k \end{pmatrix}$$

называется *минором* (k-го порядка) матрицы A. Таким образом, минор — это определитель матрицы, элементы которой стоят на пересечении выбранных строк и столбцов матрицы A. Если $i_1 = 1, \ldots, i_k = k$, $j_1 = 1, \ldots, j_k = k$, то минор называется *угловым*. Если $i_1 = j_1, \ldots i_k = j_k$, то минор называется *главным*. Если

$$m = n, \quad \{i_{k+1}, i_{k+2}, \ldots, i_n\} = \{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_k\},$$
$$\{j_{k+1}, j_{k+2}, \ldots, j_n\} = \{1, 2, \ldots, n\} \setminus \{j_1, j_2, \ldots, j_k\}$$

и

$$i_{k+1} < i_{k+2} < \ldots < i_n, \quad j_{k+1} < j_{k+2} < \ldots < j_n,$$

то минор

$$A\begin{pmatrix} i_{k+1}, i_{k+2}, \ldots, i_n \\ j_{k+1}, j_{k+2}, \ldots, j_n \end{pmatrix},$$

т. е. определитель матрицы, полученной вычеркиванием из A заданных строк и столбцов, называется *дополнительным* к минору

$$A\begin{pmatrix} i_1, i_2, \ldots, i_k \\ j_1, j_2, \ldots, j_k \end{pmatrix};$$

при этом

$$(-1)^{i_{k+1}+i_{k+2}+\ldots+i_n+j_{k+1}+j_{k+2}+\ldots+j_n} A\begin{pmatrix} i_{k+1}, i_{k+2}, \ldots, i_n \\ j_{k+1}, j_{k+2}, \ldots, j_n \end{pmatrix}$$

называется *алгебраическим дополнением* к этому минору.

Лемма 7.23. *Произведение минора определителя на алгебраическое дополнение после умножения членов минора на члены алгебраического дополнения есть алгебраическая сумма, слагаемые которой являются некоторыми членами определителя, причем их знаки в этой сумме совпадают со знаками, с которыми они входят в состав определителя.*

Доказательство. Легко видеть, что все слагаемые алгебраической суммы, получающейся после умножения членов минора на члены алгебраического дополнения, встречаются в разложении исходного определителя. Остается показать, что знаки слагаемых этой суммы совпадают со знаками соответствующих членов исходного определителя.

Вначале рассмотрим случай $i_1 = 1, \ldots, i_k = k$, $j_1 = 1, \ldots, j_k = k$. Произвольный член $a_{1j_1}a_{2j_2}\ldots a_{kj_k}$ минора $A\begin{pmatrix}1,2,\ldots,k\\1,2,\ldots,k\end{pmatrix}$ имеет в этом члене знак, определяемый четностью подстановки

$$\begin{pmatrix} 1 & 2 & \ldots & k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix}, \tag{7.17}$$

в то время как член $a_{k+1,k+l_1}a_{k+2,k+l_2}\ldots a_{n,k+l_k}$ минора $A\begin{pmatrix}k+1,k+2,\ldots,n\\k+1,k+2,\ldots,n\end{pmatrix}$ имеет знак, определяемый четностью подстановки

$$\begin{pmatrix} 1 & 2 & \ldots & n-k \\ l_1 & l_2 & \ldots & l_{n-k} \end{pmatrix}. \tag{7.18}$$

Произведение этих членов $a_{1j_1}a_{2j_2}\ldots a_{kj_k}a_{k+1,k+l_1}a_{k+2,k+l_2}\ldots a_{n,k+l_{n-k}}$ поэтому имеет знак

$$(-1)^{\sigma(j_1,j_2,\ldots,j_k)}\cdot(-1)^{\sigma(l_1,l_2,\ldots,l_{n-k})}.$$

Однако именно с этим знаком член $a_{1j_1}a_{2j_2}\ldots a_{kj_k}a_{k+1,k+l_1}a_{k+2,k+l_2}\ldots a_{n,k+l_{n-k}}$ входит в разложение исходного определителя, так как число инверсий в подстановке

$$\begin{pmatrix} 1 & 2 & \ldots & k & k+1 & k+2 & \ldots & n \\ j_1 & j_2 & \ldots & j_k & k+l_1 & k+l_2 & \ldots & k+l_{n-k} \end{pmatrix}.$$

равно сумме инверсий в подстановках (7.17) и (7.18), т. е. $\sigma(j_1 + j_2 + \ldots + j_k) + \sigma(l_1 + l_2 + \ldots + l_k)$.

Теперь рассмотрим общий случай. Переставим строки и столбцы определителя так, чтобы рассматриваемый минор стоял в левом верхнем углу. Для этого переставим i_1-ю строку с (i_1-1)-й, затем с (i_1-2)-й и т. д., пока i_1-я строка не займет место первой; затем переставим i_2-ю строку с (i_2-1)-й, затем с (i_2-2)-й и т. д., пока i_2-я строка не займет место второй строки. Аналогичные действия произведем с i_3-й строкой и т. д., пока строки с номерами $i_1, i_2, \ldots, i_k$ не займут верхние k мест. Аналогичные действия произведем со столбцами $j_1, j_2, \ldots, j_k$. В конце всех действий минор $A\begin{pmatrix}i_1,i_2,\ldots,i_k\\j_1,j_2,\ldots,j_k\end{pmatrix}$ окажется в левом верхнем углу, при этом взаимное расположение его строк и столбцов останется неизменным.

Дополнительный к нему минор будет стоять в правом нижнем углу, взаимное расположение его строк и столбцов также останется неизменным. Всего будет произведено

$$(i_1 - 1) + (i_2 - 2) + \ldots + (i_k - k) + (j_1 - 1) + (j_2 - 2) + \ldots + (j_k - k)$$

транспозиций строк и столбцов матрицы A. Легко видеть, что четность этого числа совпадает с четностью $i_1 + i_2 + \ldots + i_k + j_1 + j_2 + \ldots + j_k$. □

Теорема 7.24 (Разложение определителя по строке). *Определитель матрицы равен сумме произведений элементов выбранной, скажем, i-й, строки на их алгебраические дополнения*:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij},$$

где M_{ij} — определитель матрицы, полученной из A вычеркиванием i-й строки и j-го столбца. Аналогичное утверждение справедливо и для любого столбца матрицы.

Доказательство. По лемме 7.23 произведение a_{ij} на алгебраическое дополнение $(-1)^{i+j}M_{ij}$ является суммой части членов, входящих в разложение $\det A$, с теми же знаками, с которыми они встречаются в этом разложении. Очевидно, количество этих членов равно числу членов минора M_{ij}, т. е. $(n-1)!$. С другой стороны, все члены, встречающиеся в произведениях

$$a_{i1}M_{i1}, a_{i2}M_{i2}, \ldots, a_{in}M_{in}$$

попарно различны и общее их число равно $n \cdot (n-1)! = n!$, т. е. совпадает с общим числом членов в $\det A$, откуда получаем требуемое. □

Пример 7.25.

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

(разложение по первой строке).

Пусть $x_1, x_2, \ldots, x_n$ — некоторые числа. *Определителем Вандермонда* называется

$$W(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix}. \tag{7.19}$$

Утверждение 7.26.

$$W(x_1, x_2, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i). \tag{7.20}$$

Доказательство. Формулу (7.20) докажем индукцией по n. При $n = 2$ имеем

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

Теперь предположим, что формула (7.20) верна для определителей $(n-1)$-го порядка. Докажем ее для определителей n-го порядка. Из n-го столбца определителя (7.19) вычтем $(n-1)$-й, умноженный на x_1, затем из $(n-1)$-го столбца вычтем $(n-2)$-й, умноженный на x_1 и т. д., наконец, из 2-го столбца вычтем 1-й, умноженный на x_1. Получим:

$$W(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_2 x_1 & \ldots & x_2^{n-1} - x_2^{n-2} x_1 \\ \multicolumn{5}{c}{\dotfill} \\ 1 & x_n - x_1 & x_n^2 - x_n x_1 & \ldots & x_n^{n-1} - x_n^{n-2} x_1 \end{vmatrix}.$$

Разлагая определитель по первой строке, придем к определителю $(n-1)$-го порядка. Вынесем из каждой строки этого определителя общий множитель:

$$W(x_1, x_2, \ldots, x_n) = (x_2 - x_1)(x_3 - x_1)\ldots(x_n - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \ldots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \ldots & x_3^{n-2} \\ \multicolumn{5}{c}{\dotfill} \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-2} \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)\ldots(x_n - x_1) \cdot W(x_2, x_3, \ldots, x_n).$$

Применяя к $W(x_2, x_3, \ldots, x_n)$ предположение индукции:

$$W(x_1, x_2, \ldots, x_n) = (x_2 - x_1)(x_3 - x_1)\ldots(x_n - x_1) \prod_{2 \le i < j \le n} (x_j - x_i),$$

получаем требуемое. $\square$

Следствие 7.27. *Определитель Вандермонда $W(x_1, x_2, \ldots, x_n)$ равен нулю тогда и только тогда, когда среди чисел $x_1, x_2, \ldots, x_n$ есть по крайней мере два равных.*

Пример 7.28.

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

Приведем другой способ вывода формулы (7.20). Снова воспользуемся индукцией по n. Будем считать, что формула справедлива для порядка определителя, меньшего n, и докажем ее для определителя Вандермонда n-го порядка. Предположим, что $x_i \neq x_j$ при $i \neq j$ (в противном случае определитель равен нулю и доказываемая формула верна). Рассмотрим $W(x_1, x_2, \ldots, x_n)$ как многочлен относительно x_1. Степень этого многочлена равна $n-1$ и старший коэффициент равен $W(x_2, x_3, \ldots, x_{n-1}) \neq 0$ — в этом можно легко убедиться, если разложить $W(x_1, x_2, \ldots, x_n)$ по первой строке. Заметим однако, что числа $x_2, x_3, \ldots, x_n$ являются корнями этого многочлена: при $x_1 = x_i$ ($i = 2, 3, \ldots, n$) определитель равен нулю, поэтому

$$W(x_1, x_2, \ldots, x_{n-1}) = W(x_2, x_3, \ldots, x_{n-1})(x_1 - x_2)(x_1 - x_3) \ldots (x_1 - x_{n-1}).$$

Применяя теперь к $W(x_2, x_3, \ldots, x_{n-1})$ предположение индукции, получаем требуемое.

Пример 7.29. Вычислим определитель

$$\Delta = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{s-1} & x_1^{s+1} & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^{s-1} & x_2^{s+1} & \ldots & x_2^n \\ \multicolumn{8}{c}{\dotfill} \\ 1 & x_n & x_n^2 & \ldots & x_n^{s-1} & x_n^{s+1} & \ldots & x_n^n \end{vmatrix}.$$

Рассматриваемый определитель равен взятому со знаком $(-1)^s$ коэффициенту многочлена $W(x, x_1, x_2, \ldots, x_n)$, стоящему при x^s: чтобы убедиться в этом, достаточно разложить $W(x, x_1, x_2, \ldots, x_n)$ по первой строке. Однако по утверждению 7.26

$$W(x, x_1, x_2, \ldots, x_n) = \prod_{j=1}^{n}(x_j - x) \prod_{1 \le i < j \le n}(x_j - x_i).$$

В многочлене, стоящем справа, коэффициент при x^s равен

$$(-1)^s \prod_{1 \le < j_1 < j_2 < \ldots < j_{n-s} \le n} x_{j_1} x_{j_2} \ldots x_{j_{n-s}} \prod_{1 \le i < j \le n}(x_j - x_i).$$

Полученное значение без множителя $(-1)^s$ и есть Δ.

Определитель

$$C(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n) = \begin{vmatrix} \dfrac{1}{a_1 + b_1} & \dfrac{1}{a_1 + b_2} & \ldots & \dfrac{1}{a_1 + b_n} \\ \dfrac{1}{a_2 + b_1} & \dfrac{1}{a_2 + b_2} & \ldots & \dfrac{1}{a_2 + b_n} \\ \multicolumn{4}{c}{\dotfill} \\ \dfrac{1}{a_n + b_1} & \dfrac{1}{a_n + b_2} & \ldots & \dfrac{1}{a_n + b_n} \end{vmatrix}$$

называется *определителем Коши*.

Утверждение 7.30.

$$C(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n) = \frac{\prod\limits_{j<i}(a_i - a_j)(b_i - b_j)}{\prod\limits_{i,j}(a_i + b_j)}.$$

Доказательство. Воспользуемся индукцией по n. При $n = 1$ утверждение справедливо. Теперь предположим, что доказываемая формула верна для определителей $(n-1)$-го порядка. Вычтем в определителе $C = C(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n)$ последний столбец из всех предыдущих:

$$C(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n) =$$

$$\begin{vmatrix} \dfrac{b_n - b_1}{(a_1 + b_1)(a_1 + b_n)} & \dfrac{b_n - b_2}{(a_1 + b_2)(a_1 + b_n)} & \ldots & \dfrac{b_n - b_{n-1}}{(a_1 + b_{n-1})(a_1 + b_n)} & \dfrac{1}{a_1 + b_n} \\ \dfrac{b_n - b_1}{(a_2 + b_1)(a_2 + b_n)} & \dfrac{b_n - b_2}{(a_2 + b_2)(a_2 + b_n)} & \ldots & \dfrac{b_n - b_{n-1}}{(a_2 + b_{n-1})(a_2 + b_n)} & \dfrac{1}{a_2 + b_n} \\ \multicolumn{5}{c}{\dotfill} \\ \dfrac{b_n - b_1}{(a_n + b_1)(a_n + b_n)} & \dfrac{b_n - b_2}{(a_n + b_2)(a_n + b_n)} & \ldots & \dfrac{b_n - b_{n-1}}{(a_n + b_{n-1})(a_n + b_n)} & \dfrac{1}{a_n + b_n} \end{vmatrix}.$$

Из каждой строки вынесем множитель $\dfrac{1}{a_i+b_n}$ $(i=1,2,\dots,n)$. Из каждого столбца, кроме последнего, вынесем множитель b_n-b_j $(j=1,2,\dots,n-1)$:

$$C=\frac{\prod\limits_{j=1}^{n-1}(b_n-b_j)}{\prod\limits_{i=1}^{n}(a_i+b_n)}\cdot\begin{vmatrix}\dfrac{1}{a_1+b_1} & \dfrac{1}{a_1+b_2} & \dots & \dfrac{1}{a_1+b_{n-1}} & 1\\ \dfrac{1}{a_2+b_1} & \dfrac{1}{a_2+b_2} & \dots & \dfrac{1}{a_2+b_{n-1}} & 1\\ \dots & \dots & \dots & \dots & \dots\\ \dfrac{1}{a_n+b_1} & \dfrac{1}{a_n+b_2} & \dots & \dfrac{1}{a_n+b_{n-1}} & 1\end{vmatrix}.$$

В полученном определителе вычтем последнюю строку из всех остальных:

$$C=\frac{\prod\limits_{j=1}^{n-1}(b_n-b_j)}{\prod\limits_{i=1}^{n}(a_i+b_n)}\times$$

$$\begin{vmatrix}\dfrac{a_n-a_1}{(a_1+b_1)(a_n+b_1)} & \dfrac{a_n-a_1}{(a_1+b_2)(a_n+b_2)} & \dots & \dfrac{a_n-a_1}{(a_1+b_{n-1})(a_n+b_{n-1})} & 0\\ \dfrac{a_n-a_2}{(a_2+b_1)(a_n+b_1)} & \dfrac{a_n-a_2}{(a_2+b_2)(a_n+b_2)} & \dots & \dfrac{a_n-a_2}{(a_2+b_{n-1})(a_n+b_{n-1})} & 0\\ \dots & \dots & \dots & \dots & \dots\\ \dfrac{a_n-a_{n-1}}{(a_{n-1}+b_1)(a_n+b_1)} & \dfrac{a_n-a_{n-1}}{(a_{n-1}+b_2)(a_n+b_2)} & \dots & \dfrac{a_n-a_{n-1}}{(a_{n-1}+b_{n-1})(a_n+b_{n-1})} & 0\\ \dfrac{1}{a_n+b_1} & \dfrac{1}{a_n+b_2} & \dots & \dfrac{1}{a_n+b_{n-1}} & 1\end{vmatrix}.$$

Полученный определитель разложим по последнему столбцу и вынесем общие множители:

$$C=\frac{\prod\limits_{j=1}^{n-1}(b_n-b_j)\prod\limits_{i=1}^{n-1}(a_n-a_i)}{\prod\limits_{i=1}^{n}(a_i+b_n)\prod\limits_{j=1}^{n-1}(a_n+b_j)}\times\begin{vmatrix}\dfrac{1}{a_1+b_1} & \dfrac{1}{a_1+b_2} & \dots & \dfrac{1}{a_1+b_{n-1}}\\ \dfrac{1}{a_2+b_1} & \dfrac{1}{a_2+b_2} & \dots & \dfrac{1}{a_2+b_{n-1}}\\ \dots & \dots & \dots & \dots\\ \dfrac{1}{a_{n-1}+b_1} & \dfrac{1}{a_{n-1}+b_2} & \dots & \dfrac{1}{a_{n-1}+b_{n-1}}\end{vmatrix}.$$

Осталось применить к полученному определителю предположение индукции. □

Пример 7.31.

$$\begin{vmatrix}\dfrac{1}{a_1+b_1} & \dfrac{1}{a_1+b_2}\\ \dfrac{1}{a_2+b_1} & \dfrac{1}{a_2+b_2}\end{vmatrix}=\frac{(b_1-b_2)(a_1-a_2)}{(a_1+b_1)(a_1+b_2)(a_2+b_1)(a_2+b_2)}.$$

Следствие 7.32 (Определитель Гильберта).

$$\begin{vmatrix}1 & \dfrac{1}{2} & \dfrac{1}{3} & \dots & \dfrac{1}{n}\\ \dfrac{1}{2} & \dfrac{1}{3} & \dfrac{1}{4} & \dots & \dfrac{1}{n+1}\\ \dfrac{1}{3} & \dfrac{1}{4} & \dfrac{1}{5} & \dots & \dfrac{1}{n+2}\\ \dots & \dots & \dots & \dots & \dots\\ \dfrac{1}{n} & \dfrac{1}{n+1} & \dfrac{1}{n+2} & \dots & \dfrac{1}{2n-1}\end{vmatrix}=\frac{\Bigl(1!\,2!\,3!\,\dots\,(n-1)!\Bigr)^3}{n!\,(n+1)!\,(n+2)!\,\dots\,(2n-1)!}.$$

Теорема 7.24 позволяет дать следующее *индуктивное определение определителя* (2-я точка зрения на определители). Определителем матрицы 1-го порядка называется (единственный) элемент этой матрицы. Определителем матрицы $A = (a_{ij})$ n-го порядка, $n > 2$, называется

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} \cdot a_{1j} \cdot M_{1j},$$

где M_{ij} — определитель матрицы, полученной из A вычеркиванием i-го строки и j-го столбца. В силу теоремы 7.24 это определение эквивалентно комбинаторному определению.

Теорема Лапласа. Обобщением теоремы 7.24 является теорема Лапласа, позволяющая раскладывать определитель по нескольким строкам (столбцам).

Теорема 7.33 (Лаплас)**.** *Сумма произведений миноров, стоящих в фиксированных k строках матрицы A, на их алгебраические дополнения равна определителю матрицы A:*

$$\det A = \sum_{j_1 < \ldots < j_k} (-1)^{j_1+\ldots+j_k+i_1+\ldots+i_k} \cdot A\begin{pmatrix} i_1, \ldots, i_k \\ j_1, \ldots, j_k \end{pmatrix} \cdot A\begin{pmatrix} i_{k+1}, \ldots, i_n \\ j_{k+1}, \ldots, j_n \end{pmatrix}. \quad (7.21)$$

Аналогичное утверждение справедливо и для любого набора выбранных столбцов.

Доказательство. По лемме 7.23 произведение минора на алгебраическое дополнение является суммой части членов, входящих в разложение $\det A$, взятых с теми же знаками, с которыми они встречаются в этом разложении. Количество этих членов равно произведению числа членов в миноре на число членов в алгебраическом дополнении: $k!(n-k)!$. Общее число миноров, стоящих в выбранных строках равно $\binom{n}{k}$, поэтому общее число членов, встречающихся в произведениях этих миноров на их алгебраические дополнения, равно

$$\binom{n}{k} k!(n-k)! = n!$$

Легко видеть, что все члены попарно различны и поэтому их сумма равна $\det A$. $\square$

Пример 7.34. Разложим определитель 4-го порядка по 1-й и 4-й строкам, применим теорему 7.33 при $i_1 = 1$, $i_2 = 4$:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = (-1)^{1+4+1+2} \cdot \begin{vmatrix} a_{11} & a_{12} \\ a_{41} & a_{42} \end{vmatrix} \cdot \begin{vmatrix} a_{23} & a_{24} \\ a_{33} & a_{34} \end{vmatrix} +$$

$$+(-1)^{1+4+1+3} \begin{vmatrix} a_{11} & a_{13} \\ a_{41} & a_{43} \end{vmatrix} \cdot \begin{vmatrix} a_{22} & a_{24} \\ a_{32} & a_{34} \end{vmatrix} + (-1)^{1+4+1+4} \begin{vmatrix} a_{11} & a_{14} \\ a_{41} & a_{44} \end{vmatrix} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} +$$

$$+(-1)^{1+4+2+3}\begin{vmatrix} a_{12} & a_{13} \\ a_{42} & a_{43} \end{vmatrix}\cdot\begin{vmatrix} a_{21} & a_{24} \\ a_{31} & a_{34} \end{vmatrix} + (-1)^{1+4+2+4}\begin{vmatrix} a_{12} & a_{14} \\ a_{42} & a_{44} \end{vmatrix}\cdot\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} +$$

$$+(-1)^{1+4+3+4}\begin{vmatrix} a_{13} & a_{14} \\ a_{43} & a_{44} \end{vmatrix}\cdot\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Пример 7.35. Вычислим определитель

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

Разложим его по 2-й и 4-й строкам с большим количеством нулевых элементов. В этих строках содержится $\binom{5}{2} = 10$ миноров 2-го порядка, однако 7 из них будут содержать по крайней мере один нулевой столбец, поэтому заведомо равны нулю, поэтому в сумме (7.21) останется только 3 члена:

$$\Delta = (-1)^{2+4+1+3}\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}\cdot\begin{vmatrix} 1 & 7 & 2 \\ 2 & 3 & -4 \\ -1 & 2 & -1 \end{vmatrix} + (-1)^{2+4+1+4}\begin{vmatrix} 1 & -2 \\ 2 & 9 \end{vmatrix}\cdot\begin{vmatrix} 1 & 2 & 2 \\ 2 & 4 & -4 \\ -1 & -2 & -1 \end{vmatrix} +$$

$$+(-1)^{2+4+3+4}\begin{vmatrix} 2 & -2 \\ 3 & 9 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & 2 \\ 2 & 2 & -4 \\ 1 & -1 & -1 \end{vmatrix} = (-1)\cdot 61 - 13\cdot 0 - 24\cdot(-28) = 611.$$

Пример 7.36. Вычислим определитель

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

Из четвертого столбца вычтем первый; из пятого столбца вычтем удвоенный первый:

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

Из пятой строки вычтем четвертую:

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

Разложим определитель по последним двум столбцам:

$$\Delta = (-1)^{3+4+4+5}\cdot\begin{vmatrix} 5 & 1 \\ -2 & 0 \end{vmatrix}\cdot\begin{vmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 2 & 1 & -5 \end{vmatrix} = -2\cdot 5 = -10.$$

Пример 7.37. Пусть $A \in F^{n\times n}$, $B \in F^{m\times m}$, $C \in F^{n\times m}$, тогда

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det A \cdot \det B.$$

Для доказательства достаточно разложить определитель по первым n столбцам.

7.6. Полилинейные знакопеременные функции

Рассмотрим функцию $f(a_1, \ldots, a_n)$ от n аргументов $a_1, \ldots, a_n$, где $a_j \in F^n$ $(j = 1, \ldots, n)$. Такая функция называется *линейной по каждому аргументу* (или *полилинейной*), если

$$f(a_1, \ldots, a'_j + a''_j, \ldots, a_n) = f(a_1, \ldots, a'_j, \ldots, a_n) + f(a_1, \ldots, a''_j, \ldots, a_n),$$
$$f(a_1, \ldots, \alpha a_j, \ldots, a_n) = \alpha f(a_1, \ldots, a_j, \ldots, a_n)$$

для любых входящих сюда векторов, любых чисел α_1, α_2 из поля F и любого $j \in \{1, \ldots, n\}$. Функция f называется *знакопеременной*, если

$$f(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n) = -f(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n)$$

для любых векторов и любых $i, j \in \{1, \ldots, n\}$, $i \neq j$. Функция f называется *нормированной*, если

$$f(e_1, \ldots, e_n) = 1,$$

где e_j — столбец, у которого j-ая компонента равна 1, а остальные равны 0.

Из доказанных результатов следует, что определитель есть полилинейная знакопеременная нормированная функция от своих столбцов. Верно и обратное утверждение:

Теорема 7.38. *Любая полилинейная знакоперменная нормированная функция $f(a_1, \ldots, a_n)$ есть определитель матрицы, составленной из столбцов $a_1, \ldots, a_n$. Аналогичное утверждение справедливо и для строк матрицы.*

Доказательство. По индукции для полилинейной функции мы можем доказать

$$f\left(a_1, \ldots, a_{i-1}, \sum_{i=1}^{m} \alpha_i b_i, a_{i+1}, \ldots, a_n\right) = \sum_{i=1}^{m} \alpha_i f(a_1, \ldots, a_{i-1}, b_i, a_{i+1}, \ldots, a_n). \tag{7.22}$$

Далее, пусть $i_1, \ldots, i_n$ — некоторая перестановка чисел $1, \ldots, n$. Тогда, легко видеть, что

$$f(e_{i_1}, \ldots, e_{i_n}) = (-1)^{\sigma(i_1, \ldots, i_n)}. \tag{7.23}$$

Докажем теперь, что если $a_i = a_j$, где $i \neq j$, то знакопеременная функция $f(a_1, \ldots, a_n)$ равна нулю. Действительно, при транспозиции i-го и j-го аргумента функция не меняется, но по свойству знакопеременности меняет знак, следовательно,

$$f(a_1, \ldots, a_i, \ldots, a_i, \ldots, a_n) = 0. \tag{7.24}$$

Пусть $a_j = (a_{1j}, \ldots, a_{nj})^\top$ $(j = 1, \ldots, n)$, тогда

$$a_j = \sum_{i=1}^{n} a_{ij} e_i. \tag{7.25}$$

Из (7.22)–(7.25) получаем

$$f(a_1, \ldots, a_n) = f\left(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \ldots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\right) =$$
$$= \sum_{i_1=1}^{n} \ldots \sum_{i_n=1}^{n} a_{i_1 1}, \ldots, a_{i_n n} f(e_{i_1}, \ldots, e_{i_n}) = \sum_{i_1, \ldots, i_n} (-1)^{\sigma(i_1, \ldots, i_n)} a_{i_1 1}, \ldots, a_{i_n n} = \det A.$$

Теорема доказана. □

Мы получаем так называемое *аксиоматическое определение определителя*: определителем матрицы называется полилинейная нормированная знакопеременная функция от столбцов (строк) матрицы (*третья точка зрения на определитель*).

7.7. Минорный ранг матрицы

Назовем *базисным минором* матрицы ненулевой минор наибольшего порядка. Базисный минор матрицы в общем случае определяется неоднозначно, но однозначно определяется его порядок, называемый *минорным рангом* матрицы. В этом разделе мы увидим, что определение минорного ранга эквивалентно другим определениями ранга матрицы и, таким образом,

$$\text{столбцовый ранг} = \text{строчечный ранг} = \text{минорный ранг}.$$

Пусть $i \notin \{i_1, \ldots, i_r\}$, $j \notin \{j_1, \ldots, j_r\}$. Минор $A\begin{pmatrix} i_1, \ldots, i_r, i \\ j_1, \ldots, j_r, j \end{pmatrix}$ называется *окаймляющим* по отношению к $A\begin{pmatrix} i_1, \ldots, i_r \\ j_1, \ldots, j_r \end{pmatrix}$.

Теорема 7.39 (об окаймляющих минорах)**.** *Пусть некоторый минор матрицы не равен нулю, а все окаймляющие его миноры равны нулю, тогда строки и столбцы, на пересечении которых стоит исходный минор, образуют соответственно строчечную и столбцовую базы матрицы.*

Доказательство. Для простоты обозначений будем считать, что рассматривается ненулевой минор, стоящий в первых r строках и r столбцах. Минор $A\begin{pmatrix} 1, 2, \ldots, r, i \\ 1, 2, \ldots, r, j \end{pmatrix}$ разло́жим по последней строке:

$$0 = \begin{vmatrix} a_{11} & \ldots & a_{1r} & a_{1j} \\ \ldots & \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} & a_{rj} \\ a_{i1} & \ldots & a_{ir} & a_{ij} \end{vmatrix} = a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{ir}c_r + a_{ij}c, \tag{7.26}$$

где c_j — соответствующий дополнительный минор ($j = 1, 2, \ldots, r$), и

$$c = A\begin{pmatrix} 1, \ldots, r \\ 1, \ldots, r \end{pmatrix} \neq 0.$$

Из (7.26) получаем

$$a_{ij} = -\frac{1}{c}(c_1 a_{i1} + c_2 a_{i2} + \ldots + c_r a_{ir}). \tag{7.27}$$

Заметим, что величины c_j не зависят от i, поэтому из (7.27) следует

$$a_j = -\frac{1}{c}(c_1 a_1 + c_2 a_2 + \ldots + c_r a_r),$$

где a_j — j-й столбец матрицы A. Итак, произвольный столбец матрицы A линейно выражается через линейно независимую систему столбцов с номерами $1, \ldots, r$, поэтому эта система является столбцовой базой. Утверждение про строки матрицы доказывается аналогично. □

Следствие 7.40. *Строки и столбцы, на пересечении которых стоит базисный минор, образуют соответственно строчечную и столбцовую базы матрицы.*

Следствие 7.41. *Если в матрице некоторый минор не равен нулю, а все окаймляющие его миноры равны нулю, то исходный минор — базисный.*

Следствие 7.42. *Минорный ранг равен столбцовому и строчечному рангам матрицы.*

Теорема 7.43. *Минор, стоящий на пересечении строк строчечной базы и столбцов столбцовой базы матрицы, является базисным.*

Доказательство. Пусть $i_1, \ldots, i_r$, $j_1, \ldots, j_r$ — номера строк и столбцов матрицы A соответственно. Обозначим B подматрицу, образованную элементами, стоящими на пересечении базисных строк и столбцов. Обозначим C подматрицу, образованную всеми строками матрицы A и столбцами $j_1, \ldots, j_r$. По условию столбцы матрицы C линейно независимы, поэтому $\operatorname{rank} C = r$. По условию каждая строка матрицы A (и, следовательно, C) линейно выражается через строки $i_1, \ldots, i_r$. Так как $\operatorname{rank} C = r$, то строки $i_1, \ldots, i_r$ линейно независимы, откуда $\operatorname{rank} B = r$, т. е. $\det B \neq 0$. Так как очевидно, что все окаймляющие миноры равны нулю, то исходный минор — базисный. □

7.8. Обратная матрица

Матрица $B \in F^{n\times n}$ называется *обратной* к матрице $A \in F^{n\times n}$, если $AB = BA = E$, где E — единичная матрица. Обратная матрица обозначается A^{-1}.

Обозначим A_{ij} алгебраическое дополнение к элементу a_{ij} матрицы A. Матрица

$$\operatorname{Adj} A = \begin{pmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{pmatrix},$$

называется *присоединенной*, или *взаимной*, к матрице A.

Утверждение 7.44.

$$A \cdot \operatorname{Adj} A = \operatorname{Adj} A \cdot A = \det A \cdot E.$$

Доказательство. Покажем, что $A \cdot \operatorname{Adj} A = \det A \cdot E$. Для этого рассмотрим элемент матрицы $A \cdot \operatorname{Adj} A$, стоящий в i-ой строке и j-м столбце, и покажем, что он равен $\det A$ при $i = j$ и 0 в остальных случаях, т. е. покажем, что

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \begin{cases} \det A, & \text{при } i = j, \\ 0, & \text{при } i \neq j. \end{cases} \tag{7.28}$$

Действительно, при $i = j$ левая часть равенства (7.28) есть разложение определителя $\det A$ по i-й строке. При $i \neq j$ левая часть этого равенства есть разложение определителя матрицы, получающейся из A заменой j-й строки на i-ю — очевидно, этот определитель равен нулю.

Равенство $\operatorname{Adj} A \cdot A = \det A \cdot E$ доказывается аналогично. □

Теорема 7.45. *Если матрица $A \in F^{n\times n}$ невырождена, т. е. $\det A \neq 0$, то A^{-1} существует, единственна и равна*

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{Adj} A. \tag{7.29}$$

Если $\det A = 0$, то A^{-1} не существует.

Доказательство. Из утверждения 7.44 следует, что если $\det A \neq 0$, то матрица A^{-1} существует и определяется формулой (7.29). Докажем единственность обратной матрицы. Пусть B_1, B_2 — матрицы, обратные к A. Тогда $B_1AB_2 = B_1$, так как $AB_2 = E$, а с другой стороны, $B_1AB_2 = B_2$, так как $B_1A = E$, поэтому $B_1 = B_2$.

Осталось показать, что если $\det A = 0$, то обратной матрицы не существует. Предположим, что это не так и B — обратная к A матрица, тогда $AB = E$. Но по следствию 6.39 $\operatorname{rank} E \leq \operatorname{rank} A$, что невозможно, так как $\operatorname{rank} A < n$, а $\operatorname{rank} E = n$. □

Пример 7.46.

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

Пример 7.47. Вычислим матрицу, обратную к

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

Имеем $\det A = 41$,

$$A_{11} = \begin{vmatrix} 2 & 3 \\ -5 & -3 \end{vmatrix} = 9, \quad A_{12} = -\begin{vmatrix} -1 & 3 \\ 3 & -3 \end{vmatrix} = 6, \quad A_{13} = \begin{vmatrix} -1 & 2 \\ 3 & -5 \end{vmatrix} = -1,$$

$$A_{21} = -\begin{vmatrix} 4 & 1 \\ -5 & -3 \end{vmatrix} = 7, \quad A_{22} = \begin{vmatrix} 2 & 1 \\ 3 & -3 \end{vmatrix} = -9, \quad A_{23} = -\begin{vmatrix} 2 & 4 \\ 3 & -5 \end{vmatrix} = 22,$$

$$A_{31} = \begin{vmatrix} 4 & 1 \\ 2 & 3 \end{vmatrix} = 10, \quad A_{32} = -\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} = -7, \quad A_{33} = \begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} = 8,$$

Поэтому

$$A^{-1} = \frac{1}{41} \cdot \begin{pmatrix} 9 & 7 & 10 \\ 6 & -9 & -7 \\ -1 & 22 & 8 \end{pmatrix}.$$

Следствие 7.48. *Если $A \in F^{n\times n}$, $B \in F^{n\times n}$, $\det A \neq 0$ и $AB = E$, то $BA = E$, т. е. $A^{-1} = B$. Аналогично, если $BA = E$, то $AB = E$, т. е. $A^{-1} = B$.*

Доказательство. Домножая обе части равенства $AB = E$ слева на A^{-1}, получаем $A^{-1}AB = A^{-1}$, т. е. $B = A^{-1}$. Вторая часть теоремы доказывается аналогично. □

Утверждение 7.49. *Пусть* $A \in F^{n\times n}$, $B \in F^{n\times n}$, $\det A \neq 0$, $\det B \neq 0$, *тогда*

$$(A^{-1})^{-1} = A, \quad (AB)^{-1} = B^{-1}A^{-1}, \quad (A^\top)^{-1} = (A^{-1})^\top.$$

Доказательство. Все три свойства выводятся из определения обратной матрицы. Докажем, например, что $(A^\top)^{-1} = (A^{-1})^\top$. Это равенство означает, что матрица $(A^{-1})^\top$ является обратной к $A^\top$. Проверим это: $(A^{-1})^\top A^\top = (AA^{-1})^\top = E$. □

Замечание 7.50. Для практического нахождения обратной матрицы формула (7.29) пригодна лишь при небольших значениях порядка матрицы, либо для матриц специального вида. Как правило, для нахождения обратной матрицы используют метод элементарных преобразований. Пусть $\det A \neq 0$, тогда матрицу A элементарными преобразованиями строк можно привести к единичной матрице E. Оказывается, что если над единичной матрицей проделать те же элементарные преобразования, что и над матрицей A, то на месте единичной матрицы получим A^{-1}. Действительно, по утверждению 6.42 каждое элементарное преобразование строк матрицы A эквивалентно домножению слева матрицы A на некоторую матрицу, поэтому последовательность таких преобразований эквивалентна домножению слева на матрицы $S_1, S_2, \ldots, S_t$. Если эта последовательность элементарных преобразований приводит матрицу к единичной, то

$$S_t \cdot (S_{t-1} \cdot \ldots \cdot (S_2 \cdot (S_1 \cdot A)) \ldots) = E. \tag{7.30}$$

Обозначим $B = S_t S_{t-1} \ldots S_2 S_1$, тогда из (7.30) получаем $B = A^{-1}$. Поэтому проделав те же элементарные преобразования над единичной, получим матрицу $S_t \cdot (S_{t-1} \cdot \ldots \cdot (S_2 \cdot (S_1 \cdot E)) \ldots) = B = A^{-1}$.

Как правило, для реализации этого алгоритма записывают блочную матрицу (A, E). С помощью элементарных преобразований строк этой матрицы на месте A получают единичную матрицу, тогда на месте E получим A^{-1}.

Пример 7.51. Найдем обратную матрицу к матрице из примера 7.47. Припишем справа к A единичную матрицу:

$$(A, E) = \left(\begin{array}{ccc|ccc} 2 & 4 & 1 & 1 & 0 & 0 \\ -1 & 2 & 3 & 0 & 1 & 0 \\ 3 & -5 & -3 & 0 & 0 & 1 \end{array}\right).$$

Будем проводить преобразования со строками блочной матрицы (A, E). Переставим вторую строку на первое место и вычтем ее из оставшихся с подходящими множителями, чтобы занулить внедиагональные элементы первого столбца:

$$\left(\begin{array}{ccc|ccc} 1 & -2 & -3 & 0 & -1 & 0 \\ 0 & 8 & 7 & 1 & 2 & 0 \\ 0 & 1 & 6 & 0 & 3 & 1 \end{array}\right).$$

Переставим вторую и третью строки и вычтем ее из остальных строк с подходящими множителями так, чтобы занулить внедиагональные элементы второго столбца:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 9 & 0 & 5 & 2 \\ 0 & 1 & 6 & 0 & 3 & 1 \\ 0 & 0 & -41 & 1 & -22 & -8 \end{array}\right).$$

Умножим последнюю строку на $-\dfrac{1}{41}$ и вычтем ее из остальных строк с подходящими множителями, чтобы занулить внедиагональные элементы третьего столбца:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \dfrac{9}{41} & \dfrac{7}{41} & \dfrac{10}{41} \\ 0 & 1 & 0 & \dfrac{6}{41} & -\dfrac{9}{41} & -\dfrac{7}{41} \\ 0 & 0 & 1 & -\dfrac{1}{41} & \dfrac{22}{41} & \dfrac{8}{41} \end{array}\right).$$

Справа от черты стоит A^{-1}.

Утверждение 7.52. *Пусть ε — первообразный корень n-й степени из* 1, *тогда*

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \varepsilon & \varepsilon^2 & \dots & \varepsilon^{n-1} \\ 1 & \varepsilon^2 & \varepsilon^4 & \dots & \varepsilon^{2(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \varepsilon^{n-1} & \varepsilon^{(n-1)2} & \dots & \varepsilon^{(n-1)^2} \end{pmatrix}^{-1} = \frac{1}{n}\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \varepsilon^{-1} & \varepsilon^{-2} & \dots & \varepsilon^{-(n-1)} \\ 1 & \varepsilon^{-2} & \varepsilon^{-4} & \dots & \varepsilon^{-2(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \varepsilon^{-(n-1)} & \varepsilon^{-(n-1)2} & \dots & \varepsilon^{-(n-1)^2} \end{pmatrix}.$$

Доказательство. Рассмотрим элемент, стоящий в i-й строке и j-м столбце произведения $AB = (c_{ik})$, где $A = (a_{ik})$, $B = (b_{kj})$, $a_{ik} = \varepsilon^{(i-1)(k-1)}$, $b_{kj} = \dfrac{1}{n}\varepsilon^{-(k-1)(j-1)}$. исходная матрица, B — матрица, стоящая в правой части доказываемого равенства:

$$c_{ik} = \frac{1}{n}\sum_{k=1}^{n}\varepsilon^{(i-1)(k-1)}\varepsilon^{-(k-1)(j-1)} = \frac{1}{n}\sum_{k=0}^{n-1}\varepsilon^{k(i-j)} = \begin{cases} 1 & \text{при } i = k, \\ \dfrac{1-\varepsilon^{n(i-j)}}{1-\varepsilon^{i-j}} = 0 & \text{при } i \neq k. \end{cases}$$

Таким образом, $AB = E$, откуда $B = A^{-1}$. □

Утверждение 7.53 (Формула Фробениуса)**.** *Пусть $A \in F^{n\times n}$, $B \in F^{n\times m}$, $C \in F^{m\times n}$, $D \in F^{m\times m}$. Если матрицы A и $T = D - CA^{-1}B$ невырождены, то $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ тоже невырождена и*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BT^{-1}CA^{-1} & -A^{-1}BT^{-1} \\ -T^{-1}CA^{-1} & T^{-1} \end{pmatrix}.$$

Доказательство. Проверяется умножением. □

Формулы Фробениуса позволяют свести вычисление обратной матрицы порядка $n+m$ к обращению одной матрицы порядка n и одной матрицы порядка m.

Утверждение 7.54 (Формула Шермана–Моррисона)**.** *Пусть $A \in F^{n\times n}$, $U \in F^{n\times m}$, $V \in F^{m\times n}$,* $\det A \neq 0$. *Если матрица $E_m + VA^{-1}U$ невырождена, то матрица $A+UV$ также невырождена и*

$$(A+UV)^{-1} = A^{-1} - A^{-1}U(E_m + VA^{-1}U)^{-1}VA^{-1}.$$

Доказательство. Перемножим матрицы:

$$(A+UV)\Big(A^{-1} - A^{-1}U(E_m + VA^{-1}U)^{-1}VA^{-1}\Big) =$$
$$= E_n - U(E_m + VA^{-1}U)^{-1}VA^{-1} + UVA^{-1} - UVA^{-1}U(E_m + VA^{-1}U)^{-1}VA^{-1} =$$
$$= E_n - U\Big(E_m - (E_m + VA^{-1}U) - VA^{-1}U\Big)(E_m + VA^{-1}U)^{-1}VA^{-1} = E_n,$$

что доказывает требуемое. □

Пример 7.55. Найдем матрицу, обратную к

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

Матрица B отличается от матрицы A из примера 7.47 лишь элементом $b_{12} = 3$, в то время как $a_{12} = 4$. Для матрицы A уже была вычислена обратная:

$$A^{-1} = \frac{1}{41} \cdot \begin{pmatrix} 9 & 7 & 10 \\ 6 & -9 & -7 \\ -1 & 22 & 8 \end{pmatrix},$$

поэтому для вычисления B^{-1} воспользуемся утверждением 7.54. Пусть

$$U = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad V = (0, -1, 0),$$

тогда $B = A + UV$. Последовательно вычисляем:

$$A^{-1}U = \frac{1}{41} \cdot \begin{pmatrix} 9 \\ 6 \\ -1 \end{pmatrix}, \qquad VA^{-1} = \frac{1}{41} \cdot (-6, 9, 7), \qquad VA^{-1}U = -\frac{6}{41},$$

откуда

$$B^{-1} = A^{-1} - (A^{-1}U)(1 + VA^{-1}U)^{-1}(VA^{-1}) = \frac{1}{35} \cdot \begin{pmatrix} 9 & 4 & 7 \\ 6 & -9 & -7 \\ -1 & 19 & 7 \end{pmatrix}.$$

Упражнение 7.56. Вычислить

$$\begin{pmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ b & b & a & \dots & b \\ \dots & \dots & \dots & \dots & \dots \\ b & b & a & \dots & a \end{pmatrix}^{-1}$$

при $a \neq b$ и $a \neq b(1-n)$.

Утверждение 7.57. *Пусть* $A \in F^{n\times n}$, $B \in F^{n\times m}$, $C \in F^{m\times n}$, $\det A \neq 0$. *Тогда каждое из матричных уравнений*

$$AX = B, \qquad YA = C \tag{7.31}$$

имеет единственное решение:

$$X = A^{-1}B, \qquad Y = CA^{-1}. \tag{7.32}$$

Доказательство. Непосредственной подстановкой правых частей равенств (7.32) в (7.31) убеждаемся, что формулы (7.32) задают решения рассматриваемых уравнений. Для доказательства единственности решения уравнения $AX = B$ предположим, что оно имеет два решения X_1, X_2, тогда $AX_1 = AX_2$, откуда, домножая обе части этого равенства слева на A^{-1}, получаем $X_1 = X_2$. Единственность решения уравнения $YA = C$ доказывается аналогично. □

Замечание 7.58. Как правило при реальных вычислениях решение $X = A^{-1}B$ матричного уравнения $AX = B$ удобнее находить с помощью элементарных преобразований строк блочной матрицы (A, B), приводящих матрицу A к единичному виду. Легко доказать, что при на месте матрицы B мы получим $A^{-1}B$ (ср. замечание 7.50). Также для решения матричного уравнения $YA = C$ удобно рассмотреть блочную матрицу $\begin{pmatrix} A \\ C \end{pmatrix}$ и элементарными преобразованиями ее столбцов привести верхнюю часть к единичному виду, тогда на месте матрицы C мы получим CA^{-1}.

Пример 7.59. Решим матричное уравнение

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

Записываем блочную матрицу

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

Из второй строки вычитаем первую, умноженную на 3:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & -1 \\ 0 & -5 & -6 & 6 \end{array}\right).$$

Вторую строку умножим на $-\dfrac{1}{5}$ и вычтем ее из первой с множителем 2:

$$\left(\begin{array}{cc|cc} 1 & 0 & -\frac{7}{5} & \frac{7}{5} \\ 0 & 1 & \frac{6}{5} & -\frac{6}{5} \end{array}\right),$$

откуда

$$X = \frac{1}{5} \cdot \begin{pmatrix} -7 & 7 \\ 6 & -6 \end{pmatrix}.$$

Пример 7.60. Решим матричное уравнение

$$Y \cdot \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}.$$

Запишем блочную матрицу

$$\left(\begin{array}{cc} 1 & 2 \\ 3 & 1 \\ \hline 1 & -1 \\ -3 & 3 \end{array}\right).$$

Из второго столбца вычтем первый, умноженный на 2:

$$\left(\begin{array}{cc} 1 & 0 \\ 3 & -5 \\ \hline 1 & -3 \\ -3 & 9 \end{array}\right).$$

Второй столбец умножим на $-\frac{1}{5}$ и вычтем его из первого с множителем 3:

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline -\frac{4}{5} & \frac{3}{5} \\ \frac{12}{5} & -\frac{9}{5} \end{array}\right),$$

откуда

$$Y = \frac{1}{5} \cdot \begin{pmatrix} -4 & 3 \\ 12 & -9 \end{pmatrix}.$$

7.9. Формулы Крамера

Система линейных уравнений $Ax = b$, в которой число уравнений равно числу неизвестных, т. е. $A \in F^{n\times n}$, называется *квадратной*. В этом случае $\det A$ называется *определителем* этой системы.

Утверждение 7.61. *Квадратная система имеет единственное решение тогда и только тогда, когда матрица системы невырождена, т. е. ее определитель не равен нулю.*

Доказательство. Следует из теорем 6.20, 6.24. □

Квадратная система с ненулевым определителем называется *крамеровской*. В разделе 7.1 были найдены формулы (7.7), (7.11), (7.12), используя которые можно найти (единственное) решение крамеровской системы второго и третьего порядка. Эти формулы могут быть обобщены на случай систем произвольного порядка.

Теорема 7.62 (Формулы Крамера). *Пусть $A \in F^{n\times n}$ и $\det A \neq 0$, тогда система $Ax = b$ имеет единственное решение $x \in F^n$, компоненты которого можно найти по формулам*

$$x_j = \frac{\Delta}{\Delta_j} \quad (j = 1, 2, \ldots, n), \tag{7.33}$$

где Δ — определитель матрицы A, а Δ_j — определитель матрицы, получающийся из A заменой ее j-го столбца столбцом b.

Доказательство. Так как $\det A = n$, то $\operatorname{rank} A = n$, откуда $\operatorname{rank}(A, b) = n$, следовательно система $Ax = b$ совместна и имеет единственное решение. Обозначим $a_1, a_2, \ldots, a_n$ — столбцы матрицы A, тогда $b = \sum\limits_{k=1}^{n} x_k a_k$ и

$$\Delta_j = \det(a_1, \ldots, b, \ldots, a_n) = \det\left(a_1, \ldots, \sum_{k=1}^{n} x_k a_k, \ldots, a_n\right) =$$

$$= \sum_{k=1}^{n} x_k \det(a_1, \ldots, a_k, \ldots, a_n) = x_j \det(a_1, \ldots, a_j, \ldots, a_n) = x_j \Delta,$$

откуда получаем формулу (7.33). □

Замечание 7.63. По утверждению 7.57 решение крамеровской системы $Ax = b$ дает формула $x = A^{-1}b$. Формулы Крамера можно получить теперь из этой формулы и формул для определения коэффициентов обратной матрицы через алгебраические дополнения.

Рассмотрим однородную линейную систему, в которой число уравнений на 1 меньше числа неизвестных:

$$\begin{cases} a_{11}x_1 \;+\; a_{12}x_2 \;+\ldots\; a_{1n}x_n \;= 0; \\ a_{21}x_1 \;+\; a_{22}x_2 \;+\ldots\; a_{2n}x_n \;= 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n-1,1}x_1 + a_{n-1,2}x_2 + \ldots\; a_{n-1,n}x_n = 0. \end{cases} \tag{7.34}$$

Утверждение 7.64. *Частное решение $x = (x_1, x_2, \ldots, x_n)^\top$ системы (7.34) можно вычислить по формуле*

$$x_j = (-1)^{j+1} \begin{vmatrix} a_{11} & \ldots & a_{1,j-1} & a_{1,j+1} & \ldots & a_{1,n} \\ a_{21} & \ldots & a_{2,j-1} & a_{2,j+1} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & \ldots & a_{n-1,j-1} & a_{n-1,,j+1} & \ldots & a_{n-1,n} \end{vmatrix} \qquad (j = 1, 2, \ldots, n).$$

Если ранг системы равен $n-1$, то $x \neq 0$, и любое решение получается умножением найденного x на некоторое (произвольное) число.

Доказательство. Для доказательства рассмотрим определитель

$$\begin{vmatrix} a_{i1} & a_{i2} & \ldots & a_{i,n} \\ a_{11} & a_{12} & \ldots & a_{1,n} \\ a_{21} & a_{22} & \ldots & a_{2,n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n} \end{vmatrix} = 0 \qquad (i = 1, 2, \ldots, n-1).$$

Разложенив его по первой строке получим тождество, показывающее, что вычисленный по доказываемым формулам набор $x = (x_1, x_2, \ldots, x_n)$ удовлетворяет i-му уравнению.

Если ранг системы равен $n-1$, то по крайней мере один из миноров $n-1$-го порядка не равен нулю, поэтому $x \neq 0$. Размерность пространства решений равна $n - (n-1) = 1$, поэтому x образует фундаментальную систему решений. □

7.10. Сумма определителей

Из свойства линейности определителя получаем.

Теорема 7.65 (о сумме определителей)**.** *Определитель суммы двух матриц A и B порядка n равен сумме всех 2^n определителей матриц, получающихся из A заменой части столбцов A соответствующими столбцами (т. е. столбцами с теми же номерами) из B (включая сами матрицы A и B). Аналогичное утверждение справедливо и для строк матриц.*

Пример 7.66. Для матриц второго порядка:

$$\det(A+B) = \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12}+b_{12} \\ a_{21} & a_{22}+b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & a_{12}+b_{12} \\ b_{21} & a_{22}+b_{22} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & b_{12} \\ a_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & a_{12} \\ b_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

Пример 7.67. Вычислим определитель

$$\Delta = \begin{vmatrix} a_1b_1+x_1 & a_1b_2 & a_1b_3 & \dots & a_1b_n \\ a_2b_1 & a_2b_2+x_2 & a_2b_3 & \dots & a_2b_n \\ a_3b_1 & a_3b_2 & a_3b_3+x_3 & \dots & a_3b_n \\ \dots & \dots & \dots & \dots & \dots \\ a_nb_1 & a_nb_2 & a_nb_3 & \dots & a_nb_n+x_n \end{vmatrix}.$$

Заметим, что Δ есть определитель суммы двух матриц:

$$A = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x_n \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \dots & \dots & \dots & \dots \\ a_nb_1 & a_nb_2 & \dots & a_nb_n \end{pmatrix}$$

Применим к Δ теорему 7.65. Определитель матрицы, получающейся из A заменой более 1 столбца столбцами матрицы B, равен нулю, так как содержит по крайней мере два пропорциональных столбца. Поэтому Δ равен определителю матрицы A плюс сумма n определителей вида

$$\begin{vmatrix} x_1 & 0 & 0 & a_1b_j & 0 & 0 \\ 0 & x_2 & 0 & a_2b_j & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & x_{j-1} & a_{j-1}b_j & 0 & 0 \\ 0 & 0 & 0 & a_jb_j & 0 & 0 \\ 0 & 0 & 0 & a_{j+1}b_j & x_{j+1} & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_nb_j & 0 & x_n \end{vmatrix} = x_1x_2\dots x_{j-1}a_jb_jx_{j+1}\dots x_n$$

(раскрыли по первым $j-1$ столбцам). Таким образом,

$$\Delta = x_1x_2\dots x_n + \sum_{j=1}^{n} x_1x_2\dots x_{j-1}a_jb_jx_{j+1}\dots x_n.$$

7.11. Теорема Бине–Коши

Теорема 7.68 (Бине–Коши). *Пусть* $A = (a_{ij}) \in F^{m\times n}$, $B = (b_{ij}) \in F^{n\times m}$. *Тогда*

$$\det(AB) = \sum_{1\le k_1<k_2<\dots<k_m\le n} A\begin{pmatrix} k_1, k_2, \dots, k_m \\ 1, 2, \dots, m \end{pmatrix} \cdot B\begin{pmatrix} 1, 2, \dots, m \\ k_1, k_2, \dots, k_m \end{pmatrix}.$$

В частности, если $m > n$, *то* $\det(AB) = 0$. *Если* $m = n$, *то*

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

(*последнее равенство составляет содержание* теоремы об умножении определителей).

Доказательство. Пусть $A = (a_1, a_2, \dots, a_n)$, $C = A \cdot B = (c_{ij})$, где $a_1, a_2, \dots, a_n$ — столбцы матрицы A, $c_{ij} = \sum\limits_{k=1}^{n} a_{ik} b_{kj}$ $(i, j = 1, \dots, m)$. Тогда

$$C = \left(\sum_{i_1=1}^{n} b_{i_1 1} a_{i_1}, \sum_{i_2=1}^{n} b_{i_2 2} a_{i_2}, \dots, \sum_{i_m=1}^{n} b_{i_m m} a_{i_m} \right).$$

Пользуясь свойством линейности определителя, получаем

$$\det C = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \dots \sum_{i_m=1}^{n} b_{i_1 1} b_{i_2 1} \dots b_{i_m m} \det(a_1, a_2, \dots, a_m).$$

В правой части равенства имеем сумму n^m определителей; сумма берется по всем наборам $(i_1, \dots, i_m)$, где $1 \le i_j \le n$ $(j = 1, 2, \dots, m)$. Определители с одинаковыми столбцами заведомо равны нулю, поэтому

$$\det C = \sum_{i_1, \dots, i_m} b_{i_1 1} b_{i_2 2} \dots b_{i_m m} \det(a_{i_1}, a_{i_2}, \dots, a_{i_m}).$$

В правой части равенства имеем сумму $n(n-1)\dots(n-m+1)$ определителей по всем размещениям $i_1, \dots, i_m$ элементов $1, \dots, n$. Каждое такое размещение есть перестановка элементов $k_1, \dots, k_m$, где $k_1 \le \dots \le k_m$. Поэтому

$$\det C = \sum_{1 \le k_1 < k_2 < \dots < k_m \le n} \; \sum_{i_1, \dots, i_m} b_{i_1 1} b_{i_2 2} \dots b_{i_m m} \det(a_{i_1}, a_{i_2}, \dots, a_{i_m}).$$

Внешняя сумма берется по всем сочетаниям $k_1, \dots, k_m$ элементов $1, \dots, n$, а внутренняя — по всем перестановкам $i_1, \dots, i_m$ элементов $k_1, \dots, k_m$. Переставим столбцы определителей в порядке возрастания $k_1, \dots, k_m$ их индексов $i_1, \dots, i_m$:

$$\det C = \sum_{1 \le k_1 < k_2 < \dots < k_m \le n} \; \sum_{i_1, \dots, i_m} b_{i_1 1} b_{i_2 2} \dots b_{i_m m} (-1)^{\sigma(i_1, \dots, i_m)} \det(a_{k_1}, a_{k_2}, \dots, a_{k_m}).$$

Так как

$$\sum_{i_1, \dots, i_m} b_{i_1 1} b_{i_2 2} \dots b_{i_m m} (-1)^{\sigma(i_1, \dots, i_m)} = \begin{vmatrix} b_{k_1 1} & \dots & b_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & b_{k_m m} \end{vmatrix},$$

то

$$\det C = \sum_{1 \le k_1 < k_2 < \dots < k_m \le n} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix} \cdot \begin{vmatrix} b_{k_1 1} & \dots & a_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & a_{k_m m} \end{vmatrix}.$$

□

Пример 7.69.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix} =$$

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

Пример 7.70. Пусть a, b, c, x, y, z — векторы геометрического трехмерного пространства. Далее $(\cdot, \cdot)$, $[\cdot, \cdot]$, $(\cdot, \cdot, \cdot)$ обозначают скалярное, векторное и смешанное произведение векторов соответственно. Справедливы равенства:

$$(a, b, c)(x, y, z) = \begin{vmatrix} (a, x) & (a, y) & (a, z) \\ (b, x) & (b, y) & (b, z) \\ (c, x) & (c, y) & (c, z) \end{vmatrix},$$

$$([a, b], [x, y]) = \begin{vmatrix} (a, x) & (a, y) \\ (b, x) & (b, y) \end{vmatrix},$$

$$(a, b, c)[x, y] = \begin{vmatrix} a & b & c \\ (a, x) & (b, x) & (c, x) \\ (a, y) & (b, y) & (c, y) \end{vmatrix}.$$

Следствие 7.71. *Если матрица $A \in F^{n\times n}$ невырождена, то*

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

Доказательство. Из равенства $AA^{-1} = E$ вытекает $\det(AA^{-1}) = 1$, откуда получаем требуемое. □

Пример 7.72. Для вычисления определителя матрицы

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

умножим A на $A^\top$:

$$AA^\top = \begin{pmatrix} a^2+b^2+c^2+d^2 & 0 & 0 & 0 \\ 0 & a^2+b^2+c^2+d^2 & 0 & 0 \\ 0 & 0 & a^2+b^2+c^2+d^2 & 0 \\ 0 & 0 & 0 & a^2+b^2+c^2+d^2 \end{pmatrix},$$

откуда $\det(AA^\top) = (\det A)^2 = (a^2+b^2+c^2+d^2)^4$, поэтому $\det A = (a^2+b^2+c^2+d^2)^2$. Заодно мы показали, что

$$A^{-1} = \frac{1}{a^2+b^2+c^2+d^2} \cdot A^\top.$$

Циркулянтом называется определитель

$$R(b_0, b_1, \ldots, b_{n-1}) = \begin{vmatrix} b_0 & b_1 & b_2 & \ldots & b_{n-2} & b_{n-1} \\ b_{n-1} & b_0 & b_1 & \ldots & b_{n-3} & b_{n-2} \\ b_{n-2} & b_{n-1} & b_0 & \ldots & b_{n-4} & b_{n-3} \\ \multicolumn{6}{c}{\dotfill} \\ b_2 & b_3 & b_4 & \ldots & b_0 & b_1 \\ b_1 & b_2 & b_3 & \ldots & b_{n-1} & b_0 \end{vmatrix}.$$

Утверждение 7.73. *Пусть* ε — *первообразный корень* n*-й степени из* 1 *и пусть* $f(x) = b_0 + b_1x + \ldots + b_{n-1}x^{n-1}$, *тогда*

$$R(b_0, b_1, \ldots, b_{n-1}) = f(1)f(\varepsilon)f(\varepsilon^2)\ldots f(\varepsilon^{n-1}).$$

Доказательство. Для доказательства умножим циркулянт на определитель Вандермонда $W(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1})$. Рассмотрим, например, случай $n = 3$. Пользуясь теоремой об умножении определителей, получаем

$$\begin{vmatrix} b_0 & b_1 & b_2 \\ b_2 & b_0 & b_1 \\ b_1 & b_2 & b_0 \end{vmatrix} \cdot \begin{vmatrix} 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon^2 \\ 1 & \varepsilon^2 & \varepsilon^4 \end{vmatrix} = \begin{vmatrix} f(1) & f(\varepsilon) & f(\varepsilon^2) \\ f(1) & \varepsilon f(\varepsilon) & \varepsilon^2 f(\varepsilon^2) \\ f(1) & \varepsilon^2 f(\varepsilon) & \varepsilon^4 f(\varepsilon^2) \end{vmatrix} =$$

$$= f(1)f(\varepsilon)f(\varepsilon^2) \cdot \begin{vmatrix} 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon^2 \\ 1 & \varepsilon^2 & \varepsilon^4 \end{vmatrix}.$$

Итак, $R(b_0, b_1, \ldots, b_{n-1})W(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}) = f(1)f(\varepsilon)\ldots f(\varepsilon^{n-1})W(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1})$. Так как $W(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}) \neq 0$, то получаем требуемое. В общем случае рассуждения аналогичные. □

Пример 7.74. Рассмотрим циркулянт

$$R = R(\underbrace{1, 1, \ldots, 1}_{k}, \underbrace{0, 0, \ldots, 0}_{n-k}) = \begin{vmatrix} 1 & 1 & 1 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 1 & \ldots & 1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 1 & 1 & 1 & \ldots & 0 & 0 \\ \multicolumn{10}{c}{\dotfill} \\ 1 & 1 & 1 & \ldots & 0 & 0 & 0 & \ldots & 1 & 1 \\ 1 & 1 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 \end{vmatrix}.$$

Согласно утверждению 7.73 имеем $f(x) = 1 + x + \ldots + x^{k-1}$ и

$$R = k(1+\varepsilon+\varepsilon^2+\ldots+\varepsilon^{k-1})(1+\varepsilon^2+\varepsilon^4+\ldots+\varepsilon^{2(k-1)})\ldots(1+\varepsilon^{n-1}+\varepsilon^{(n-1)2}+\ldots+\varepsilon^{(n-1)(k-1)}) =$$

$$= k \cdot \frac{1-\varepsilon^k}{1-\varepsilon} \cdot \frac{1-\varepsilon^{2k}}{1-\varepsilon^2} \cdot \frac{1-\varepsilon^{(n-1)k}}{1-\varepsilon^{n-1}}, \tag{7.35}$$

где ε — первообразный корень n-й степени из единицы. Обозначим $d = \text{НОД}(n, k)$. Имеем $\varepsilon^{kn/d} = 1$. При $d > 1$ отсюда следует, что один из множителей в (7.35) равен нулю, поэтому $R = 0$. Если же $d = 1$, то ε^k является первообразным корнем n-й степени и числа $1, \varepsilon^k, \varepsilon^{2k}, \ldots, \varepsilon^{(n-1)k}$ исчерпывают все значения корня n-й степени из 1, поэтому $R = k$.

7.12. Приложения

Интерполяционный многочлен. Задачу нахождения интерполяционного многочлена можно свести к задаче решения крамеровской системы специального вида. Пусть требуется найти многочлен $f(x)$ степени меньшей n, удовлетворяющий условиям

$$f(x_i) = y_i \qquad (i = 1, 2, \ldots, n), \tag{7.36}$$

где x_i, y_i — заданные числа, причем $x_i \neq x_j$ при $i \neq j$. По теореме 4.39 решение этой задачи существует и единственно. Дадим другое доказательство этой теоремы.

Пусть

$$f(x) = a_0 + a_1x + \ldots + a_{n-1}x^{n-1}.$$

— искомый интерполяционный многочлен, в котором необходимо отыскать значения коэффициентов a_j. Покажем, что эти коэффициенты определяются, причем единственным образом из условий (7.36). Действительно, эти условия приводят к квадратной системе

$$\begin{cases} a_0 + x_1 a_1 + x_1^2 a_2 + \ldots + x_1^{n-1} a_{n-1} = y_1, \\ a_0 + x_2 a_1 + x_2^2 a_2 + \ldots + x_2^{n-1} a_{n-1} = y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + x_n a_1 + x_n^2 a_2 + \ldots + x_n^{n-1} a_{n-1} = y_n, \end{cases} \tag{7.37}$$

рассматриваемой относительно неизвестных $a_0, \ldots, a_{n-1}$. Определитель этой системы — определитель Вандермонда $W(x_1, x_2, \ldots, x_n)$. Так как $x_i \neq x_j$ при $i \neq j$, то по следствию 7.27 он не равен нулю, поэтому система (7.37) имеет единственное решение.

Приведенное доказательство дает еще один способ отыскания интерполяционного многочлена.

Пример 7.75. Найдем многочлен $f(x)$ степени не выше 4, удовлетворяющий условиям

$$f(-2) = 41, \quad f(-1) = 8, \quad f(0) = 3, \quad f(1) = 2, \quad f(2) = 5. \tag{7.38}$$

Пусть $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$, где a_j — коэффициенты, которые требуется найти. Полагая в $f(x)$ $x = -2, -1, 0, 1, 2$, получим систему уравнений относительно неизвестных a_j $(j = 0, 1, \ldots, 4)$:

$$\begin{cases} a_0 - 2a_1 + 4a_2 - 8a_3 + 16a_4 = 41, \\ a_0 - a_1 + a_2 - a_3 + a_4 = 8, \\ a_0 = 3, \\ a_0 + a_1 + a_2 + a_3 + a_4 = 2, \\ a_0 + 2a_1 + 4a_2 + 8a_3 + 16a_4 = 5. \end{cases}$$

Решая эту систему находим $a_0 = 3$, $a_1 = -1$, $a_2 = 1$, $a_3 = -2$, $a_4 = 1$. Таким образом, искомый многочлен — это $f(x) = 3 - x + x^2 - 2x^3 + x^4$.

Интерполяционный многочлен Лагранжа–Эрмита. Рассмотрим одно обобщение задачи об интерполяционном многочлене.

Пример 7.76. Найдем многочлен $f(x)$ степени не выше 4, удовлетворяющий условиям

$$f(1) = 2, \quad f'(1) = 7, \quad f''(1) = 26, \quad f(2) = 35, \quad f'(2) = 75. \tag{7.39}$$

Пусть $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$, где a_j — коэффициенты, которые требуется найти. Имеем $f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3$, $f''(x) = 2a_2 + 6a_3 x + 12a_4 x^2$. Подставляя в $f(x)$, $f'(x)$, $f''(x)$ условия (7.39), приходим к системе уравнений относительно неизвестных a_j $(j = 0, 1, \ldots, 4)$:

$$\begin{cases} a_0 + a_1 + a_2 + a_3 + a_4 = 2, \\ a_1 + 2a_2 + 3a_3 + 4a_4 = 7, \\ 2a_2 + 6a_3 + 12a_4 = 26, \\ a_0 + 2a_1 + 4a_2 + 8a_3 + 16a_4 = 35, \\ a_1 + 4a_2 + 12a_3 + 32a_4 = 75. \end{cases}$$

Система имеет единственное решение $a_0 = 1$, $a_1 = -1$, $a_2 = 1$, $a_3 = -2$, $a_4 = 3$. Найден многочлен $f(x) = 1 - x + x^2 - 2x^3 + 3x^4$.

Рассмотрим таблицу чисел

$$\begin{array}{|c|cccc|} \hline x_1 & y_{10} & y_{11} & \cdots & y_{1,n_1-1} \\ x_2 & y_{20} & y_{21} & \cdots & y_{2,n_2-1} \\ \cdots & \multicolumn{4}{c|}{\ldots\ldots\ldots\ldots\ldots\ldots} \\ x_s & y_{s0} & y_{s1} & \cdots & y_{s,n_s-1} \\ \hline \end{array}\ , \tag{7.40}$$

в которой $x_i \neq x_j$ при $i \neq j$. Многочлен $f(x) \in \mathbb{C}[x]$ степени меньшей $n = n_1 + n_2 + \ldots + n_s$, удовлетворяющий условиям

$$f^{(j)}(x_i) = y_{ij} \quad (i = 1, 2, \ldots, s;\ j = 0, 1, \ldots, n_i - 1) \tag{7.41}$$

назовем *интерполяционным многочленом Лагранжа–Эрмита*, относящимся к интерполяционной таблице (7.40).

Пусть

$$f(x) = a_0 + a_1 x + \ldots + a_{n-2}x^{n-2} + a_{n-1}x^{n-1}$$

— искомый интерполяционный многочлен. Из условий (7.41) получаем квадратную систему линейных уравнений

$$\begin{cases} a_0 + x_1 a_1 + \ x_1^2 a_2 + \ldots + \qquad\qquad\qquad x_1^{n-1}a_{n-1} = y_{10}, \\ \qquad a_1 + 2x_1 a_2 + \ldots + \qquad\quad (n-1)x_1^{n-2}a_{n-1} = y_{11}, \\ \qquad\qquad 2a_2 + \ldots + (n-1)(n-2)x_1^{n-3}a_{n-1} = y_{12}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + x_s a_1 + \ x_s^2 a_2 + \ldots + \qquad\qquad\qquad x_s^{n-1}a_{n-1} = y_{s0}, \\ \qquad a_1 + 2x_s a_2 + \ldots + \qquad\quad (n-1)x_s^{n-2}a_{n-1} = y_{s1}, \\ \qquad\qquad 2a_2 + \ldots + (n-1)(n-2)x_s^{n-3}a_{n-1} = y_{s2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases} \tag{7.42}$$

рассматриваемую относительно неизвестных $a_0, \ldots, a_{n-1}$. Покажем, что решение системы (7.42) существует и единственно, что будет означать, что существует и единственен интерполяционный многочлен Лагранжа–Эрмита.

Лемма 7.77. *Для любой таблицы интерполяции* (7.40), *в которой* $y_{ij} = 0$ ($i = 1, 2, \ldots, s$; $j = 0, 1, \ldots, n_i - 1$), *интерполяционный многочлен Лагранжа–Эрмита существует и единственен.*

Доказательство. Условия $f^{(j)}(x_i) = 0$ ($i = 1, 2, \ldots, s$; $j = 0, 1, \ldots, n_i - 1$) означают, что x_i является корнем кратности не меньше n_i. Таким образом, многочлен $f(x)$ степени меньше n имеет с учетом кратности $n = n_1 + n_2 + \ldots + n_s$ корней. Существует лишь один такой многочлен $f(x) = 0$. □

Лемма 7.78. *Пусть* $x_i \neq x_j$ *при* $i \neq j$. *Тогда*

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ & 1 & 2x_1 & \ldots & (n-1)x_1^{n-2} \\ & & 2 & \ldots & (n-1)(n-2)x_1^{n-3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_s & x_s^2 & \ldots & x_s^{n-1} \\ & 1 & 2x_s & \ldots & (n-1)x_s^{n-2} \\ & & 2 & \ldots & (n-1)(n-2)x_s^{n-3} \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{vmatrix} \neq 0.$$

Доказательство. Из леммы (7.77) вытекает, что система (7.42) с нулевой правой частью имеет единственное решение, откуда следует, что определитель этой системы не равен нулю. □

Теорема 7.79. *Для любой таблицы интерполяции* (7.40) *интерполяционный многочлен Лагранжа–Эрмита существует и единственен.*

Доказательство. Так как по лемме (7.12) определитель системы (7.42) не равен нулю, то эта система имеет единственное решение при любой правой части, откуда следует доказываемое. □

Замечание 7.80. Определитель из леммы 7.78 может быть вычислен явно. Он равен

$$\prod_{i=1}^{s} 1!\,2!\,\ldots\,(n_i-1)! \prod_{1 \le i < j \le s} (x_j - x_i)^{n_i n_j}.$$

Для вычисления из каждого столбца, начиная с последнего, вычтем предшествующий, умноженный на x_1, из строк вынесем множители и применим индукцию.

Линейные рекуррентные соотношения. Пусть члены последовательности комплексных чисел

$$\langle x_k \rangle = \langle x_0,\ x_1,\ x_2,\ \ldots,\ x_k,\ \ldots \rangle \tag{7.43}$$

удовлетворяют соотношению

$$x_k = d_1 x_{k-1} + d_2 x_{k-2} + \ldots + d_n x_{k-n} \qquad (k = n, n+1, \ldots), \tag{7.44}$$

где $d_j \in \mathbb{C}$ $(j = 1, 2, \ldots, n)$ и $d_n \neq 0$, а также удовлетворяют *начальным условиям*

$$x_0 = x_0', \quad x_1 = x_1', \quad \ldots, x_{n-1} = x_{n-1}'.$$

Равенство (7.44) называется *линейным рекуррентным соотношением* n-го порядка. Рассмотрим задачу получения свернутого выражение для k-го члена последовательности, удовлетворяющей заданному рекуррентному соотношению и начальным условиям.

По соотношению (7.44) построим многочлен

$$f(\lambda) = \lambda^n - d_1\lambda^{n-1} - d_2\lambda^{n-2} - \ldots - d_n, \tag{7.45}$$

называемый *характеристическим*.

Теорема 7.81. *Пусть корни $\lambda_1, \lambda_2, \ldots, \lambda_s$ характеристического многочлена* (7.45) *имеют кратности $n_1, n_2, \ldots, n_s$ соответственно. Тогда общий член последовательности, удовлетворяющей рекуррентному соотношению* (7.44), *имеет вид*

$$\begin{gathered}
x_k = c_{10}\lambda_1^k + c_{11}k\lambda_1^{k-1} + c_{12}k(k-1)\lambda_1^{k-2} + \ldots \\
\ldots + c_{1,n_1-1}k(k-1)\ldots(k-n_1+2)\lambda_1^{k-n_1+1} + \ldots \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\ldots + c_{s0}\lambda_s^k + c_{s1}k\lambda_s^{k-1} + c_{s2}k(k-1)\lambda_s^{k-2} + \ldots \\
\ldots + c_{s,n_s-1}k(k-1)\ldots(k-n_s+2)\lambda_s^{k-n_s+1},
\end{gathered}$$

где c_{ij} — некоторые числа. По начальным условиям эти числа определяются единственным образом.

Доказательству теоремы предпошлем несколько лемм.

Лемма 7.82. *Множество $\ell(F)$ последовательностей с членами из поля F относительно обычных операций сложения последовательностей и умножения их на числа из F образуют бесконечномерное линейное пространство над полем F.*

Доказательство. Легко проверяются все аксиомы линейного пространства. □

Лемма 7.83. *Множество V последовательностей, удовлетворяющих заданному рекуррентному соотношению* (7.44) (*при любых начальных условиях*), *образует в $\ell(F)$ подпространство размерности n.*

Доказательство. Соотношению (7.44) всегда удовлетворяет нулевая последовательность, поэтому $V \neq \emptyset$. Замкнутость V относительно операций сложения и умножения на число легко проверяется. Итак, V есть подпространство в $\ell(F)$. Докажем, что $\dim V = n$. Для этого проверим, что последовательности

$$\begin{aligned}
e^{(0)} &= \langle e_k^{(1)} \rangle = \langle 1,\ 0,\ \ldots,\ 0,\ e_n^{(0)},\ e_{n+1}^{(0)},\ \ldots \rangle, \\
e^{(1)} &= \langle e_k^{(2)} \rangle = \langle 0,\ 1,\ \ldots,\ 0,\ e_n^{(1)},\ e_{n+1}^{(1)},\ \ldots \rangle, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
e^{(n-1)} &= \langle e_k^{(n-1)} \rangle = \langle 0,\ 0,\ \ldots,\ 1,\ e_n^{(n-1)},\ e_{n+1}^{(n-1)},\ \ldots \rangle,
\end{aligned}$$

в которых члены, начиная с n-го, определяются по формуле (7.44), образуют базис пространства V. Действительно, линейная независимость системы $e^{(0)}, e^{(1)}, \ldots, e^{(n-1)}$ очевидна. Докажем, что эта система полна в V. Для этого достаточно показать, что любая последовательность, удовлетворяющая соотношению (7.44), может быть представлена в виде линейной комбинации последовательностей $e^{(0)}, e^{(1)}, \ldots, e^{(n-1)}$. Действительно, если последовательность (7.43) удовлетворяет (7.44), то

$$x_k = x_0 e_k^{(0)} + x_1 e_k^{(1)} + \ldots + x_{n-1} e_k^{(n-1)} \qquad (k = 0, 1, \ldots). \tag{7.46}$$

Докажем это индукцией по k. При $k = 0, 1, \ldots, n-1$ равенство (7.46) очевидно. Пусть $k \geq n$. По предположению индукции

$$x_j = x_0 e_j^{(0)} + x_1 e_j^{(1)} + \ldots + x_{n-1} e_j^{(n-1)} \quad (j = k-n, k-n+1, \ldots, k-1).$$

Умножая обе части этого равенства на d_{k-j} и складывая полученные равенства при $j = k-n, k-n+1, \ldots, k-1$ придем к (7.46). □

Лемма 7.84. *Пусть корни $\lambda_1, \lambda_2, \ldots, \lambda_s$ характеристического многочлена* (7.45) *имеют кратности $n_1, n_2, \ldots, n_s$ соответственно. Тогда n последовательностей*

$$\lambda_1^k,\ k\lambda_1^{k-1},\ k(k-1)\lambda_1^{k-2},\ \ldots,\ k\cdot(k-1)\cdot\ldots\cdot(k-n_1+2)\cdot\lambda_1^{k-n_1+1},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\lambda_s^k,\ k\lambda_s^{k-1},\ k(k-1)\lambda_s^{k-2},\ \ldots,\ k\cdot(k-1)\cdot\ldots\cdot(k-n_s+2)\cdot\lambda_s^{k-n_s+1},$$

образуют базис пространства V всех последовательностей, удовлетворяющих рекуррентному соотношению (7.44).

Доказательство. Во-первых докажем, что каждая из указанных последовательностей принадлежит V. Пусть $i \in \{1, 2, \ldots, s\}$. Так как

$$\lambda_i^n - d_1\lambda_i^{n-1} - d_2\lambda_i^{n-2} - \ldots - d_n = 0, \tag{7.47}$$

то $\lambda_i^k = d_1\lambda_i^{k-1} - d_2\lambda_i^{k-2} - \ldots - d_n\lambda_i^{k-n}$, т. е. λ_i^k удовлетворяет рекуррентному соотношению (7.44). Докажем, что если λ_i — кратный корень, то $k\lambda_i^{k-1}$ также удовлетворяет этому соотношению. Действительно, в этом случае $f'(\lambda_i) = 0$, т. е.

$$n\lambda_i^{n-1} - d_1(n-1)\lambda_i^{n-2} - d_2(n-2)\lambda_i^{n-2} - \ldots - d_{n-1} = 0. \tag{7.48}$$

Умножая (7.47) на $\lambda_i^{k-n-1}(k-n)$ и складывая с (7.48), умноженным на λ_i^{k-n}, получим

$$k\lambda_i^{k-1} - (k-1)d_1\lambda_i^{k-2} - d_2(k-1)\lambda_i^{k-2} - \ldots - d_{n-1}\lambda_i^{k-n}(k-n+1) - d_n(k-n)\lambda_i^{k-n+1} = 0,$$

т. е. $k\lambda_i^{k-1}$ удовлетворяет рекуррентному соотношению (7.44). Остальные последовательности проверяются аналогично.

Во-вторых докажем, что указанные последовательности линейно независимы и, так как $\dim V = n$, образуют базис. Для этого составим из первых n членов этих последовательностей матрицу:

$$\begin{pmatrix}
1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\
 & 1 & 2\lambda_1 & \ldots & (n-1)\lambda_1^{n-2} \\
 & & 2 & \ldots & (n-1)(n-2)\lambda_1^{n-3} \\
\multicolumn{5}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\
1 & \lambda_s & \lambda_s^2 & \ldots & \lambda_s^{n-1} \\
 & 1 & 2\lambda_s & \ldots & (n-1)\lambda_s^{n-2} \\
 & & 2 & \ldots & (n-1)(n-2)\lambda_s^{n-3} \\
\multicolumn{5}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots}
\end{pmatrix}$$

По лемме 7.78 определитель этой матрицы не равен нулю, следовательно, указанные последовательности линейно независимы. □

Доказательство теоремы 7.81. Непосредственно следует из леммы 7.84. Числа c_{ij} в формулировке теоремы — это (определяемые единственным образом) координаты произвольной последовательности, удовлетворяющей рекуррентному соотношению, в базисе из леммы 7.84. □

Пример 7.85. Рассмотрим линейное рекуррентное соотношение

$$x_k = x_{k-1} + x_{k-2} \qquad (k = 2, 3, \ldots) \tag{7.49}$$

с начальными условиями

$$x_0 = 0, \qquad x_1 = 1, \tag{7.50}$$

определяющее последовательность *чисел Фибоначчи*:

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

Характеристический многочлен имеет вид $\lambda^2 - \lambda - 1$. Его корни равны

$$\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2},$$

поэтому общий член последовательности, удовлетворяющей рекуррентному соотношению (7.49), имеет вид:

$$x_k = c_1 \cdot \left(\frac{1+\sqrt{5}}{2}\right)^k + c_2 \cdot \left(\frac{1-\sqrt{5}}{2}\right)^k \qquad (k = 0, 1, \ldots). \tag{7.51}$$

Подставляя в (7.51) при $k = 0, 1$ значения (7.50), получим систему линейных уравнений

$$\begin{cases} 0 = c_1 + c_2, \\ 1 = c_1 \cdot \dfrac{1+\sqrt{5}}{2} + c_2 \cdot \dfrac{1-\sqrt{5}}{2}, \end{cases}$$

из которой определим

$$c_1 = \frac{1}{\sqrt{5}}, \qquad c_2 = -\frac{1}{\sqrt{5}}.$$

Получаем, что общий член последовательности, удовлетворяющей рекуррентному соотношению (7.49) и начальным условиям (7.50), имеет вид:

$$x_k = \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^k - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2}\right)^k \qquad (k = 0, 1, \ldots).$$

Пример 7.86. Рассмотрим линейное рекуррентное соотношение

$$x_k = 4x_{k-1} - 4x_{k-2} \qquad (k = 2, 3, \ldots) \tag{7.52}$$

с начальными условиями

$$x_0 = 1, \qquad x_1 = 0. \tag{7.53}$$

Составляем характеристический многочлен $\lambda^2 - 4\lambda + 4$. Многочлен имеет один кратный корень $\lambda_{1,2} = 2$, поэтому общий член последовательности, удовлетворяющей рекуррентному соотношению (7.52), имеет вид:

$$x_k = c_1 \cdot 2^k + c_2 \cdot k \cdot 2^{k-1} \qquad (k = 0, 1, \ldots). \tag{7.54}$$

Подставляя в (7.54) при $k = 0, 1$ значения (7.53), получим систему линейных уравнений

$$\begin{cases} 1 = c_1, \\ 0 = 2c_1 + c_2, \end{cases}$$

из которой определим $c_1 = 1$, $c_2 = -2$. Получаем, что общий член последовательности, удовлетворяющей рекуррентному соотношению (7.52) и начальным условиям (7.53), имеет вид:

$$x_k = 2^k - 2 \cdot k \cdot 2^{k-1} = (1 - k) \cdot 2^k \qquad (k = 0, 1, \ldots).$$

Пример 7.87. Рассмотрим линейное рекуррентное соотношение

$$x_k = 5x_{k-1} - 6x_{k-2} - 4x_{k-3} + 8x_{k-4} \qquad (k = 4, 5, \ldots) \tag{7.55}$$

с начальными условиями

$$x_0 = 1, \qquad x_1 = -1, \qquad x_2 = 3, \qquad x_3 = 11. \tag{7.56}$$

Составляем характеристический многочлен:

$$\lambda^4 - 5\lambda^3 + 6\lambda^2 + 4\lambda - 8 = (x+1)(x-2)^3.$$

Многочлен имеет простой корень -1 и трехкратный корень 2, поэтому общий член последовательности, удовлетворяющей рекуррентному соотношению (7.55), имеет вид:

$$x_k = c_1(-1)^k + c_2 2^k + c_3 k 2^{k-1} + c_4 k(k-1)2^{k-2}. \tag{7.57}$$

Подставляя в (7.57) при $k = 0, 1, 2, 3$ значения (7.56), получим систему линейных уравнений

$$\begin{cases} 1 = c_1 + c_2, \\ -1 = -c_1 + 2c_2 + c_3, \\ 3 = c_1 + 4c_2 + 4c_3 + 2c_4, \\ 11 = -c_1 + 8c_2 + 12c_3 + 12c_4, \end{cases}$$

из которой определим $c_1 = 1$, $c_2 = 0$, $c_3 = 0$, $c_4 = 1$. Получаем, что общий член последовательности, удовлетворяющей рекуррентному соотношению (7.55) и начальным условиям (7.56), имеет вид:

$$x_k = (-1)^k + k(k-1)2^{k-2}. \qquad (k = 0, 1, \ldots).$$

Пример 7.88. Линейные рекуррентные соотношения появляются, например, при вычислении *трехдиагональных определителей* n-го порядка:

$$\Delta_n = \begin{vmatrix} b & c & 0 & 0 & \ldots & 0 & 0 \\ a & b & c & 0 & \ldots & 0 & 0 \\ 0 & a & b & c & \ldots & 0 & 0 \\ 0 & 0 & a & b & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & b & c \\ 0 & 0 & 0 & 0 & \ldots & a & b \end{vmatrix}.$$

Раскладывая определитель по первой строке, получаем

$$\Delta_n = b\Delta_{n-1} - c \cdot \begin{vmatrix} a & c & 0 & \ldots & 0 & 0 \\ 0 & b & c & \ldots & 0 & 0 \\ 0 & a & b & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & b & c \\ 0 & 0 & 0 & \ldots & a & b \end{vmatrix}.$$

Далее полученный определитель раскладываем по первому столбцу:

$$\Delta_n = b\Delta_{n-1} - ca \cdot \begin{vmatrix} b & c & \ldots & 0 & 0 \\ a & b & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b & c \\ 0 & 0 & \ldots & a & b \end{vmatrix} = b\Delta_{n-1} - ca\Delta_{n-2}.$$

Получили рекуррентное соотношение второго порядка $\Delta_n = b\Delta_{n-1} - ca\Delta_{n-2}$.

Пример 7.89.

$$\Delta_n = \begin{vmatrix} 6 & 5 & 0 & \ldots & 0 & 0 \\ -1 & 6 & 5 & \ldots & 0 & 0 \\ 0 & -1 & 6 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 6 & 5 \\ 0 & 0 & 0 & \ldots & -1 & 6 \end{vmatrix} = 6\Delta_{n-1} + 5\Delta_{n-2}.$$

Для решения рекуррентного соотношения составим характеристический многочлен $f(\lambda) = \lambda^2 - 6\lambda - 5 = (\lambda - 5)(\lambda + 1)$. Таким образом, $\Delta_n = c_1 5^n + c_2(-1)^n$. Учитывая начальные условия

$$\Delta_1 = 6, \qquad \Delta_2 = \begin{vmatrix} 6 & 5 \\ -1 & 6 \end{vmatrix} = 41,$$

составляем систему

$$\begin{cases} 6 = \ \ 5c_1 - c_2, \\ 41 = 25c_1 + c_2, \end{cases}$$

из которой находим $c_1 = \dfrac{47}{30}$, $c_2 = \dfrac{11}{6}$. Таким образом, $\Delta_n = \dfrac{47}{30} \cdot 5^n + \dfrac{11}{6} \cdot (-1)^n$.

Глава 8

Линейные отображения и преобразования

8.1. Определения и примеры

Рассмотрим два линейных пространства V, W, заданных над одним и тем же полем F. Отображение $\varphi : V \to W$ называется *линейным оператором* (*линейным отображением*), если выполнены следующие свойства (свойства линейности):

1. $\varphi(x + y) = \varphi x + \varphi y$ для произвольных векторов $x, y \in V$.
2. $\varphi(\alpha x) = \alpha \varphi x$ для произвольного вектора $x \in V$ и произвольного числа $\alpha \in F$.

Отображение $\varphi : V \to V$ называется *преобразованием*.

Примеры.

Нулевой оператор. Для произвольных пространств V, W, заданных над полем F определим $\varphi x = 0$ для произвольного $x \in V$. Отображение φ, очевидно, является линейным.

Тождественное преобразование. Для произвольного пространства V определим преобразование $\varepsilon : V \to V$ следующим образом: для произвольного $x \in V$ положим $\varepsilon x = x$. Отображение ε, очевидно, является линейным.

Преобразование проектирования. Преобразование $\varphi : V \to V$, ставящее в соответствие вектору x его проекцию $\mathrm{pr}_{V_1 \| V_2} x$ является линейным.

Действительно, пусть $x = x_1 + x_2$, $y = y_1 + y_2$; $x_1, y_1 \in V_1$, $x_2, y_2 \in V_2$. Тогда $x + y = (x_1 + y_1) + (x_2 + y_2)$, $x_1 + y_1 \in V_1$, $x_2 + y_2 \in V_2$ и поэтому

$$\mathrm{pr}_{V_1 \| V_2}(x + y) = x_1 + y_1 = \mathrm{pr}_{V_1 \| V_2} x + \mathrm{pr}_{V_1 \| V_2} y.$$

Таким образом, $\varphi(x + y) = \varphi x + \varphi y$.

С другой стороны, $\alpha x = \alpha x_1 + \alpha x_2$, поэтому

$$\mathrm{pr}_{V_1 \| V_2} \alpha x = \alpha\, \mathrm{pr}_{V_1 \| V_2} x.$$

Следовательно, $\varphi(\alpha x) = \alpha \varphi x$.

Упражнение 8.1. Дать (геометрическую) интерпретацию преобразования проектирования в пространствах $\mathbf{V}_2$ и $\mathbf{V}_3$.

Из свойств линейности оператора по индукции легко вывести следующее обобщение:

$$\varphi\left(\sum_{j=1}^{s}\alpha_j a_j\right)=\sum_{j=1}^{s}\alpha_j\varphi a_j. \tag{8.1}$$

В свойстве 2) полагая $\alpha=0$ получаем $\varphi 0=0$.

8.2. Матрица линейного оператора

Обозначим через $\Phi(V,W)$ множество всех линейных операторов, действующих из V в W. В этом разделе мы дадим обозрение множества $\Phi(V,W)$.

1. Пусть $e_1,e_2,\dots,e_n$ — некоторый базис пространства V. Для произвольного вектора $x=\sum_{j=1}^{n}\alpha_j e_j\in V$ из (8.1) получаем:

$$\varphi x=\sum_{j=1}^{n}\alpha_j\varphi e_j. \tag{8.2}$$

Таким образом, *линейный оператор восстанавливается однозначно по образам базисных векторов.*

2. Пусть $e_1,e_2,\dots,e_n$ и $f_1,f_2,\dots,f_m$ — базисы пространств V и W соответственно. Так как $\varphi e_j\in W$ $(j=1,2,\dots,n)$, то φe_j раскладывается по базису $f_1,f_2,\dots,f_m$. Пусть $[\varphi e_j]_{\mathfrak{f}}=(\alpha_{1j},\alpha_{2j},\dots,\alpha_{mj},)^{\top}$ — столбец координат вектора φe_j в базисе $f_1,f_2,\dots,f_m$. Матрица, j-ый столбец которой есть столбец координат $[\varphi e_j]_{\mathfrak{f}}$ $(j=1,2,\dots,n)$, называется *матрицей оператора* φ, построенной в базисах $e_1,e_2,\dots,e_n$ и $f_1,f_2,\dots,f_m$. Обозначение: $[\varphi]_{f,e}$. Из определения получаем

$$[\varphi]_{f,e}=(\alpha_{ij})\in F^{m\times n}.$$

3. Переходя в выражении (8.2) к координатам, получим:

$$[\varphi x]_{\mathfrak{f}}=\sum_{j=1}^{n}\alpha_j[\varphi e_j]_{\mathfrak{f}}.$$

Последнее равенство в матричной форме приобретет вид

$$[\varphi x]_{\mathfrak{f}}=[\varphi]_{f,g}[x]_g. \tag{8.3}$$

4. По аналогии с (8.3) рассмотрим равенство

$$[\varphi x]_{\mathbf{f}} = A[x]_{\mathbf{e}}, \tag{8.4}$$

где e, f — два фиксированных базиса пространств V, W соответственно, A — произвольная матрица из $F^{m\times n}$. Приведенная формула каждому вектору $x \in V$ ставит в соответствие вектор $\varphi x \in W$, таким образом, *определяет* оператор $\varphi : V \to W$. Исследуем его.

Лемма 8.2. *Отображение $\varphi : V \to W$, заданное формулой* (8.4), *является линейным, причем A — его матрица*: $[\varphi]_{f,e} = A$.

Доказательство. Для произвольных $x, y \in V$ имеем

$$[\varphi(x+y)]_{\mathbf{f}} = A[x+y]_{\mathbf{e}} = A([x]_{\mathbf{e}} + [y]_{\mathbf{e}}) = A[x]_{\mathbf{e}} + A[y]_{\mathbf{e}} = [\varphi x]_{\mathbf{f}} + [\varphi y]_{\mathbf{f}},$$

следовательно, $\varphi(x+y) = \varphi x + \varphi y$. Для произвольного $x \in V$ и произвольного $\alpha \in F$ имеем

$$[\varphi(\alpha x)]_{\mathbf{f}} = A[\alpha x]_{\mathbf{e}} = A(\alpha[x]_{\mathbf{e}}) = \alpha A[x]_{\mathbf{e}} = \alpha[\varphi x]_{\mathbf{f}},$$

следовательно, $\varphi(\alpha x) = \alpha\varphi x$. Итак, оператор φ — линейный.

Проверим теперь, что $[\varphi]_{f,e} = A$. Подставим e_j ($j = 1, 2, \ldots, n$) в (8.4): $[\varphi e_j]_{\mathbf{f}} = A[e_j]_{\mathbf{e}}$. Так как j-ая компонента столбца $[e_j]_{\mathbf{e}}$ равна 1, а остальные компоненты равны 0, то $A[e_j]_{\mathbf{e}}$ есть j-ый столбец матрицы A, значит A, по определению, есть матрица оператора φ. □

5. Переформулировка результатов пп. 2,4 приводит нас к следующему.

Следствие 8.3. *Отображение, ставящее в соответствие всякому оператору его матрицу, является биекцией из $\Phi(V, W)$ в $F^{m\times n}$.*

8.3. Операции с линейными отображениями

Определения.

- *Суммой двух операторов $\varphi, \psi \in \Phi(V, W)$* называется оператор χ, такой, что $\chi x = \varphi x + \psi x$ для произвольного $x \in V$. Обозначение для суммы: $\chi = \varphi + \psi$.
- *Произведением оператора $\varphi \in \Phi(V, W)$ на число $\alpha \in F$* называется оператор χ, такой, что $\chi x = \alpha(\varphi x)$ для произвольного $x \in V$. Обозначение для произведения оператора на число: $\chi = \alpha\varphi$.
- Рассмотрим три пространства U, V, W, заданных над одним и тем же полем F. Пусть ψ, φ — линейные операторы, действующие из U в V и из V в W соответственно: $U \xrightarrow{\psi} V \xrightarrow{\varphi} W$. *Произведением операторов* φ и ψ называется оператор θ, такое, что $\chi x = \varphi(\psi x)$. Обозначение для произведения операторов: $\chi = \varphi\psi$.

Итак,

$$(\varphi+\psi)x = \varphi x + \psi x,$$
$$(\alpha\varphi)x = \alpha(\varphi x),$$
$$(\varphi\psi)x = \varphi(\psi x).$$

Линейность результата операций.

Утверждение 8.4. 1) *Пусть* $\varphi, \psi \in \Phi(V, W)$, $\alpha \in F$, *тогда операторы* $\varphi+\psi$, $\alpha\varphi$ *— линейные, причем* $[\varphi+\psi] = [\varphi]+[\psi]$, $[\alpha\varphi] = \alpha[\varphi]$.

2) *Пусть* $\varphi \in \Phi(V, W)$, $\psi \in \Phi(U, V)$, *тогда оператор* $\varphi\psi$ *— линейный, причем* $[\varphi\psi] = [\varphi][\psi]$.

Доказательство. Все свойства доказываются аналогично. Приведем для примера два способа доказательства п. 2.
1 *способ* Пусть

$$\begin{array}{lll} g_1, g_2, \ldots, g_l & \text{— базис пространства} & U, \\ & & \downarrow\psi \\ e_1, e_2, \ldots, e_n & \text{— базис пространства} & V, \\ & & \downarrow\psi \\ f_1, f_2, \ldots, f_m & \text{— базис пространства} & W; \end{array} \tag{8.5}$$

$$[\varphi]_{f,e} = (\alpha_{ij}) \in F^{m\times n}, [\psi]_{e,g} = (\beta_{jk}) \in F^{n\times l}. \tag{8.6}$$

Теперь получаем $[(\varphi\psi)x]_{\mathfrak{f}} = [\varphi(\psi x)]_{\mathfrak{f}} = [\varphi]_{f,e}[\psi x]_{\mathfrak{e}} = [\varphi]_{f,e}[\psi]_{e,g}[x]_g$. Обозначим $A = [\varphi]_{f,e}[\psi]_{e,g}$, тогда $[(\varphi\psi)x]_{\mathfrak{f}} = A[x]_g$. Применяя лемму из предыдущего раздела, получаем, что оператор $\varphi\psi$ — линейный и $[\varphi\psi]_{f,g} = [\varphi]_{f,e}[\psi]_{e,g}$.
2 *способ* Докажем вначале линейность. Для произвольных $x, y \in U$ имеем $(\varphi\psi)(x+y) = \varphi(\psi(x+y)) = \varphi(\psi x+\psi y) = \varphi(\psi x)+\varphi(\psi y) = (\varphi\psi)x+(\varphi\psi)y$. Для произвольных $x \in U$, $\alpha \in F$ имеем $(\varphi\psi)(\alpha x) = \varphi(\psi(\alpha x)) = \varphi(\alpha\psi x) = \alpha\varphi(\psi x) = \alpha(\varphi\psi)x$. Итак, оператор $\varphi\psi$ — линейный.

Теперь вычислим $[\varphi\psi]_{f,g} = (\gamma_{ik})$. В системе обозначений (8.5, 8.6) для $k = 1, 2, \ldots, l$ получаем $(\varphi\psi)g_k = \varphi(\psi g_k) = \varphi\sum_{j=1}^n \beta_{jk}e_j = \sum_{j=1}^n \beta_{jk}\varphi e_j = \sum_{j=1}^n \beta_{jk}\sum_{i=1}^m \alpha_{ij}f_i = \sum_{i=1}^n f_i\sum_{j=1}^n \alpha_{ij}\beta_{jk}$. По определению матрицы линейного оператора получаем, что $\gamma_{ik} = \sum_{j=1}^n \alpha_{ij}\beta_{jk}$, или $[\varphi\psi]_{f,g} = [\varphi]_{f,e}[\psi]_{e,g}$. □

Замечание 8.5. Сумма, произведение и произведение на число линейных операторов являются операторами линейными в том числе и для бесконечномерного пространства. 2-ой способ доказательства проходит и для этого случая.

Другие свойства линейных операций. Линейные операции над операторами — это сложение операторов и умножение их на числа из поля.

Утверждение 8.6. *Для любых* $\varphi, \psi, \chi \in \Phi(V, W)$, $\alpha, \beta \in F$ *справедливы равенства*:

1. $\varphi + \psi = \psi = \varphi$,
2. $\varphi + (\psi + \chi) = (\varphi + \psi) + \chi$,
3. $\varphi + 0 = \varphi$, *где* 0 *— нулевой оператор*,
4. $\varphi + (-1)\varphi = 0$ (*таким образом*, $(-1)\varphi$ *— оператор* противоположный *оператору* φ),
5. $1\varphi = \varphi$,
6. $\alpha(\beta\varphi) = (\alpha\beta)\varphi$,
7. $(\alpha + \beta)\varphi = \alpha\varphi + \beta\varphi$,
8. $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$.

Доказательство. Доказательство приведенных выше свойств не должно вызывать затруднений. Докажем для примера двумя способами последнее свойство.
1 *способ* (*по определению*) Для произвольного $x \in V$ имеем $[\alpha(\varphi + \psi)x]_{\mathbf{f}} = [\alpha(\varphi + \psi)]_{f,e}[x]_{\mathbf{e}} = [\alpha\varphi]_{f,e}[x]_{\mathbf{e}} + [\alpha\psi]_{f,e}[x]_{\mathbf{e}} = [\alpha\varphi x]_{\mathbf{f}} + [\alpha\psi x]_{\mathbf{f}}$, откуда вытекает доказываемое.
2 *способ* (*через свойства матриц*) Для произвольного $x \in V$ имеем $\big(\alpha(\varphi + \psi)\big)x = \alpha\big((\varphi + \psi)x\big) = \alpha(\varphi x + \psi x) = \alpha(\varphi x) + \alpha(\psi x) = (\alpha\varphi)x + (\alpha\psi)x$, т. е. $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$.

$\square$

Замечание 8.7. 2-й способ доказательства проходит также и для случая бесконечномерного линейного пространства.

Следствие 8.8. *Множество* $\Phi(V, W)$ *относительно ранее введенных операций сложения операторов и умножения их на числа является линейным пространством. Отображение, ставящее в соответствие всякому преобразованию его матрицу, является изоморфизмом* $\Phi(V, W)$ *в* $F^{m\times n}$.

Доказательство. Замкнутость операций сложения операторов и умножения их на числа вытекает из утверждения о линейности результата операций. Аксиомы линейного пространства доказаны в предыдущем утверждении. Биективность указанного в формулировке следствия вытекает из утверждения п. 5 предыдущего раздела. Сохранение операций следует из утверждения о линейности результата операций. $\square$

Другие свойства операций умножения линейных операторов.

Утверждение 8.9. *Пусть* U, V, W, Q *— линейные пространства, заданные над одним полем* F. *Тогда для любых линейных операторов* $\varphi, \varphi_1 \in \Phi(V, W)$, $\psi, \psi_1 \in \Phi(U, V)$, $\chi \in \Phi(Q, U)$, *и произвольных* $\alpha, \beta \in F$ *справедливы равенства*:

1. $\varphi(\psi\chi) = (\varphi\psi)\chi$,

2. $\alpha(\varphi\psi) = (\alpha\varphi)\psi = \varphi(\alpha\psi)$,

3. $(\varphi + \varphi_1)\psi = \varphi\psi + \varphi_1\psi$,

4. $\varphi(\psi + \psi_1) = \varphi\psi + \varphi\psi_1$.

Доказательство. Как и свойства линейных операций сформулированные свойства легко могут быть выведены из определения, либо из соответствующих свойств матриц. □

8.4. Изменение матрицы оператора при замене базисов

Пусть $\varphi \in \Phi(V, W)$,

$$\left.\begin{matrix} e_1, e_2, \dots, e_n \\ e'_1, e'_2, \dots, e'_n \end{matrix}\right\} \text{два базиса пространства } V;$$
$$\left.\begin{matrix} f_1, f_2, \dots, f_m \\ f'_1, f'_2, \dots, f'_m \end{matrix}\right\} \text{два базиса пространства } W.$$

Изучим, как меняется матрица преобразования φ при переходе от базисов $e_1, e_2, \dots, e_n$ и $f_1, f_2, \dots, f_m$ к базисам $e'_1, e'_2, \dots, e'_n$ и $f'_1, f'_2, \dots, f'_m$ соответственно. Из равенств

$$\begin{aligned} [x]_{\mathbf{e}} &= Q_{e\to e'}[x]_{e'}, \\ [\varphi x]_{\mathbf{f}} &= Q_{f\to f'}[\varphi x]_{f'}, \\ [\varphi x]_{\mathbf{f}} &= [\varphi]_{f,e}[x]_e \end{aligned}$$

легко получается формула

$$[\varphi x]_{f'} = Q^{-1}_{f\to f'}[\varphi]_{f,e}Q_{e\to e'}[x]_{e'},$$

справедливая для произвольного $x \in V$. Обозначив $A = Q^{-1}_{f\to f'}[\varphi]_{f,e}Q_{e\to e'}$, получаем

$$[\varphi]_{f',e'} = Q^{-1}_{f\to f'}[\varphi]_{f,e}Q_{e\to e'}.$$

8.5. Ядро и образ оператора

Образом, или *множеством значений* оператора $\varphi \in \Phi(V, W)$, называется множество векторов из W, для которых в V существует по крайней мере один прообраз. Обозначение: φV или $\operatorname{Im}\varphi$. Итак, по определению,

$$\varphi V = \{y \in W : \exists\, x \in V, \varphi x = y\}.$$

Ядром, или *нуль-пространством*, оператора $\varphi \in \Phi(V, W)$, называется множество векторов из V, обращающихся в $0 \in W$. Обозначение: $\operatorname{Ker}\varphi$. Итак, по определению,

$$\operatorname{Ker}\varphi = \{x \in V : \varphi x = 0\}.$$

Утверждение 8.10. *Образ и ядро линейного оператора* $\varphi \in \Phi(V, W)$ *являются подпространствами в* W, V *соответственно.*

Доказательство. 1) Очевидно, что $\varphi V \neq \emptyset$. Для $y_1, y_2 \in \varphi V$ найдутся такие $x_1, x_2 \in V$, что $\varphi x_1 = y_1$, $\varphi x_2 = y_2$. Так как $y_1 + y_2 = \varphi(x_1 + x_2)$ и $x_1 + x_2 \in V$, то $y_1 + y_2 \in \varphi V$. Пусть $y = \varphi x \in V$, $\alpha \in F$, тогда $\alpha y = \alpha\varphi x = \varphi(\alpha x) \in \varphi V$.

2) Так как $\varphi 0 = \varphi 0 \cdot 0 = 0\varphi 0 = 0$, то $0 \in \varphi V$, поэтому $\varphi V \neq \emptyset$. Пусть $x_1 \in \operatorname{Ker}\varphi$, $x_2 \in \operatorname{Ker}\varphi$, тогда $\varphi(x_1+x_2) = \varphi x_1 + \varphi x_2 = 0$, следовательно, $x_1+x_2 \in \operatorname{Ker}\varphi$. Пусть $x \in \operatorname{Ker}\varphi$, $\alpha \in F$, тогда $\varphi(\alpha x) = \alpha\varphi x = 0$, следовательно, $\alpha x \in \operatorname{Ker}\varphi$. $\square$

Размерность образа линейного оператора называется его *рангом* и обозначается $\operatorname{rank}\varphi$, размерность ядра линейного оператора называется *дефектом* и обозначается $\operatorname{def}\varphi$.

Теорема 8.11 (Ранг и дефект оператора). 1. *Пусть* $e_1, e_2, \ldots, e_n$ *и* $f_1, f_2, \ldots, f_m$ *— произвольные базисы пространств* V, W *соответственно, тогда* $\operatorname{rank}\varphi = [\varphi]_{f,e}$.

2. *Имеет место равенство* $\operatorname{def}\varphi + \operatorname{rank}\varphi = \dim V$.

Доказательство. 1) Рассмотрим множество $[\varphi V]$ координатных столбцов всех векторов из $\varphi V = \{y = \varphi x : x \in V\}$. Имеем $[\varphi V] = \left\{y = [\varphi]_{f,e}x : x \in F^n\right\}$, таким образом, $[\varphi V]$ является линейной оболочкой столбцов матрицы $[\varphi]_{f,e}$ и имеет размерность $\operatorname{rank}[\varphi]_{f,e}$. Очевидно, $\dim V = \dim[\varphi V]$.

2) Рассмотрим множество $[\operatorname{Ker} V]$ координатных столбцов всех векторов из $\operatorname{Ker}\varphi$. Имеем $[\operatorname{Ker}\varphi] = \left\{x \in F^n : [\varphi]_{f,e}x = 0\right\}$, таким образом, $[\operatorname{Ker}\varphi]$ есть множество решений системы линейных уравнений, поэтому $\dim[\operatorname{Ker}\varphi] = n - \operatorname{rank}[\varphi]_{f,e}$, или $\operatorname{def}\varphi = n - \operatorname{rank}\varphi$. $\square$

8.6. Линейные преобразования

Напомним, что линейным преобразованием называется линейный оператор $\varphi : V \to V$.

- Преобразование называется *тождественным* (единичным) и обозначается ε, если $\varepsilon x = x$ для любого $x \in V$.

- Преобразование ψ_1 называется *левым обратным* для преобразования φ, если $\psi_1\varphi = \varepsilon$.

- Преобразование ψ_2 называется *правым обратным* для преобразования φ, если $\varphi\psi_2 = \varepsilon$.

- Преобразование называется *обратным* для преобразования φ и обозначается φ^{-1}, если $\varphi\varphi^{-1} = \varphi\varphi^{-1} = \varepsilon$. Преобразование φ называется *невырожденным*, если для него существует обратное преобразование.

Матрицей преобразования φ в базисе $e_1, e_2, \ldots, e_n$ называется матрица, составленная из координатных столбцов $[\varphi e_j]_{\mathbf{e}}$ ($j = 1, 2, \ldots, n$). Для матрицы преобразования используется обозначение $[\varphi]_{\mathbf{e}}$. Для произвольных φ, ψ из $\Phi(V, V)$

и произвольного α из F, используя свойства матриц линейных операторов, получаем

$$\begin{aligned} &[\alpha\varphi]_{\mathbf{e}} = \alpha[\varphi]_{\mathbf{e}}, \\ &[\varphi+\psi]_{\mathbf{e}} = [\varphi]_{\mathbf{e}} + [\psi]_{\mathbf{e}}, \\ &[\varphi\psi]_{\mathbf{e}} = [\varphi]_{\mathbf{e}}[\psi]_{\mathbf{e}}. \end{aligned} \tag{8.7}$$

В пространстве $\Phi(V, V)$ определим операцию возведения в степень:

$$\varphi^m = \underbrace{\varphi\varphi\ldots\varphi}_{m \text{ раз}}.$$

Из (8.7) по индукции получаем $[\varphi^m] = [\varphi]^m$.

Значением многочлена

$$f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \ldots + a_{m-1}\lambda + a_m \in F[\lambda] \tag{8.8}$$

*от преобразования*Значение многочлена от преобразования φ назовем преобразование

$$f(\varphi) = a_0\varphi^m + a_1\varphi^{m-1} + \ldots + a_{m-1}\varphi + a_m\varepsilon.$$

Значением многочлена от матрицы Значение многочлена от матрицы $A \in F^{n\times n}$ назовем матрицу

$$f(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mE.$$

Из (8.7) по индукции получаем $[f(\varphi)]_{\mathbf{e}} = f([\varphi]_{\mathbf{e}})$.

8.7. Собственные числа и собственные векторы преобразования

Подпространство W линейного пространства V называется *инвариантным относительно преобразования* φ, если $\varphi W \subseteq W$, т. е. $\varphi x \in W$ для любого x из W. *Сужением преобразования* (*индуцированным преобразованием*) φ на инвариантное подпространство W называется преобразование $\psi: W \to W$ такое, что $\psi x = \phi x$ для любого $x \in W$. Обозначение: $\phi \mid W = \psi$. Очевидно, преобразование $\phi \mid W$ — линейное.

Упражнение 8.12. Докажите, что образ и ядро линейного преобразования являются инвариантными пространствами. Опишите соответствующие индуцированные преобразования.

Вектор $x \in V$ называется *собственным вектором* преобразования φ, если для некоторого $\lambda \in F$ выполнены условия

$$\begin{aligned} &\varphi x = \lambda x, \\ &x \neq 0. \end{aligned} \tag{8.9}$$

Число λ при этом называется *собственным значением* (*числом*) преобразования φ. Говорят также, что собственный вектор x *принадлежит* или *относится* к собственному значению λ.

Обозначим через V_λ множество всех собственных векторов, принадлежащих собственному значению λ, дополненное нулевым вектором, иными словами:

$$V_\lambda = \{x \in V : \lambda x\}.$$

Утверждение 8.13. *Множество V_λ является подпространством пространства V.*

Доказательство. Очевидно $V_\lambda \neq \emptyset$. Далее, если $x, y \in V_\lambda$, $\alpha \in F$, то

$$\begin{aligned}
&\varphi(x+y) = \varphi x + \varphi y = \lambda x + \lambda y = \lambda(x+y),\\
&\varphi(\alpha x) = \alpha(\varphi x) = \alpha(\lambda x) = \lambda(\alpha x).
\end{aligned}$$

поэтому $x + y \in V_\lambda$, $\alpha x \in V_\lambda$. □

Подпространство V_λ называется *собственным пространством* преобразования φ.

Упражнение 8.14. Доказать что V_λ инвариантно относительно φ.

Утверждение 8.15. *Если x — собственный вектор, то для любого ненулевого $\alpha \in F$ вектор αx — тоже собственный, относящийся к тому же собственному числу.*

Доказательство.

$$\varphi(\alpha x) = \alpha(\varphi x) = \alpha(\lambda x) = \lambda(\alpha x).$$

□

Утверждение 8.16. *Собственные векторы — это в точности базисные векторы одномерных инвариантных подпространств.*

Доказательство. Пусть x — собственный вектор. По предыдущему утверждению αx для любого $\alpha \neq 0$ — тоже собственный, относящийся к тому же собственному числу, поэтому подпространство $L(x)$ инвариантно.

Пусть $L(x)$, где x — некоторый ненулевой вектор, инвариантно, т. е. $\varphi x \in L(x)$, или $\varphi x = \lambda x$ для некоторого $\lambda \in F$, следовательно, вектор x — собственный. □

Упражнение 8.17. Дайте геометрическую интерпретацию понятию собственного вектора.

Исследуем задачу нахождения собственных векторов преобразования φ. Пусть $e_1, \ldots, e_n$ — произвольный базис пространства V. Условия (8.9), очевидно, эквивалентны системе

$$\begin{cases} (\varphi - \lambda\varepsilon)x = 0, \\ x \neq 0; \end{cases} \quad \Leftrightarrow \quad \begin{cases} [\varphi - \lambda\varepsilon]_{\mathbf{e}}[x]_{\mathbf{e}} = 0, \\ x \neq 0; \end{cases}$$

или

$$\begin{cases} ([\varphi]_{\mathbf{e}} - \lambda E)\,[x]_{\mathbf{e}} = 0, \\ x \neq 0. \end{cases} \tag{8.10}$$

Итак, вектор x является собственным тогда и только тогда, когда его координатный столбец $[x]_{\mathbf{e}}$ является нетривиальным решением квадратной системы линейных уравнений (8.10). Для существования такого x необходимо и достаточно, чтобы

$$\det([\varphi]_{\mathbf{e}} - \lambda E) = 0. \tag{8.11}$$

По аналогии с (8.11) для произвольной матрицы $A \in F^{n\times n}$ относительно неизвестного λ рассмотрим уравнение

$$\det(A - \lambda E) = 0,$$

называемое *характеристическим уравнением матрицы* A. Его корни совпадают с корнями многочлена

$$\chi_A(\lambda) = \det(A - \lambda E)$$

из $F[\lambda]$. Этот многочлен называется *характеристическим многочленом матрицы* A.

Лемма 8.18. *Характеристические многочлены подобных матриц совпадают.*

Доказательство. Пусть $B = Q^{-1}AQ$ для некоторой невырожденной Q. Тогда

$$\begin{aligned}
&\det(\lambda E - B)\\
&\quad= \det\left(Q^{-1}\lambda EQ - Q^{-1}AQ\right)\\
&\quad= \det\left(Q^{-1}(\lambda E - A)Q\right)\\
&\quad= \underbrace{\det\left(Q^{-1}\right)}_{1/\det Q}\det(\lambda E - A)\det Q\\
&\quad= \det(\lambda E - A).
\end{aligned}$$

□

Из леммы следует, что уравнение (8.11), записанное для одно и того же преобразования φ в разных базисах, имеет один и тот же вид и поэтому корректны следующие определения. Уравнение (8.11) называется *характеристическим уравнением преобразования* φ. Левая часть этого уравнения

$$\chi(\lambda) = \det(\lambda E - [\varphi]_{\mathbf{e}})$$

называется *характеристическим многочленом преобразования* φ.

Из всего вышесказанного следует

Теорема 8.19. *Для того, что $\lambda \in F$ являлось собственным числом преобразования φ необходимо и достаточно, чтобы оно являлось корнем характеристического многочлена этого преобразования.*

8.7.1. Выражение коэффициентов характеристического многочлена через главные миноры матрицы

Миноры матрицы $A \in F^{n\times n}$ вида $M_A\binom{j_1,\dots,j_k}{j_1,\dots,j_k}$ называются *главными минорами* матрицы A. Сумма диагональных элементов (сумма главных миноров первого порядка) матрицы A называется ее *следом* и обозначается $\operatorname{tr} A$.

Лемма 8.20. *Определитель суммы двух матриц A и B порядка n равен сумме всех 2^n определителей матриц, получающихся из A заменой части столбцов соответсвующими столбцами из B.*

Доказательство. Пусть $a_1, \dots, a_n$ — столбцы матрицы A, и $b_1, \dots, b_n$ — столбцы матрицы B. Имеем

$$
\begin{aligned}
&\det(A + B)\\
&= \det(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)\\
&\quad= \det(a_1, a_2 + b_2, \dots, a_n + b_n) + \det(b_1, a_2 + b_2, \dots, a_n + b_n)\\
&\quad= \det(a_1, a_2, \dots, a_n + b_n) + \det(a_1, b_2, \dots, a_n + b_n)\\
&\qquad+ \det(b_1, a_2, \dots, a_n + b_n) + \det(b_1, b_2, \dots, a_n + b_n)\\
&\quad= \text{и т. д.}
\end{aligned}
$$

Каждый раз расщепляя определители по очередному столбцу мы получим в конце цепочки равенств 2^n определителей, удовлетворяющих доказываемым свойствам. □

Теорема 8.21. *Коэффициент s_k многочлена*

$$\det(A - \lambda E) = (-\lambda)^n + s_1(-\lambda)^{n-1} + \ldots + s_{n-1}(-\lambda) + s_n$$

равен сумме главных миноров порядка k. В частности, $s_1 = \operatorname{tr} A$, $s_n = \det A$.

Доказательство. Применим предыдущую лемму к матрицам A и $(-\lambda E)$. В указанной сумме каждый из определителей имеет следующий вид:

- столбцы с номерами $j_1, \dots, j_k$, где $1 \le j_1 < \ldots < j_k \le n$, совпадают с соответствующими столбцами матрицы A (столбцы первой группы),

- в остальных столбцах на диагонали стоит $-\lambda$, на остальных местах 0 (столбцы второй группы).

Раскладывая каждый определитель по столбцам второй группы, мы получим $\lambda^{n-k}\binom{j_1,\dots,j_k}{j_1,\dots,j_k}$. Для окончания доказательства осталось собрать слагаемые с одинаковым множителем λ^{n-k}. □

8.7.2. Матрица Фробениуса

Матрица вида

$$F(a_1, \ldots, a_n) = \begin{pmatrix} -a_1 & -a_2 & \ldots & -a_{n-1} & -a_n \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & & 1 & 0 \end{pmatrix}$$

называется *матрицей Фробениуса*.

Утверждение 8.22. *Характеристический многочлен матрицы Фробениуса* $F(a_1, \ldots, a_n)$ *равен*

$$\Phi(a_1, \ldots, a_n; \lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n.$$

Доказательство. Доказательство проведем индукцией по n. Основание индукции легко проверяется. Найдем теперь $\det(\lambda E - F(a_1, \ldots, a_n))$. Раскладывая этот определитель по последнему столбцу и пользуясь предположением индукции, получаем

$$\begin{aligned} &\det(\lambda E - F(a_1, \ldots, a_n)) \\ &\quad = a_n(-1)^{1+n}(-1)^{n-1} + (-\lambda)((-\lambda)^{n-1} + a_1(-\lambda)^{n-2} + \ldots + a_{n-1}) \\ &\quad = \Phi(a_1, \ldots, a_n; \lambda). \end{aligned}$$

□

Ввиду предыдущего утверждения матрица Фробениуса $F(a_1, \ldots, a_n)$ называется также *сопровождающей матрицей* многочлена $\Phi(a_1, \ldots, a_n; \lambda)$.

Следствие 8.23. *Всякий многочлен степени n со старшим коэффициентом* 1 *может быть характеристическим многочленом некоторой квадратной матрицы порядка n.*

8.8. Диагонализируемость линейного преобразования

Преобразование называется *диагонализируемым*, если существует базис, в котором матрица преобразования имеет диагональный вид.

Теорема 8.24 (Очень простое, но очень важное утверждение)**.** *Преобразование диагонализируемо тогда и только тогда, когда существует базис из собственных векторов преобразования.*

Доказательство. Необходимость Пусть в базисе $e_1, \ldots, e_n$

$$[\varphi]_{\mathbf{e}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Так как j-ый столбец матрицы $[\varphi]_{\mathbf{e}}$ есть

$$[\varphi e_j]_{\mathbf{e}} = (0, \ldots, 0, \lambda_j, 0, \ldots, 0)^{\top},$$

то $\varphi e_j = \lambda_j e_j$, т. е. e_j — собственный вектор ($j = 1, \ldots, n$).
Достаточность Пусть базис $e_1, \ldots, e_n$ состоит из собственных векторов, тогда $\varphi e_j = \lambda_j e_j$ ($j = 1, \ldots, n$), откуда

$$[\varphi]_{\mathbf{e}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

□

Пусть характеристический многочлен преобразования φ линейного пространства V, заданного над полем F, имеет вид

$$\chi(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s} p(\lambda),$$
$$\text{где } \lambda_i \in F, \lambda_i \neq \lambda_j \text{ при } i \neq j,$$
$$f(\lambda) \text{ корней в } F \text{ не имеет.}$$

Назовем *алгебраической кратностью* собственного значения λ_i величину k_i (т. е. его кратность, как корня характеристического многочлена).

Назовем *геометрической кратностью* собственного значения λ_i размерность собственного подпространства, принадлежащего собственному значению λ_i:

$$d_i = \dim V_{\lambda_i},$$

иными словами, величину $\operatorname{def}(\varphi - \lambda_i \varepsilon)$, или, что то же, максимальное число линейно независимых решений системы

$$([\varphi]_{\mathbf{e}} - \lambda_i E)x = 0.$$

Лемма 8.25. *Геометрическая кратность собственного числа не превосходит алгебраической.*

Доказательство. Пусть геометрическая кратность собственного значения λ_0 равна d_0. Следовательно, найдется система линейно независимых собственных векторов $e_1, \ldots, e_{d_0}$, относящихся к собственному значению λ_0. Дополним ее до базиса и в нем построим матрицу преобразования:

$$[\varphi]_{\mathbf{e}} = \left(\begin{array}{ccc|c} \lambda_0 & & & \\ & \ddots & & B \\ & & \lambda_0 & \\ \hline & 0 & & C \end{array}\right).$$

Очевидно, что

$$\chi_{\varphi} = \det(\lambda E - [\varphi]_{\mathbf{e}}) = (\lambda - \lambda_0)^{d_0} p(\lambda), \tag{8.12}$$

где $p(\lambda) \in F[\lambda]$. Из (8.12) получаем, что алгебраическая кратность собственного значения λ_0 не меньше d_0. □

Лемма 8.26. *Собственные векторы, относящиеся к разным собственным числам, линейно независимы.*

Доказательство. Пусть $\lambda_1, \ldots, \lambda_s$ — некоторые попарно различные собственные числа преобразования φ. В i-ой строке следующей таблицы выписана произвольная линейно независимая система собственных векторов, относящихся к собственному значению λ_i:

$$\begin{array}{lcl} e_{11}, e_{12}, \ldots, e_{1d_1} & \to & \lambda_1; \\ e_{21}, e_{22}, \ldots, e_{2d_2} & \to & \lambda_2; \\ \ldots & & \\ e_{s1}, e_{s2}, \ldots, e_{sd_s} & \to & \lambda_s. \end{array} \tag{8.13}$$

Покажем, что объединенная система собственных векторов (8.13) линейно независима.

Доказательство проведем индукцией по s. При $s = 1$ система линейно независима по предположению. Предположим теперь, что векторы, стоящие в первых $s - 1$ строках таблицы (8.13), линейно независимы и покажем, что все векторы в (8.13) линейно независимы. Рассмотрим произвольную равную нулю линейную комбинацию этих векторов:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} e_{ij} = 0. \tag{8.14}$$

Применим к обеим частям (8.14) преобразование φ:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} \underbrace{\varphi e_{ij}}_{\lambda_i e_{ij}} = 0, \tag{8.15}$$

откуда

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} \lambda_i e_{ij} = 0. \tag{8.16}$$

Теперь домножим обе части (8.14) на λ_s:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} \lambda_s e_{ij} = 0. \tag{8.17}$$

Вычитая (8.17) из (8.16), получаем

$$\sum_{i=1}^{\cancel{s}\, s-1} \sum_{j=1}^{d_i} \alpha_{ij} (\lambda_i - \lambda_s) e_{ij} = 0.$$

В последнем равенстве имеем нулевую линейную комбинацию векторов, по предположению индукции линейно независимых. Поэтому все коэффициенты этой линейной комбинации равны нулю:

$$\alpha_{ij}\underbrace{(\lambda_i - \lambda_s)}_{\neq 0} = 0 \quad (i = 1, \ldots, s-1),$$

Откуда

$$\alpha_{ij} = 0 \quad (i = 1, \ldots, s-1;\ j = 1, \ldots, k_i).$$

Теперь в правой части (8.14) остается лишь одна сумма:

$$\sum_{j=1}^{d_s} \alpha_{sj} e_{sj} = 0.$$

Так как векторы $e_{s1}, \ldots, e_{s1}$ линейно независимые, то и

$$\alpha_{sj} = 0 \quad (j = 1, \ldots, k_s).$$

Итак, все коэффициенты в произвольной линейной комбинации (8.14) равны нулю, поэтому система (8.13) линейно независима. □

Из двух предыдущих лемм и утверждения получаем следующий результат.

Теорема 8.27 (Критерий диагонализируемости)**.** *Преобразование диагонализируемо тогда и только тогда, когда сумма геометрических кратностей всех собственных чисел совпадает с размерностью пространства.*

8.9. Аннулирующий многочлен

Говорят, что многочлен $f(\lambda)$ *аннулирует* (обращает в ноль) преобразование φ на подпространстве $W \subseteq V$, если $W \subseteq \operatorname{Ker} f(\varphi)$, иными словами, $f(\varphi)x = 0$ для произвольного вектора $x \in W$. В данном случае $f(\lambda)$ называется также *аннулирующим многочленом* преобразования φ на подпространстве (относительно подпространства) W.

Аннулирующий многочлен минимальной степени со старшим коэффициентом 1 называется *минимальным аннулирующим*, или просто *минимальным*.

Пример 8.28. a) Минимальным многочленом, аннулирующим тождественное преобразование ε на произвольном подпространстве, является, очевидно, $f(\lambda) = \lambda - 1$.

b) Пусть

$$[\varphi] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

В данном случае первый базисный вектор e_1 собственный, поэтому относительно подпространства $L(e_1)$ аннулирующим (и минимальным) многочленом, является $\lambda - 1$. Относительно всего пространства, легко проверить, аннулирующим является многочлен $f(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Действительно,

$$[f(\varphi)] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^2 - 2\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Заметим, что многочлен $f(\lambda)$ совпадает с характеристическим.

Утверждение 8.29. *Минимальный многочлен существует и единственен.*

Доказательство. Существование Вначале мы докажем существование аннулирующего многочлена. В линейном пространстве $\Phi(V, V)$ рассмотрим систему векторов $\varepsilon, \varphi, \varphi^2, \ldots, \varphi^{n^2}$. Так как $\dim \Phi(V, V) = n^2$, то рассматриваемая система линейно зависима, т. е. найдутся такие коэффициенты $\alpha_0, \alpha_1, \ldots, \alpha_{n^2}$, что

$$\alpha_0\varepsilon + \alpha_1\varphi + \ldots \alpha_{n^2}\varphi^{n^2} = 0.$$

Очевидно, что многочлен $f(\lambda) = \alpha_0 + \alpha_1\lambda + \ldots \alpha_{n^2}\lambda^{n^2}$ аннулирует преобразование φ на любом подпространстве $W \subseteq V$. Отсюда следует существование минимального аннулирующего многочлена.

Единственность Пусть $f_1(\lambda), f_2(\lambda)$ — минимальные многочлены преобразования φ на подпространстве W. Докажем, что многочлен $f_1(\lambda) - f_2(\lambda)$ является аннулирующим. Действительно,

$$(f_1(\varphi) - f_2(\varphi))x = f_1(\varphi)x - f_2(\varphi)x = 0$$

для произвольного вектора $x \in W$. У разности $f_1(\lambda) - f_2(\lambda)$ старшие члены взаимно уничтожаются, следовательно, либо $f_1(\lambda) - f_2(\lambda) = 0$, либо степень $f_1(\lambda) - f_2(\lambda)$ меньше степени многочленов $f_1(\lambda)$, $f_1(\lambda)$. Последнее однако невозможно, так как $f_1(\lambda), f_2(\lambda)$ — минимальные, следовательно, $f_1(\lambda) - f_2(\lambda) = 0$, и поэтому $f_1(\lambda) = f_2(\lambda)$. □

Утверждение 8.30. *Пусть $f(\lambda)$ — минимальный многочлен преобразования φ относительно подпространства W. Тогда множество всех многочленов, аннулирующих φ на W есть множество $\{f(\lambda)p(\lambda) : p(\lambda) \in F[\lambda]\}$.*

Доказательство. Так как $p(\varphi)f(\varphi)x = p(\varphi)0 = 0$ для произвольного $x \in W$, то $f(\lambda)p(\lambda)$ аннулирующий.

Покажем, что других аннулирующих многочленов, кроме многочленов вида $f(\lambda)p(\lambda)$, нет. Для этого поделим произвольный аннулирующий многочлен $g(\lambda)$ с остатком на $f(\lambda)$:

$$g(\lambda) = f(\lambda)p(\lambda) + r(\lambda), \tag{8.18}$$

причем либо $r(\lambda) = 0$, либо $\deg r(\lambda) < \deg f(\lambda)$. Так как $r(\varphi)x = g(\varphi)x - p(\lambda)f(\lambda)x = 0$ для произвольного $x \in W$, то $r(\lambda)$ — также аннулирующий, поэтому $r(\lambda) = 0$. Теперь из (8.18) получаем $g(\lambda) = f(\lambda)p(\lambda)$. □

Утверждение 8.31. *Пусть многочлены $f_1(\lambda)$, $f_2(\lambda)$ являются минимальными аннулирующими многочленами для преобразования φ на подпространствах W_1, W_2 соответственно, причем $W_1 \subseteq W_2$. Тогда $f_2 \vdots f_1$.*

Доказательство. Так как $W_1 \subseteq W_2$, то $f_2(\lambda)$ аннулирует φ на W_1. Доказываемое теперь следует из предыдущего утверждения. □

Утверждение 8.32. *Пусть многочлены $f_1(\lambda)$, $f_2(\lambda)$ являются минимальными аннулирующими многочленами для преобразования φ на подпространствах W_1, W_2 соответственно. Тогда* НОК$(f_1(\lambda), f_2(\lambda))$ *является минимальным аннулирующим многочленом для преобразования φ на подпространстве $W_1 + W_2$.*

Доказательство. Сначала докажем, что $f(\lambda) = \text{НОК}(f_1(\lambda), f_2(\lambda))$ аннулирует преобразование φ на подпространстве $W_1 + W_2$. Пусть

$$f(\lambda) = p_1(\lambda)f_1(\lambda) = p_2(\lambda)f_2(\lambda).$$

Для произвольного вектора $x = x_1 + x_2 \in W$, где $x_1 \in W_1$, $x_2 \in W_2$, имеем $f(\varphi)x = f(\varphi)(x_1+x_2) = f(\varphi)x_1 + f(\varphi)x_2 = p_1(\varphi)f_1(\varphi)x_1 + p_2(\varphi)f_2(\varphi)x_2 = 0$. Итак, многочлен $f(\lambda)$ — аннулирующий на $W_1 + W_2$.

Докажем теперь, что $f(\lambda)$ — минимальный многочлен, т. е. из всех аннулирующих многочленов многочлен $f(\lambda)$ имеет минимальную степень. Так как $W_1 \subseteq W_1 + W_2$ и $W_2 \subseteq W_1 + W_2$, то по предыдущему утверждению произвольный аннулирующий на подпространстве $W_1 + W_2$ многочлен должен делиться без остатка как на многочлен $f_1(\lambda)$, так и на многочлен $f_2(\lambda)$. Из всех многочленов, удовлетворяющих этим свойствам, минимальную степень имеет $f(\lambda)$. Следовательно, этот многочлен минимальный аннулирующий. □

8.9.1. Метод Крылова построения минимального многочлена

В данном разделе мы опишем способ нахождения минимального многочлена, аннулирующего преобразование φ на всем пространстве V.

Пусть $e_1, e_2, \ldots, e_n$ — базис пространства V. Для произвольного $i \in \{1, 2, \ldots, n\}$ в силу конечномерности пространства найдется натуральное s, такое, что система

$$e_i, \varphi e_i, \varphi^2 e_i, \ldots, \varphi^{s-1} e_i$$

линейно независима, а система

$$e_i, \varphi e_i, \varphi^2 e_i, \ldots, \varphi^{s} e_i$$

линейно зависима. Пусть

$$\alpha_0 e_i + \alpha_1 \varphi e_i + \ldots + \alpha_{s-1}\varphi^{s-1} e_i + \varphi^s e_i = 0$$

для некоторых α_j $(j = 0, 1, \ldots, s-1)$. Легко видеть, что многочлен

$$f_i(\lambda) = \alpha_0 + \alpha_1\lambda + \ldots + \alpha_{s-1}\lambda^{s-1} + \lambda^s$$

является минимальным для $L(e_i)$.

По утверждению 4 минимальным многочленом, аннулирующим φ на всем пространстве V, является $\text{НОК}(f_1(\lambda), \ldots, f_n(\lambda))$.

Пример 8.33. Построить минимальный аннулирующий многочлен преобразования φ, заданного в некотором базисе матрицей

$$[\varphi] = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

Для базиса e_1, e_2, e_3, в котором задана матрица преобразования, имеем

$$[e_1] = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad [\varphi e_1] = \begin{pmatrix} 0 \\ -4 \\ -2 \end{pmatrix}, \quad [\varphi^2 e_1] = \begin{pmatrix} -4 \\ -16 \\ -8 \end{pmatrix};$$

$$[e_2] = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad [\varphi e_2] = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, \quad [\varphi^2 e_2] = \begin{pmatrix} 4 \\ 12 \\ 4 \end{pmatrix};$$

$$[e_3] = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad [\varphi e_3] = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

В каждой строке приведенной таблицы векторы линейно зависимые. Найдем их нулевые нетривиальные комбинации:

$$\begin{array}{lll} 4e_1 - 4\varphi e_1 + \varphi^2 e_1 = 0 & \to & f_1(\lambda) = 4 - 4\lambda + \lambda^2 = (\lambda - 2)^2; \\ 4e_2 - 4\varphi e_2 + \varphi^2 e_2 = 0 & \to & f_2(\lambda) = 4 - 4\lambda + \lambda^2 = (\lambda - 2)^2; \\ 4e_3 - 4\varphi e_3 = 0 & \to & f_3(\lambda) = 4 - 4\lambda + \lambda^2 = \lambda - 2. \end{array}$$

Минимальный многочлен имеет вид $f(\lambda) = \text{НОК}(f_1(\lambda), f_2(\lambda), f_3(\lambda)) = (\lambda - 2)$. Как и в примере 8.28 в данном случае характеристический многочлен $\chi(\lambda) = (\lambda - 2)^3$ является аннулирующим. В следующей теореме утверждается, что это справедливо для любого преобразования.

Теорема 8.34 (Гамильтон—Кэли). *Любое линейное преобразование является корнем своего характеристического многочлена*

Доказательство. Мы покажем, что произвольная матрица A является корнем своего характеристического многочлена $\chi_A(\lambda)$, т. е. $\chi_A(A) = 0$.

Рассмотрим квадратную матрицу $(\lambda E - A)^* = (\alpha_{ij})$, в которой $\alpha_{ij} = (-1)^{i+j} M_A\binom{j}{i}$. Имеем

$$(\lambda E - A)^*(\lambda E - A) = \det(\lambda E - A)E.$$

В полученное равенство подставим $\lambda = A$, тогда в его правой части получаем $\chi_A(A)$, а в левой 0. □

Замечание 8.35. Доказанная теорема равносильна утверждению, что характеристический многочлен преобразования является аннулирующим многочленом того же преобразования во всем пространстве.

Замечание 8.36. Приведенное доказательство не основывалось на результатах данного раздела и является другим доказательством существования аннулирующего многочлена.

Теорема 8.37 (Теорема о корнях аннулирующего многочлена). *Пусть многочлен*

$$\chi(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s}$$
$$(\lambda_i \neq \lambda_j \text{ при } i \neq j)$$

является характеристическим многочленом преобразования φ. Тогда минимальным многочленом, аннулирующим φ на всем пространстве, является многочлен

$$\mu(\lambda) = (\lambda - \lambda_1)^{l_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{l_s}, \tag{8.19}$$

где $l_1, \ldots, l_s$ — некоторые натуральные числа, такие, что $1 \leq l_i \leq k_i$ $(i = 1, 2, \ldots, s)$.

Доказательство. По теореме Гамильтона-Кэли характеристический многочлен является аннулирующим. Минимальный многочлен является его делителем и, поэтому, имеет вид (8.19), причем $l_i \le k_i$ $(i = 1, 2, \dots, s)$.

Теперь докажем, что $l_i \ge 1$. Для этого рассмотрим собственный вектор x. Имеем $\varphi x = \lambda_i x$. Легко видеть, что многочлен $f_i(\lambda) = \lambda - \lambda_i$ является минимальным аннулирующим преобразование φ на подпространстве $L(x)$. Необходимое неравенство следует теперь из утверждения 4. □

Замечание 8.38. Из доказанной теоремы следует, что множество корней минимального многочлена без учета их кратностей есть в точности множество характеристических чисел. Кратности, с которыми они встречаются в минимальном многочлене не превосходят алгебраических кратностей характеристических чисел.

8.10. Канонический вид преобразования комплексного линейного пространства (Жорданова форма линейного преобразования)

8.10.1. Определения

Жордановой клеткой называется квадратная матрица порядка n, следующего вида:

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_0 & \ddots & 0 & 0 \\ \multicolumn{3}{c}{\dots\dots\dots\dots} & \ddots & \ddots & \dots \\ 0 & 0 & 0 & \dots & \lambda_0 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_0 \end{pmatrix}.$$

Жордановой матрицей J называется блочно-диагональная матрица, составленная из жордановых клеток:

$$J = \operatorname{diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_1), \dots, J_{k_s}(\lambda_1)).$$

Базис пространства V называется *Жордановым*, если в этом базисе матрица преобразования является Жордановой (имеет *Жорданову форму*). В данном случае базис также называют каноническим базисом преобразования, а саму матрицу преобразования в этом базисе — *каноническим видом*, или *Жордановой формой* преобразования.

8.10.2. Цель

Теорема 8.39 (Жордан). *Для любого преобразования φ комплексного линейного n-мерного пространства существует Жорданов базис. Жорданова форма определена единственным с точностью до перестановки жордановых клеток образом.*

Теорема 8.40 (Жордан (матричная формулировка)). *Для любой матрицы* $A \in \mathbb{C}^{n\times n}$ *найдется такая невырожденная матрица* Q, *что* $J = Q^{-1}AQ$ *есть жорданова матрица. Матрица* J *определяется единственным образом с точностью до перестановки жордановых клеток.*

Следствие 8.41 (Критерий подобия матриц на полем C). *Для того, чтобы матрицы* A *и* B *из* $A \in \mathbb{C}^{n\times n}$ *были подобны над полем* $\mathbb{C}$ *необходимо и достаточно совпадение их жордановых форм.*

8.10.3. Разложение пространства в прямую сумму корневых подпространств

Рассмотрим минимальный многочлен

$$f(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}\dots(\lambda - \lambda_s)^{k_s},$$

аннулирующий преобразование φ на всем пространстве V.

Множество $P_i = \mathrm{Ker}(\varphi - \lambda_i)^{k_i}$ называется *корневым подпространством*, принадлежащим собственному числу (корню) λ_i $(i = 1, 2, \dots, s)$.

Теорема 8.42. *Пространство* V *есть прямая сумма корневых подпространств* P_i:

$$V = P_1 \dot{+} P_2 \dot{+} \dots \dot{+} P_s.$$

Доказательство. Легко видеть, что многочлены

$$f_i(\lambda) = \frac{f(\lambda}{(\lambda - \lambda_i)^{k_i}}, \quad (i = 1, 2, \dots, s)$$

взаимно просты, поэтому найдутся такие $u_1(\lambda), u_2(\lambda), \dots, u_s(\lambda)$, что

$$1 = \sum_{i=1}^{s} f_i(\lambda)u_i(\lambda),$$

откуда

$$\varepsilon = \sum_{i=1}^{s} f_i(\varphi)u_i(\varphi),$$

следовательно, для любого $x \in V$

$$x = \sum_{i=1}^{s} \underbrace{f_i(\varphi)u_i(\varphi)x}_{x_i}.$$

Обозначим $x_i = f_i(\varphi)u_i(\varphi)x$, тогда

$$x = \sum_{i=1}^{s} x_i. \tag{8.20}$$

Докажем, что формула (8.20) задает разложение произвольного вектора x по прямой сумме $P_1 \dotplus \ldots \dotplus P_s$.

Сперва покажем, что $x_i \in P_i$. Действительно,

$$(\varphi - \lambda_i \varepsilon)^{k_i} x_i = \underbrace{(\varphi - \lambda_i \varepsilon)^{k_i} f_i(\varphi)}_{f(\varphi)=0} u_i(\varphi) x = 0.$$

Теперь докажем, что рассматриваемая сумма прямая. Для этого достаточно установить единственность разложения по пространствам P_i $(i = 1, 2, \ldots, s)$ для нулевого вектора. Пусть

$$0 = \sum_{j=1}^{s} x_j, \text{ где } x_j \in P_j \qquad (8.21).$$

Применим к обеим частям равенства (8.21) преобразование $f_i(\varphi)$, тогда

$$0 = \sum_{j=1}^{s} f_i(\varphi) x_j = f_i(\varphi) x_i,$$

последнее равенство справедливо в силу

$$f_i(\varphi) x_j = \prod_{\substack{\nu = 1 \\ \nu \neq i}} (\varphi - \lambda_\nu \varepsilon)^{k_{nu}} x_j = 0.$$

Так как многочлены $f_i(\lambda)$ и $(\lambda - \lambda_i)^{k_i}$ взаимно простые, то существуют такие $u(\lambda)$, $v(\lambda)$, что

$$1 = u(\lambda) f_i(\lambda) + v(\lambda)(\lambda - \lambda_i)^{k_i},$$

поэтому

$$x_i = u(\varphi) \underbrace{f_i(\varphi) x_i}_{0} + v(\varphi) \underbrace{(\varphi - \lambda_i \varepsilon)^{k_i} x_i}_{0, \text{ т.к. } x_i \in P_i} = 0.$$

Итак, компонента разложения $x_i = 0$ $(i = 1, 2 \ldots, s)$ определяется однозначно, поэтому рассматриваемая сумма прямая. □

Замечание 8.43. Вместо $f(\lambda)$ в теореме можно рассматривать любой аннулирующий многочлен преобразования φ. Формулировка теоремы и доказательство при этом никак не изменятся.

Замечание 8.44. Корневые пространства P_i $(i = 1, 2, \ldots, s)$ инвариантны относительно преобразования φ.

Замечание 8.45. Сужение преобразования φ на подпространство P_i (индуцированное преобразование) имеет одно собственное значение λ_i.

8.10.4. Построение Жорданова базиса

В силу замечаний 2, 3 предыдущего пункта мы можем на время ограничиться рассмотрением линейного преобразования φ пространства V с одним собственным числом λ_0. В этом случае минимальный многочлен имеет вид $f(\lambda) = (\lambda-\lambda_0)^k$. Обозначим $\psi = \varphi - \lambda_0\varepsilon$, тогда
$$\psi^k = 0. \tag{8.22}$$

Назовем *высотой* вектора $x \in V$ такое $h \geq 0$, что $\psi^{h-1}x \neq 0$, но $\psi^h x = 0$. В нашем случае высота произвольного вектора не превосходит k. Существует единственный вектор высоты 0 — нулевой вектор. Векторы высоты 1 это в точности собственные векторы преобразования.

Обозначим $L_h = \mathrm{Ker}\,\psi^h$. Иными словами, L_h — это множество векторов высоты $\leq h$. Имеем

$$\{0\} = L_0 \subseteq \underbrace{L_1}_{\substack{\text{собственное}\\ \text{подпространство}}} \subseteq \ldots \subseteq L_k = L_{k+1} = \ldots = V.$$

Следующая лемма показывает, что $L_h \neq L_{h+1}$ при $h \leq k$, и поэтому

$$\{0\} = L_0 \subset L_1 \subset \ldots \subset L_k = L_{k+1} = \ldots = V.$$

Лемма 8.46. *Если $L_h = L_{h+1}$, то $L_h = L_l$ для любого $l \geq h$.*

Доказательство. Воспользуемся методом доказательства от противного. Пусть

$$L_h = L_{h+1} \neq L_{h+2},$$

тогда найдется такой $x \in V$, что

$$x \notin L_h = L_{h+1}, \quad x \in L_{h+2}.$$

Используя определение пространств L_{h+2} и L_{h+1}, соответственно получаем

$$\begin{aligned} \psi^{h+2}x = 0 \quad &\Rightarrow \quad \psi^{h+1}\psi x = 0 \quad \Rightarrow \quad \psi x \in L_{h+1}, \\ \psi^{h+1}x \neq 0 \quad &\Rightarrow \quad \psi^{h}\psi x \neq 0 \quad \Rightarrow \quad \psi x \notin L_{h}. \end{aligned}$$

Итак, $\psi x \in L_{h+1}$, однако $\psi x \notin L_h$, что невозможно, так как $L_h = L_{h+1}$. □

Опишем алгоритм построения Жорданова базиса.

- на предварительном шаге необходимо построить базисы пространств $L_1, L_2, \ldots L_k$;
- далее найти такую линейно независимую систему $a_{11}, \ldots, a_{1t_1}$, для которой справедливо соотношение

$$L_k = L_{k-1} \dotplus L(a_{11}, \ldots, a_{1t_1}); \tag{8.23}$$

Таблица 8.1: Схема построения Жорданова базиса

$$
\begin{aligned}
L_k &= L_{k-1} \dotplus L(a_{11}, \ \dots, \ a_{1t_1}),\\
L_{k-1} &= L_{k-2} \dotplus L(\psi a_{11}, \ \dots, \ \psi a_{1t_1}, \ a_{21}, \ \dots, \ a_{2t_2}),\\
L_{k-2} &= L_{k-3} \dotplus L(\psi^2 a_{11}, \ \dots, \ \psi^2 a_{1t_1}, \ \psi a_{21}, \ \dots, \ \psi a_{2t_2}, \ a_{31}, \ \dots, \ a_{3t_3}),\\
&\vdots\\
L_{k-h+2} &= L_{k-h+1} \dotplus L(\psi^{h-2} a_{11}, \ \dots, \ \psi^{h-2} a_{1t_1}, \ \dots, \ a_{h-1,1}, \ \dots, \ a_{h-1,t_{h-1}}),\\
L_{k-h+1} &= L_{k-h} \dotplus L(\psi^{h-1} a_{11}, \ \dots, \ \psi^{h-1} a_{1t_1}, \ \dots, \ \psi a_{h-1,1}, \ \dots, \ \psi a_{h-1,t_{h-1}}, a_{h1}, \ \dots, a_{ht_h}),\\
&\vdots\\
L_1 &= L_0 \dotplus L(\psi^{k-1} a_{11}, \dots, \psi^{k-1} a_{1t_1}, \dots, \psi a_{k-1,1}, \dots, \psi a_{k-1,t_k}, a_{k1}, \dots, a_{kt_k}).
\end{aligned}
$$

- для каждого $h = 2, 3, \dots, k$ необходимо найти линейно независимую систему $a_{h1}, \dots, a_{ht_h}$ (возможно, $t_h = 0$). такую, что

$$
\begin{aligned}
&L_{k-h+1} = L_{k-h}\\
&\quad \dotplus L(\psi^{h-1} a_{11}, \dots, \psi^{h-1} a_{1t_1}, \dots, \psi a_{h-1,1}, \dots, \psi a_{h-1,t_{h-1}}, a_{h1}, \dots, a_{ht_h}).
\end{aligned} \tag{8.24}
$$

Необходимые соотношения еще раз приведены в таблице 8.1 В настоящем пункте мы докажем, что построенная система векторов

$$
\begin{array}{cccccccc}
a_{11} \dots & a_{1t_1} & & & & & & \\
\psi a_{11} \dots & \psi a_{1t_1} & \psi a_{21} \dots & \psi a_{2t_2} & & & & \\
\vdots & \vdots & \vdots & \vdots & & & & \\
\psi^{k-1} a_{11} \dots & \psi^{k-1} a_{1t_1} & \psi^{k-2} a_{21} \dots & \psi^{k-2} a_{2t_2} & \dots & a_{k1} \dots a_{kt_k}
\end{array} \tag{8.25}
$$

образует Жорданов базис пространства V.

Сперва докажем возможность построения системы (8.25).

Лемма 8.47. *Системы векторов*

$$a_{h1}, \dots, a_{ht_h} \quad (h = 1, 2, \dots, k),$$

удовлетворяющие условиям (8.23–8.24), *существуют.*

Доказательство. Лемма доказывается индукцией по h.
Основание индукции $h = 1$ Система $a_{11}, \dots a_{1t_1}$ является дополнением базиса пространства L_{k-1} до базиса пространства L_k. Так как $L_{k-1} \subset L_k$, то такая система существует.
Индуктивный переход $h - 1 \to h$ Пусть $u_1, \dots, u_p$ — базис подпространства L_{k-h}, а $v_1, \dots, v_r$ — базис подпространства L_{k-h+1}. Обозначим векторы

$$\psi^{h-2} a_{11}, \dots, \psi^{h-2} a_{1t_1}, \dots, a_{h-1,1}, \dots, a_{h-1,t_{h-1}}$$

через $w_1, \dots, w_q$ соответственно. По предположению индукции, векторы

$$v_1, \dots, v_r, w_1, \dots, w_q$$

составляют базис подпространства L_{k-h+2}. Итак, имеем

$$\begin{aligned} &L_{k-h+2} = L_{k-h+1} \dotplus L(w_1, \dots, w_q), \\ &L_{k-h+1} = L(v_1, \dots, v_r), \\ &L_{k-h} = L(u_1, \dots, u_p). \end{aligned}$$

Докажем, что векторы

$$u_1, \dots, u_p, \psi w_1, \dots, \psi w_q$$

образуют линейно независимую систему в L_{k-h+1} и, следовательно, могут быть дополнены (векторами $a_{h1}, \dots, a_{ht_h}$) до базиса пространства L_{k-h+1}. Для этого рассмотрим линейную комбинацию

$$\sum_{j=1}^{p} \alpha_j u_j + \sum_{i=1}^{q} \beta_i \psi w_i = 0. \tag{8.26}$$

Применим преобразование ψ^{k-h} к обеим частям (8.26):

$$\sum_{j=1}^{p} \alpha_j \underbrace{\psi^{k-h} u_j}_{0} + \psi^{k-h+1} \sum_{i=1}^{q} \beta_i w_i = 0.$$

Получим

$$\psi^{k-h+1} \sum_{i=1}^{q} \beta_i w_i = 0 \quad \Rightarrow \quad \sum_{i=1}^{q} \beta_i w_i \in L_{k-h+1}.$$

Раскладывая эту сумму по базису пространства L_{k-h+1}, получаем

$$\sum_{i=1}^{q} \beta_i w_i = \sum_{j=1}^{r} \gamma_j v_j \tag{8.27}$$

для некоторых $\gamma_1, \dots, \gamma_r$. Так как по предположению индукции векторы $v_1, \dots, v_r, w_1, \dots, w_q$ линейно независимы, то в (8.27)

$$\beta_i = 0 \quad (i = 1, \dots, q); \qquad \gamma_j = 0 \quad (j = 1, \dots, r).$$

Теперь из (8.26) получаем

$$\sum_{j=1}^{p} \alpha_j u_j = 0.$$

Так как векторы $u_1, \dots, u_p$ линейно независимы, то $\alpha_j = 0$ $(j = 1, \dots, r)$. Итак, в линейной комбинации (8.26) коэффициенты равны нулю и, следовательно, векторы $v_1, \dots, v_r, \psi w_1, \dots, \psi w_q$ линейно независимы. □

Лемма 8.48. *Система* (8.25) *образует базис пространства* V.

Доказательство. По построению L_1 есть линейная оболочка векторов, стоящих в нижней строке таблицы (8.25). Подпространство L_2 есть линейная оболочка векторов, стоящих в двух последних строках. Поднимаясь так далее снизу вверх по таблице (8.25) мы получаем, что пространство $L_k = V$ есть линейная оболочка векторов системы (8.25). $\square$

Вертикальный ряд векторов в таблице (8.25) назовем *цепочкой*. Векторы в одной цепочке будем рассматривать в последовательности от собственного вектора к вектору максимальной высоты (снизу вверх). Покажем, что каждой цепочке соответствует одна Жорданова клетка. Таким образом, вся совокупность векторов (8.25) является Жордановым базисом.

Лемма 8.49. *Каждой цепочке векторов* $a_1, \ldots a_r$ *таблицы* (8.25) *в матрице преобразования соответствует Жорданова клетка.*

Доказательство. Для доказательства достаточно заметить, что вектор a_1 — собственный, поэтому $\varphi a_1 = \lambda_0 a_1$, и $a_j = \psi a_{j+1}$ ($j = 2, 3, \ldots, r$), поэтому $\varphi a_{j+1} = \lambda_0 a_{j+1} + a_j$. $\square$

В заключение пункта отметим, что в случае нескольких собственных значений преобразования φ необходимо проделать те же вычисления с каждым собственным числом.

8.10.5. Единственнось Жордановой формы

Лемма 8.50. *Для произвольного натурального* i

$$\operatorname{rank}(J_n(\lambda_0) - \lambda E))^i = \begin{cases} n - i, & \textit{если} \quad \lambda = \lambda_0, \quad n - i > 0, \\ 0, & \textit{если} \quad \lambda = \lambda_0, \quad n - i \leq 0, \\ n, & \textit{если} \quad \lambda \neq \lambda_0, \quad n - i \leq 0. \end{cases}$$

Обозначим через $l_j = l_j(\lambda_0)$ число Жордановых клеток вида $J_n(\lambda_0)$ в Жордановой форме J преобразования φ. Пусть $r_j = r_j(\lambda_0) = \operatorname{rank}(\varphi - \lambda_0\varepsilon)^j$. Следующая лемма показывает, что Жорданова форма преобразования определена однозначно с точностью до перестановки диагональных клеток.

Теорема 8.51. *Для числа Жордановых клеток порядка* j *в Жордановой форме* J *справедливо соотношение*

$$l_j(\lambda_0) = r_{j-1}(\lambda_0) - 2r_j(\lambda_0) + r_{j+1}(\lambda_0).$$

Доказательство. Выразим величины

$$r_j(\lambda_0) = \operatorname{rank}(\varphi - \lambda_0\varepsilon)^j = \operatorname{rank}(J - \lambda_0 E)^j$$

через $l_j(\lambda_0)$. При возведении блочно-диагональной матрицы в степень каждый блок возводится в степень независимо и матрица сохраняет блочную структуру.

Таким образом, величина $r_j(\lambda_0)$ равна сумме рангов каждого блока. Обозначим через $\bar{r}$ сумму рангов блоков, соответствующих собственным числам, отличным от λ_0. Пользуясь леммой, получаем

$$r_j(\lambda_0) = l_{j+1} + 2l_{j+2} + \ldots + (p-j)l_p + \bar{r} = \sum_{i=j+1}^{p} (i-j)l_i + \bar{r},$$

откуда

$$\begin{aligned} & r_{j-1} - 2r_j + r_{j+1} \\ & \quad = \sum_{i=j}^{p} (i-j)l_i - 2\sum_{i=j+1}^{p} (i-j)l_i + \sum_{i=j+2}^{p} (i-j)l_i \\ & \quad = l_j + 2l_{j+1} + \sum_{i=j+2}^{p} (i-j)l_i \\ & \quad -2\left(l_{j+1} + \sum_{i=j+2}^{p} (i-j)l_i\right) \\ & \quad + \sum_{i=j+2}^{p} (i-j)l_i \\ & \quad = l_j. \end{aligned}$$

□

Итак, число Жордановых клеток заданного порядка с заданным числом на диагонали полностью определяется формулами из предыдущей теоремы и зависит только от преобразования, но никак не от базиса. Поэтому *Жорданова форма (но не базис) линейного преобразования определены однозначно с точностью до перестановок диагональных клеток.*

8.11. Матричные функции

8.11.1. Значение многочлена от матрицы

Утверждение 8.52. *Если матрицы A, B подобны, причем $B = Q^{-1}AQ$, тогда для любого многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ выполнено соотношение*

$$g(B) = Q^{-1}g(A)Q. \tag{8.28}$$

Утверждение 8.53. *Для произвольного многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ и произвольной блочно-диагональной матрицы* $\mathrm{diag}(A_1, \ldots, A_p$

$$g(\mathrm{diag}(A_1, \ldots, A_p)) = \mathrm{diag}(g(A_1), \ldots, g(A_p)). \tag{8.29}$$

Утверждение 8.54.

$$(J_n(\lambda_0))^m = \begin{pmatrix} \lambda_0^m & \binom{m}{1}\lambda_0^{m-1} & \binom{m}{2}\lambda_0^{m-2} & \ldots & \binom{m}{n-1}\lambda_0^{m-n+1} \\ 0 & \lambda_0^m & \binom{m}{1}\lambda_0^{m-1} & \ldots & \binom{m}{n-2}\lambda_0^{m-n+2} \\ \ldots & & & & \\ 0 & 0 & 0 & \ldots & \lambda_0^m \end{pmatrix}. \tag{8.30}$$

Доказательство. Утверждение легко доказывается индукцией по m. □

Утверждение 8.55. *Для произвольного многочлена* $g(\lambda) \in \mathbb{C}[\lambda]$ *выполняется соотношение*

$$g(J_n(\lambda_0)) = \begin{pmatrix} g(\lambda_0) & \dfrac{g'(\lambda_0)}{1!} & \dfrac{g''(\lambda_0)}{2!} & \dots & \dfrac{g^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & g(\lambda_0) & \dfrac{g'(\lambda_0)}{1!} & \dots & \dfrac{g^{(n-2)}(\lambda_0)}{(n-2)!} \\ \dots & & & & \\ 0 & 0 & 0 & \dots & g(\lambda_0) \end{pmatrix}. \tag{8.31}$$

Доказательство. Если $g(\lambda) = \lambda^m$, то формула 8.30 превращается в 8.31. Далее доказательство не должно вызывать затруднений. □

Пример 8.56. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найти A^{2000}.

Характеристический многочлен в данном случае совпадает с минимальным и имеет вид $f(\lambda) = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, поэтому

$$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \quad Q = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \quad J = Q^{-1}AQ.$$

Отсюда

$$A = QJQ^{-1}, \quad A^{2000} = QJ^{2000}Q^{-1} = Q\begin{pmatrix} 2^{2000} & 2000 \cdot 2^{1999} \\ 0 & 2^{2000} \end{pmatrix} Q^{-1}.$$

8.11.2. Функции от матриц

Пусть

$$f(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s}$$

— минимальный многочлен матрицы A. Будем говорить, что произвольная функция $g(\lambda) : \mathbb{C} \to \mathbb{C}$ *задана на спектре матрицы* A, если существует набор значений

$$\begin{aligned} &g(\lambda_1), g'(\lambda_1), \dots, g^{k_1-1}(\lambda_1); \\ &g(\lambda_2), g'(\lambda_2), \dots, g^{k_2-1}(\lambda_1); \\ &\dots \\ &g(\lambda_s), g'(\lambda_s), \dots, g^{k_s-1}(\lambda_1). \end{aligned} \tag{8.32}$$

В данном случае говорят также, что *функция* $g(\lambda)$ *на спектре матрицы* A *принимает значения* (8.32). Обозначим через $g_A(\lambda)$ многочлен минимальной степени, совпадающий на спектре матрицы A с функцией $g(\lambda)$ (так называемый интерполяционный многочлен Лагранжа–Эрмита, заданный на спектре матрицы A). Значение $g(A)$ функции $g(\lambda)$ определим как значение многочлена $g_A(\lambda)$ от матрицы A:

$$g(A) = g_A(A).$$

Сразу отметим, что, по предыдущей теореме, *степень интерполяционного многочлена Лагранжа–Эрмита, заданного на спектре матрицы A, меньше степени минимального многочлена этой матрицы.*

Пример 8.57. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найти e^A.

Минимальный многочлен $f(\lambda) = (\lambda - 2)^2$ имеет вторую степень, поэтому ищем интерполяционный многочлен в виде

$$g_A(\lambda) = f_1\lambda + f_2.$$

На спектре матрицы A функция $g(\lambda) = e^\lambda$ принимает следующие значения: $g(2) = e^2$, $g'(2) = e^2$, поэтому

$$\begin{aligned} g_A(2) &= 2f_1 + f_2 = e^2, \\ g'_A(2) &= f_1 = e^2; \end{aligned}$$

откуда получаем $f_1 = e^2$, $f_2 = -e^2$, следовательно, $g_A(\lambda) = e^2\lambda - e^2$, поэтому

$$e^A = g_A(A) = e^2A - e^2E = \begin{pmatrix} 0 & e^2 \\ -e^2 & 2e^2 \end{pmatrix}.$$

Следующее утверждение показывает корректность введенного определения для того случая, когда функция $g(\lambda)$ сама является многочленом.

Теорема 8.58. *Пусть $g(\lambda)$, $h(\lambda)$ — многочлены из $\mathbb{C}[\lambda]$, A — матрица из $\mathbb{C}^{n\times n}$. Для того, чтобы $g(A) = h(A)$ необходимо и достаточно, чтобы значения многочленов $g(\lambda)$, $h(\lambda)$ совпадали на спектре матрицы A.*

Доказательство. Пусть

$$f(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s}$$

— минимальный многочлен матрицы A, тогда

$$\begin{aligned} & g(A) = h(A) \\ \Leftrightarrow\quad & g(A) - h(A) = 0 \\ \Leftrightarrow\quad & g(\lambda) - h(\lambda) \text{ — аннулирующий многочлен матрицы } A \\ \Leftrightarrow\quad & g(\lambda) - h(\lambda) = p(\lambda)f(\lambda) \text{ для некоторого } p(\lambda) \in \mathbb{C}[\lambda] \\ \Leftrightarrow\quad & \lambda_i \text{ — корень } g(\lambda) - h(\lambda) \text{ кратности } \geq k_i \ (i = 1, 2, \ldots, s) \\ \Leftrightarrow\quad & g^{(j)}(\lambda_i) - h^{(j)}(\lambda_i) = 0 \ (i = 1, \ldots, s;\, j = 0, \ldots, k_i - 1). \end{aligned}$$

□

Пример 8.59. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найти A^{2000}.

Этот пример уже рассматривался. Приведем другой способ его решения. Для функции $g(\lambda) = \lambda^{2000}$ найдем интерполяционный многочлен

$$g_A(\lambda) = f_1\lambda + f_2,$$

заданный на спектре матрицы A. Имеем

$$g_A(2) = 2f_1 + f_2 = 2^{2000},$$
$$g'_A(2) = f_1 = 2000 \cdot 2^{1999};$$

откуда $f_1 = 2000 \cdot 2^{1999}$, $f_2 = -1999 \cdot 2^{2000}$, следовательно,

$$g_A(\lambda) = 2000 \cdot 2^{1999}\lambda - 1999 \cdot 2^{2000} = 2^{2000}(1000\lambda - 1999).$$

Подставляя $\lambda = A$ в $g_A(\lambda)$, получаем

$$A^{2000} = 2^{2000} \cdot \begin{pmatrix} 1 & 1000 \\ -1000 & 1001 \end{pmatrix}.$$

Следующие утверждения являются аналогами утверждений 1, 2, 4 из раздела 8.11.1.

Утверждение 8.60. 5 *Если матрицы A, B подобны, причем $B = Q^{-1}AQ$, тогда для любой функции $g(\lambda)$, определенной на спектре матрицы A, $g(B)$ существует, причем*

$$g(B) = Q^{-1}g(A)Q.$$

Доказательство. Так как A и B подобны, то у них совпадают минимальные многочлены и, следовательно, наборы значений функции $g(\lambda)$ на их спектре, поэтому $g(B)$ существет и $g_A(\lambda) = g_B(\lambda)$. Теперь по утверждению 1 получаем

$$g(B) = g_B(B) = g_A(B) = Q^{-1}g_A(A)Q == Q^{-1}g(A)Q.$$

□

Утверждение 8.61. *Для произвольной функции $g(\lambda)$, определенной на спектре блочно-диагональной матрицы* $\operatorname{diag}(A_1, \ldots, A_p)$ *функция $g(A_i)$ существует для каждого $i = 1, 2 \ldots, p$ и*

$$g(\operatorname{diag}(A_1, \ldots, A_p)) = \operatorname{diag}(g(A_1), \ldots, g(A_p)).$$

Доказательство. Пусть $f_i(\lambda)$ — минимальный многочлен матрицы A_i ($i = 1, 2, \ldots, p$), а $f(\lambda)$ — минимальный многочлен матрицы A. Так как $f(\lambda) = \text{НОК}(f_1(\lambda), \ldots, f_p(\lambda))$, то все корни многочлена $f_i(\lambda)$ являются корнями многочлена $f(\lambda)$, причем в последнем каждый из корней имеет не меньшую кратность (спектр матрицы A_i содержится в спектре матрицы A). Следовательно, $g(A_i)$ существует, $g_A(\lambda) = g_{A_i}(\lambda)$ и поэтому, используя утверждение 2, получаем

$$\begin{aligned} &g(\operatorname{diag}(A_1, \ldots, A_p)) \\ &\quad = g_A(\operatorname{diag}(A_1, \ldots, A_p)) \\ &\quad = \operatorname{diag}(g_A(A_1), \ldots, g_A(A_p)) \\ &\quad = \operatorname{diag}(g_{A_1}(A_1), \ldots, g_{A_p}(A_p)) \\ &\quad = \operatorname{diag}(g(A_1), \ldots, g(A_p)). \end{aligned}$$

□

Утверждение 8.62. *Пусть функция* $g(\lambda)$ *определена на спектре матрицы* $J_n(\lambda_0)$, *тогда справедлива формула* (8.31).

Доказательство. Для краткости обозначим $A = J_n(\lambda_0))$. Минимальный многочлен матрицы A равен $f(\lambda) = (\lambda - \lambda_0)^n$, поэтому интерполяционный многочлен $g_A(\lambda)$ удовлетворяет n условиям вида

$$g_A^{(j)}(\lambda_0) = g^{(j)}(\lambda_0) \quad (j = 0, 1, \ldots, n-1),$$

откуда по утверждению 4 получаем доказываемое. □

Глава 9

Билинейные функции

см. пособие Золотых, Ильичев, Таланов "Билинейные функции и их применение"

Глава 10

Евклидовы пространства

см. пособие Золотых, Ильичев, Таланов "Билинейные функции и их применение"

Глава 11

Линейные преобразования унитарных и евклидовых пространств

11.1. Сопряженные преобразования

Рассмотрим линейное преобразование φ унитарного или евклидова пространства V. Отображение $V \to V$ называется *сопряженным* к преобразованию φ и обозначается φ^*, если

$$(\varphi x, y) = (x, \varphi^* y) \tag{11.1}$$

для произвольных x, y из V.

Теорема 11.1. *Для любого преобразования φ унитарного или евклидова пространства V преобразование φ^* существует, единственно и является линейным, причем*

$$[\varphi^*]_e = \overline{\Gamma_e^{-1} [\varphi]_e^\top \Gamma_e}. \tag{11.2}$$

Доказательство. Существование Покажем, что линейное преобразование ψ с матрицей

$$[\psi]_e = \overline{\Gamma_e^{-1} [\varphi]_e^\top \Gamma_e}$$

является сопряженным к преобразованию φ. Используя правила выражения скалярного произведения через координаты векторов и матрицу Грама и правила выражения координатного столбца образа вектора через координаты прообраза и матрицу преобразования, для произвольных x, y из V получаем

$$\begin{aligned}
(x, \psi y) &= [x]^\top \Gamma \overline{[\psi][y]} \\
&= [x]^\top \Gamma \overline{\overline{\Gamma_e^{-1} [\varphi]_e^\top \Gamma_e} [y]} \\
&= ([\varphi][x])^\top \Gamma \overline{[y]} \\
&= (\varphi x, y).
\end{aligned}$$

Единственность и проч Выражая скалярное произведение через координаты векторов, равенство (11.1) перепишем в виде

$$([\varphi][x])^\top \Gamma\overline{[y]} = [x]^\top \Gamma\overline{[\varphi^* y]},$$

откуда $$[x]^\top[\varphi]^\top \Gamma\overline{[y]} = [x]^\top \Gamma\overline{[\varphi^* y]},$$

так как x — любой, то $$[\varphi]^\top \Gamma\overline{[y]} = \Gamma\overline{[\varphi^* y]},$$

следовательно, $$[\varphi^* y] = \underbrace{\overline{\Gamma^{-1}[\varphi]^\top \Gamma}}_{A=[\varphi^*]}[y].$$

Таким образом, преобразование φ^* единственно, является линейным и имеет матрицу (11.2). □

Замечание Если базис ортонормированный, то матрица Грама единичная и формула (11.2) приобретает особенно простой вид:

$$[\varphi^*]_e = \overline{[\varphi]}_e^\top.$$

Матрица $\overline{A}^\top$ называется *сопряженной* к матрице $A \in F^{n\times n}$ и обозначается A^*. Таким образом, в *ортонормированном базисе матрица сопряженного преобразования сопряжена с матрицей исходного преобразавания*.

Теорема 11.2 (Свойства операции сопряжения)**.** *Для произвольных преобразований φ, ψ и любого скаляра $\alpha \in F$ справедливы следующие свойства*:

1. $(\varphi^*)^* = \varphi$;
2. $(\varphi + \psi)^* = \varphi^* + \psi^*$;
3. $(\alpha\varphi)^* = \overline{\alpha}\varphi^*$;
4. $(\varphi\psi)^* = \psi^*\varphi^*$.

Доказательство. Свойства можно вывести из соответствующих свойств матриц. Проиллюстрируем на примере свойства 1) другой способ доказательства. Для произвольных x, y из V имеем:

$$\begin{aligned}(\varphi x, y) = (x, \varphi^* y) = \overline{(\varphi^* y, x)}\\ = \overline{(y, \varphi^{**} x)} = (\varphi^{**} x, y).\end{aligned}$$

Откуда $(\varphi x - \varphi^{**}x, y) = 0$. Последнее равенство справедливо для произвольного вектора y. В частности, для $y = \varphi x - \varphi^{**}$ имеем

$$(\varphi x - \varphi^{**}x, \varphi x - \varphi^{**}x) = 0,$$

поэтому $\varphi x = \varphi^{**}x$, где x — любой вектор из V, поэтому $\varphi = \varphi^{**}$. □

11.2. Теорема Шу́ра

Лемма 11.3. *Ортогональное дополнение к собственному вектору линейного преобразования φ унитарного или евклидового пространства является инвариантным подпространством для сопряженного преобразования φ^*.*

Доказательство. Пусть x — собственный вектор, λ — соответствующее ему собственное число: $\varphi x = \lambda x$. Покажем, что

$$\varphi^* x^{\perp} \subseteq x^{\perp}.$$

Действительно, для произвольного $y \in x^{\perp}$ имеем

$$(x, \varphi^* y) = (\varphi x, y) = \lambda(x, y) = 0.$$

□

Лемма 11.4. *Для любого линейного преобразования φ комплексного n-мерного линейного пространства V существует инвариантное подпространство размерности $n-1$.*

Доказательство. Способ 1 См. теорему Жордана.
Способ 2 В пространстве V введем скалярное произведение. Для преобразования φ^* найдем собственный вектор x. Подпространство $x^{\perp}$, разумеется, имеет размерность $n-1$ и по предыдущей лемме инвариантно относительно преобразования $\varphi^{**} = \varphi$. □

Замечание 11.5. Второй способ доказательства дает, конечно, более простой алгоритм нахождения $(n-1)$-мерного инвариантного подпространства.

Лемма 11.6. *Для любого линейного преобразования φ комплексного n-мерного линейного пространства V существует система инвариантных подпространств V_i $(i = 1, 2, \dots, n)$, такая, что* $\dim V_i = i$ *и*

$$V_1 \subset V_2 \subset \ldots \subset V_n = V.$$

Доказательство. Способ 1 См. теорему Жордана.
Способ 2 В V найдем $(n-1)$-мерное инвариантное (относительно φ) подпространство. Рассмотрим сужение $\varphi \mid V_{n-1}$ преобразования φ на подпространство V_{n-1}. В нем по лемме 2 снова найдем инвариантное подпространство V_{n-2} размерности на 1 меньше и т. д. □

Теорема 11.7 (Шур). *Для любого линейного преобразования φ унитарного пространства существует ортонормированный базис, в котором матрица преобразования верхнетреугольная.*

Доказательство. Способ 1 Построим жорданов базис $e_1, \dots, e_n$ преобразования φ. С помощью процесса ортогонализации ортонормируем этот базис. Полученный базис обозначим через $f_1, \dots, f_n$. Заметим, что матрица перехода $Q_{e \to f}$ верхнетреугольная. Имеем

$$[\varphi]_f = Q_{e \to f}^{-1} [\varphi]_e Q_{e \to f},$$

причем все три матрицы в правой части верхнетреугольные. Поэтому матрица $[\varphi]_f$ тоже верхнетреугольная и $f_1, \ldots, f_n$ — искомый базис.
Способ 2 По предыдущей лемме существуют инвариантные подпространства V_i, такие, что $V_1 \subset V_2 \subset \ldots \subset V_n$ и $\dim V_i = i$. Пусть f_1 — произвольный нормированный вектор из V_1, f_2 — нормированный вектор из V_2, ортогональный V_1, $\ldots$, f_n — нормированный вектор из V_n, ортогональный V_{n-1}. Легко видеть, что базис $f_1, \ldots, f_n$ удовлетворяет условиям теоремы. □

11.3. Нормальные преобразования

Преобразование φ евклидова или унитарного пространства называется *нормальным*, если
$$\varphi^*\varphi = \varphi\varphi^*.$$
Матрица $A \in F^{n\times n}$ называется *нормальной*, если $A^*A = AA^*$.

Теорема 11.8. *Критерий нормального преобразования унитарного пространства*] *Преобразование φ унитарного пространства является нормальным тогда и только тогда, когда существует ортонормированный базис пространства, состоящий из собственных векторов.*

Доказательство. Достаточность Пусть $e_1, \ldots, e_n$ — система собственных векторов преобразования φ, составляющая ортонормированный базис пространства, тогда
$$[\varphi]_e = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \quad [\varphi^*]_e = \operatorname{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_n).$$
Легко видеть, что $[\varphi^*]_e[\varphi]_e = [\varphi]_e[\varphi^*]_e$, и, следовательно, $\varphi^*\varphi = \varphi\varphi^*$.
Необходимость Пусть φ — нормальное преобразование. По теореме Шура для него существует ортонормированный базис $e_1, \ldots, e_n$, в котором $[\varphi]_e = (\beta_{ij})$ — верхнетреугольная матрица, т. е. $\beta_{ij} = 0$ при $i > j$. Итак,
$$[\varphi]_e = \begin{pmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1n} \\ 0 & \beta_{22} & \ldots & \beta_{2n} \\ \ldots & & & \\ 0 & 0 & \ldots & \beta_{nn} \end{pmatrix} \quad \text{и} \quad [\varphi^*]_e = \begin{pmatrix} \overline{\beta}_{11} & 0 & \ldots & 0 \\ \overline{\beta}_{12} & \overline{\beta}_{22} & \ldots & 0 \\ \ldots & & & \\ \overline{\beta}_{1n} & \overline{\beta}_{2n} & \ldots & \overline{\beta}_{nn} \end{pmatrix}.$$
Из определения нормального преобразования получаем
$$[\varphi][\varphi^*] - [\varphi^*][\varphi] = 0.$$
Выразим через β_{ij} диагональные элементы матрицы, стоящей в правой части приведенного равенства:
$$\begin{aligned} |\beta_{11}|^2 + |\beta_{12}|^2 + \ldots + |\beta_{1n}|^2 - |\beta_{11}|^2 &= 0, \\ |\beta_{22}|^2 + |\beta_{23}|^2 + \ldots + |\beta_{2n}|^2 - |\beta_{12}|^2 - |\beta_{22}|^2 &= 0, \\ &\ldots \\ |\beta_{nn}|^2 - |\beta_{1n}|^2 - |\beta_{2n}|^2 - \ldots - |\beta_{nn}|^2 &= 0. \end{aligned}$$
Последовательно получаем, что все внедиагональные элементы равны нулю. □

Следствие 11.9. *Собственные подпространства нормального преобразования унитарного пространства попарно ортогональны.*

Следствие 11.10. *Для произвольной нормальной матрицы* $A \in \mathbb{C}^{n\times n}$ *существует ортогональная* $Q \in \mathbb{C}^{n\times n}$, *такая, что* $Q^{-1}AQ$ *диагональна (также говорят, что матрица* A *ортогонально подобна диагональной).*

Теорема 11.11. *Критерий нормального преобразования евклидова пространства Преобразование* φ *евклидова пространства является нормальным тогда и только тогда, когда существует ортонормированный базис пространства, в котором матрица преобразования* φ *блочно диагональная с блоками двух типов:*

1. *блоки первого порядка;*
2. *блоки второго порядка вида*

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}.$$

Доказательство. Достаточность Пусть в некотором ортонормированном базисе $e_1, \dots, e_n$ матрица преобразования φ имеет блочно-диагональный вид

$$[\varphi]_e = \operatorname{diag}\left(\lambda_1, \dots, \lambda_k, \begin{pmatrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{pmatrix}, \dots, \begin{pmatrix} \alpha_m & -\beta_m \\ \beta_m & \alpha_m \end{pmatrix}\right).$$

Тогда

$$[\varphi]_e = \operatorname{diag}\left(\overline{\lambda}_1, \dots, \overline{\lambda}_k, \begin{pmatrix} \overline{\alpha}_1 & \overline{\beta}_1 \\ -\overline{\beta}_1 & \overline{\alpha}_1 \end{pmatrix}, \dots, \begin{pmatrix} \overline{\alpha}_m & \overline{\beta}_m \\ -\overline{\beta}_m & \overline{\alpha}_m \end{pmatrix}\right).$$

Поэтому

$$[\varphi\varphi^*]_e = \left(|\lambda_1|^2, \dots, |\lambda_k|^2, \begin{pmatrix} |\alpha_1|^2 + |\beta_1|^2 & 0 \\ 0 & |\alpha_1|^2 + |\beta_1|^2 \end{pmatrix}, \dots, \begin{pmatrix} |\alpha_1|^2 + |\beta_1|^2 & 0 \\ 0 & |\alpha_1|^2 + |\beta_1|^2 \end{pmatrix}\right) = [\varphi^*\varphi]_e,$$

Следовательно, $\varphi\varphi^* = \varphi^*\varphi$.

Необходимость Пусть φ — нормальное преобразование евклидова пространства, $e_1, \dots, e_n$ — произвольный ортонормированный базис этого пространства. Рассмотрим преобразование $\Phi : \mathbb{C}^n \to \mathbb{C}^n$, заданное формулой $\Phi u = Au$, где $u \in \mathbb{C}^n$, $A = [\varphi]_e$. Пусть в пространстве $\mathbb{C}^n$ определено стандартное скалярное произведение. Так как матрица A — нормальная, то преобразование Φ — нормальное. Так как матрица A — вещественная, то, не нарушая общности можно считать, что система собственных значений имеет вид

$$\begin{aligned} \lambda_1, \dots, \lambda_t, \lambda_{t+1}, \dots, \lambda_k, \\ \overline{\lambda}_{t+1}, \dots, \overline{\lambda}_k, \end{aligned}$$

где $\lambda_i \in \mathbb{R}$ $(i = 1, \ldots, t)$, $\lambda_i \neq \lambda_j$ $(i \neq j)$.

Для каждого вещественного собственного значения λ_i $(i = 1, \ldots, t)$ построим ортонормированный базис $u_{i1}, \ldots, u_{is_i}$ собственного подпространства, составленный из векторов с вещественными компонентами. Так как матрица A и собственное значение λ_i — вещественные, то это возможно.

Для пары комплексно сопряженных собственных значений $\alpha_i \pm i\beta_i \notin \mathbb{R}$ $(i = t+1, \ldots, k)$ построим ортогональный базис собственного подпространства

$$u_{i1} = x_{i1} + iy_{i1}, \ldots, u_{is_i} = x_{is_i} + iy_{is_i} \tag{11.3}$$

и

$$v_{i1} = x_{i1} - iy_{i1}, \ldots, v_{is_i} = x_{is_i} - iy_{is_i} \tag{11.4}$$

соответственно, где $x_{ij}, y_{ij} \in \mathbb{R}^n$ $(i = t+1, \ldots, k;\ j = 1, \ldots, s_i)$.

Таким образом, имеем:

Собственные числа	Собственные векторы
λ_1 :	$u_{11}, \ldots, u_{1s_1}$
$\vdots$	$\vdots$
λ_t :	$u_{t1}, \ldots, u_{ts_t}$
$\alpha_{t+1} \pm i\beta_{t+1}$:	$x_{t+1\,1} \pm iy_{t+1\,1}, \ldots, x_{t+1\,s_{t+1}} \pm iy_{t+1\,s_{t+1}}$
$\vdots$	$\vdots$
$\alpha_k \pm i\beta_k$:	$x_{k1} \pm iy_{k1}, \ldots, x_{ks_k} \pm iy_{ks_k}$

1. Из формул (11.3,11.4) следует, что

$$x_{ij} = \frac{u_{ij} + v_{ij}}{2}, \quad y_{ij} = \frac{u_{ij} - v_{ij}}{2i},$$

поэтому

$$L(u_{ij}, v_{ij}) = L(x_{ij}, y_{ij}). \tag{11.5}$$

2а. Докажем, что $(x_{ij}, y_{ij}) = 0$. Имеем

$$0 = (u_{ij}, v_{ij}) = (x_{ij} + iy_{ij}, x_{ij} - iy_{ij}) = (x_{ij}, x_{ij}) + i(x_{ij}, y_{ij}) + i(y_{ij}, x_{ij}) - (y_{ij}, y_{ij}),$$

откуда, приравнивая нулю действительные и мнимые части, получаем $(x_{ij}, x_{ij}) = (y_{ij}, y_{ij})$ и $(x_{ij}, y_{ij}) = 0$.

2б. При $i \neq p$ или(и) $j \neq q$ имеем $L(u_{ij}, v_{ij}) \perp L(u_{pq}, v_{pq})$, откуда, в силу (11.5), получаем

$$(x_{ij}, x_{pq}) = (x_{ij}, y_{pq}) = (y_{ij}, y_{pq}) = 0.$$

3. Имеем

$$\Phi x_{ij} + i\Phi y_{ij} = \Phi u_{ij} = (\alpha_{ij} + i\beta_{ij})(x_{ij} + iy_{ij}) = (\alpha_{ij}x_{ij} - \beta_{ij}y_{ij}) + i(\alpha_{ij}y_{ij} + \beta_{ij}x_{ij}),$$

откуда, приравнивая действительные и мнимые части, получаем

$$\Phi x_{ij} = \alpha_{ij}x_{ij} - \beta_{ij}y_{ij}, \quad \Phi y_{ij} = \alpha_{ij}y_{ij} + \beta_{ij}x_{ij}.$$

4. Теперь легко видеть, что система f_{ij} ($i = 1, \dots, k$; $j = 1, \dots, s_i$), f'_{ij} ($i = t + 1, \dots, k$; $j = 1, \dots, s_i$), где

$$\begin{aligned} [f_{ij}]_e &= u_{ij} && (i = 1, \dots, t; j = 1, \dots, s_i), \\ [f_{ij}]_e &= x_{ij}/|x_{ij}| && (i = t+1, \dots, k; j = 1, \dots, s_i), \\ [f_{ij}]_e &= y_{ij}/|y_{ij}| && (i = t+1, \dots, k; j = 1, \dots, s_i), \end{aligned}$$

составляет искомый базис. □

11.4. Унитарные и ортогональные преобразования

Преобразование φ унитарного (соответственно евклидова) пространства называется *унитарным* (соответственно *ортогональным*), если

$$\varphi\varphi^* = \varphi^*\varphi = \varepsilon. \tag{11.6}$$

Легко видеть, что унитарное (ортогональное) преобразование является нормальным. Условие (11.6), очевидно, эквивалентно утверждению, что

$$(\varphi x, \varphi y) = (x, y) \tag{11.7}$$

для любых векторов x, y и поэтому *унитарные* (*ортогональные*) *преобразования — это в точности те преобразования, которые сохраняют*[1] *скалярное произведение*. Из (11.6) также следует, что для унитарного (ортогонального) преобразования φ обратное φ^{-1} существует и $\varphi^{-1} = \varphi^*$.

Матрица $A \in \mathbb{C}^{n\times n}$ называется *унитарной*, если $A^\top A = E$.

Теорема 11.12. *Преобразование φ унитарного* (*евклидова*) *пространства унитарно* (*ортогонально*) *тогда и только тогда, когда*

$$(\varphi x, \varphi x) = (x, x) \tag{11.8}$$

для любого вектора x.

Доказательство. Необходимость очевидна. Докажем достаточность. Для унитарного (евклидова) пространства легко проверить равенства

$$(x, y) = \frac{1}{4}\left((x + y, x + y) - (x - y, x - y) + i(x + iy, x + iy) - i(x - iy, x - iy)\right)$$

и

$$(x, y) = \frac{1}{4}\left((x + y, x + y) - (x - y, x - y)\right)$$

соответственно, откуда $(\varphi x, \varphi y) = (x, y)$. □

[1] в смысле формулы (11.7)

Замечание 11.13. Для случая евклидова геометрического пространства $\mathbf{V}_3$ можно дать следующее доказательство достаточности условий теоремы. Формула (11.8) означает, что преобразование сохраняет модули (т. е. длины) векторов, а значит расстояние между точками. Сохранение углов следует теперь из теоремы косинусов. Скалярное произведение выражается через длины векторов (которые не изменяются) и косинус угла между ними (который также остается постоянным).

Теорема 11.14. *Преобразование φ унитарного* (*евклидова*) *пространства унитарно* (*ортогонально*) *тогда и только тогда, когда образы векторов произвольного* (*какого-нибудь*) *ортонормированного базиса ортонормированы.*

Доказательство. Необходимость Пусть φ — унитарное (ортогональное) преобразование, $e_1, \ldots, e_n$ — ортонормированный базис. Тогда $(\varphi e_i, \varphi e_j) = (e_i, e_j)$, т. е. система $\varphi e_1, \ldots, \varphi e_n$ — ортонормирована.
Достаточность Пусть $e_1, \ldots, e_n$, $\varphi e_1, \ldots, \varphi e_n$ — две ортонормированные системы, $x = \alpha_1 e_1 + \ldots + \alpha_n e_n$, $y = \beta_1 e_1 + \ldots + \beta_n e_n$ — произвольные векторы. Тогда

$$\begin{aligned}(\varphi x, \varphi y) &= (\alpha_1 \varphi e_1 + \ldots + \alpha_n \varphi e_n, \beta_1 \varphi e_1 + \ldots + \beta_n \varphi e_n) = \alpha_1 \overline{\beta}_1 + \ldots + \alpha_n \overline{\beta}_n \\ &= (\alpha_1 e_1 + \ldots + \alpha_n e_n, \beta_1 e_1 + \ldots + \beta_n e_n) = (x, y).\end{aligned}$$

□

Теорема 11.15 (Свойства собственных чисел ортогонального (унитарного) преобразования)**.** *Все собственные числа ортогонального* (*унитарного*) *преобразования по абсолютной величине равны* 1.

Доказательство. Для собственного вектора x, такого, что $\varphi x = \lambda x$, имеем

$$(x, x) = (\varphi x, \varphi x) = (\lambda x, \lambda x) = \lambda \overline{\lambda} (x, x),$$

откуда $\lambda \overline{\lambda} = |\lambda|^2 = 1$. □

Теорема 11.16 (Критерий унитарного преобразования)**.** *Если все собственные числа нормального преобразования φ унитарного пространства по абсолютной величине равны* 1, *то φ — унитарное.*

Доказательство. Так как φ — нормальное, то существует ортонормированный базис $e_1, \ldots, e_n$, для которого

$$[\varphi]_e = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

откуда

$$[\varphi^*]_e = \operatorname{diag}(\overline{\lambda}_1, \ldots, \overline{\lambda}_n).$$

По условию $|\lambda_i| = 1$, поэтому $[\varphi]_e [\varphi^*]_e = E$, т. е. $\varphi \varphi^* = \varepsilon$. □

Теорема 11.17 (Критерий ортогонального преобразования). *Преобразование φ евклидова пространства является ортогональным тогда и только тогда, когда существует ортонормированный базис пространства, в котором матрица преобразования φ блочно диагональная с блоками двух типов*:

1. *блоки первого порядка вида*
$$(\pm 1);$$
2. *блоки второго порядка вида*
$$\begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}.$$

Доказательство. Теорема доказывается аналогично доказательству критерия нормального преобразования евклидова пространства. □

11.5. Эрмитовы (самосопряженные) и симметрические преобразования

Преобразование φ унитарного (соответственно евклидова) пространства называется *эрмитовым≡самосопряженным* (соответственно *симметрическим*), если $\varphi = \varphi^*$. Легко видеть, что эрмитово (симметрическое) преобразование является нормальным.

Теорема 11.18 (Свойства собственных чисел эрмитового преобразования). *Все собственные значения эрмитового преобразования вещественны.*

Доказательство. Для собственного вектора x, такого, что $\varphi x = \lambda x$, имеем
$$\lambda(x,x) = (\varphi x, x) = (x, \varphi x) = \overline{\lambda}(x,x),$$
откуда $\lambda = \overline{\lambda} \in \mathbb{R}$. □

Теорема 11.19 (Критерий эрмитового преобразования). *Если все собственные числа нормального преобразования φ унитарного пространства вещественны, то φ эрмитово.*

Доказательство. Так как φ — нормальное, то существует ортонормированный базис $e_1, \ldots, e_n$, для которого
$$[\varphi]_e = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$
Так как по условию $\lambda_i \in \mathbb{R}$, то $[\varphi]_e = [\varphi^*]_e$, т. е. $\varphi = \varphi^*$. □

Теорема 11.20 (Критерий симметрического преобразования). *Преобразование φ евклидова пространства тогда и только тогда является симметрическим, когда для него существует ортонормированный базис из собственных векторов.*

Доказательство. Ну, вообще-то, очевидно. □

Эрмитово (симметрическое) преобразование называется

- *положительным*, или *положительно определенным*, если $(x, \varphi x) > 0$ для любого ненулевого вектора x,
- *неотрицательным*, или *положительно полуопределенным*, если $(x, \varphi x) \geq 0$ для любого вектора x,
- *отрицательным*, или *отрицательно определенным*, если $(x, \varphi x) < 0$ для любого ненулевого вектора x,
- *неположительным*, или *отрицательно полуопределенным*, если $(x, \varphi x) \geq 0$ для любого вектора x.

Теорема 11.21 (Критерий положительного (неотрицательного) преобразования). *Эрмитово или симметрическое преобразование положительно определено* (*соответственно полуопределено*) *тогда и только тогда, когда все его собственные значения положительны* (*соответственно неотрицательны*).

Доказательство. Пусть $e_1, \ldots, e_n$ — ортонормированный базис из собственных векторов эрмитова или симметрического преобразования φ, причем $\varphi e_i = \lambda_i e_i$. Пусть $[x]_e = (x_1, \ldots, x_n)^\top$, $[y]_e = (y_1, \ldots, y_n)^\top$, тогда $(x, y) = x_1 \overline{y}_1 + \ldots + x_n \overline{y}_n$. и $(x, \varphi x) = \lambda_1 |x_1|^2 + \ldots + \lambda_n |x_n|^2$, откуда и следует утверждение теоремы. □

11.6. Косоэрмитово преобразование

Преобразование φ унитарного пространства называется *косоэрмитовым*, если $\varphi^* = -\varphi$.

Утверждение 11.22. *Преобразование φ унитарного пространства является косоэрмитовым тогда и только тогда, когда $\varphi = i\psi$ для некоторого эрмитова преобразования ψ.*

Доказательство. Необходимость Пусть $\psi = -i\varphi$, где ϕ — косоэрмитово преобразование, тогда

$$(\psi)^* = (-i\varphi)^* = i\varphi^* = -i\varphi = \psi,$$

таким образом, ψ — эрмитово. Легко видеть, что $\varphi = i\psi$.
Достаточность Достаточность проверяется непосредственно. □

11.7. Разложения преобразований

Произвольное комплексное число можно представить либо в алгебраической форме $\alpha + i\beta$, либо в тригонометрической $r(\cos\varphi + i \sin\varphi)$. Следующие две теоремы о разложении пробразований являются аналогами этих представлений.

Теорема 11.23 (Разложение преобразования в прямую сумму эрмитова и косо-эрмитова преобразований)**.** *Для любого преобразования φ унитарного пространства существуют единственные эрмитовы преобразования ψ_1, ψ_2, такие, что $\varphi = \psi_1 + i\psi_2$.*

Доказательство. Существование Пусть

$$\psi_1 = \frac{1}{2}(\varphi + \varphi^*), \quad \psi_2 = \frac{1}{2i}(\varphi - \varphi^*). \tag{11.9}$$

Легко проверить, что данные преобразования эрмитовы и $\varphi = \psi_1 + i\psi_2$.
Единственность Пусть $\varphi = \psi_1 + i\psi_2$ и преобразования ψ_1, ψ_2 эрмитовы. Тогда $\varphi^* = \psi_1 - i\psi_2$, откуда следуют формулы (11.9). По этим формулам ψ_1, ψ_2 определяются однозначно. □

Теорема 11.24 (Полярное разложении)**.** *Для любого преобразования φ унитарного* (*соответственно евклидова*) *пространства существуют неотрицательное самосопряженное* (*соответственно симметрическое*) *преобразование ψ и унитарное* (*соответственно ортогональное*) *преобразование η, такие, что $\varphi = \psi\eta$. Если φ — невырождено, то преобразования ψ, η определяются единственным образом.*

Доказательство. Во-первых, исследуем преобразование $\varphi\varphi^*$. Докажем, что оно самосопряженное (симметрическое) и неотрицательное. Действительно,

$$(\varphi\varphi^*)^* = (\varphi^*)^* \varphi^* = \varphi\varphi^*.$$

Далее,

$$(x, \varphi\varphi^* x) = (\varphi x, \varphi x) \geq 0.$$

для любого вектора x.

Во-вторых, докажем что если φ — невырождено, то преобразования ψ, η определяются единственным образом. Пусть $\varphi = \psi\eta$, $\psi^* = \psi$, $\psi \geq 0$, $\eta^* = \eta^{-1}$. Тогда $\varphi\varphi^* = \psi\eta(\psi\eta)^* = \psi^2$, откуда положительное самосопряженное преобразование ψ определяется единственным образом. Теперь $\eta = \psi^{-1}\varphi$.

Наконец, покажем, как построить преобразования ψ и η[2]. Пусть $e_1, \ldots, e_n$ — ортонормированный базис из собственных векторов преобразования $\varphi\varphi^* \geq 0$. Имеем

$$\varphi\varphi^* e_i = k_i^2 e_i.$$

Не нарушая общности, будем считать, что

$$k_i > 0 \quad (i = 1, \ldots, m), \qquad k_i = 0 \quad (i = m+1, \ldots, n).$$

Пусть[3]

$$\psi e_i = k_i e_i \quad (i = 1, \ldots, n). \tag{$\star$}$$

[2]Из предыдущего абзаца следует способ нахождения $\psi = \sqrt{\varphi\varphi^*}$ и $\eta = \psi^{-1}\varphi$ по произвольному *невырожденному* преобразованию φ. Приведенный далее способ уже не требует невырожденности φ.

[3]Таким образом, как и для *невырожденного* преобразования имеем $\psi = \sqrt{\varphi\varphi^*}$. Далее символами $\star$ отмечены основные формулы, используемые в *алгоритме* нахождения полярного разложения.

Положим

$$w_i = \frac{\varphi^* e_i}{k_i} \quad (i = 1, \ldots, m) \tag{$\star\star$}$$

и докажем, что векторы $w_1, \ldots, w_m$ образуют ортонормированный базис пространства $W = L(w_1, \ldots, w_m)$. Действительно,

$$k_i k_j (w_i, w_j) = (\varphi^* e_i, \varphi^* e_j) = (e_i, \varphi\varphi^* e_j) = k_j^2 (e_i, e_j).$$

Пусть

$$\eta w_i = e_i \quad (i = 1, \ldots, n). \tag{$\star\star\star$}$$

Легко видеть, что ψ — неотрицательное самосопряженное (симметрическое), η — унитарное (ортогональное). Проверим, что $\varphi = \psi\eta$. Имеем

Для $i = 1, \ldots, m$: $\quad \psi\eta w_i = \psi e_i = k_i e_i \quad$ и $\quad \varphi w_i = \varphi\varphi^* e_i / k_i = k_i e_i$;

для $i = m+1, \ldots, n$: $\quad \psi\eta w_i = \psi e_i = 0 \quad$ и $\quad (\varphi w_i, \varphi w_i) = (w_i, \underbrace{\varphi^* \varphi w_i}_{\in \operatorname{Im} \varphi^* = W}) = 0.$

$\square$

Замечание 11.25. Применим доказанную теорему для преобразования φ^*, тогда $\varphi^* = \psi\eta$, откуда $\varphi = \eta^*\psi$, причем η^* — унитарное (ортогональное), ψ — самосопряженное (симметрическое) неотрицательное преобразование.

Глава 12

Кривые и поверхности второго порядка

см. пособие В.Н.Шевченко «Кривые и поверхности второго порядка»

Глава 13

Матрицы над евклидовым кольцом

13.1. Нормальная диагональная форма

В данной главе везде предполагается, что K — некоторое евклидово кольцо. Рассмотрим кольцо $K^{m\times n}$ матриц размера $m\times n$ с элементами из K. Везде далее мы будем обращаться к двум важным примерам таких колец:

1. Кольцо $F[\lambda]^{m\times n}$, где F — некоторое поле. Элементы этого кольца называются *многочленными матрицами*, или *λ-матрицами* над полем F.

2. Кольцо $\mathbb{Z}^{m\times n}$ целочисленных матриц.

Элементарными преобразованиями таких матриц назовем преобразования следующих типов:

1. перестановка i-й и j-й строк;

2. умножение i-й строки на обратимый элемент из K (т. е. для кольца λ-матриц возможно умножение на ненулевой скаляр, для кольца целочисленных матриц — умножение на ± 1);

3. прибавление к i-й строке j-й строки, умноженной на произвольный элемент из K,

и аналогичные преобразования столбцов.

Будем говорить, что матрица $A\in K^{m\times n}$ эквивалентна матрице $B\in K^{m\times n}$ и писать $A\sim B$, если от A к B можно перейти при помощи цепочки элементарных преобразований. Легко видеть, что отношение эквивалентности обладает тремя основными свойствами: рефлексивностью, симметричностью и транзитивностью. Отсюда следует, что множество всех матриц из $K^{m\times n}$ разбивается на непересекающиеся классы эквивалентных матриц. Возникает вопрос о канонической матрице, характеризующей данный класс.

Канонической или *нормальной диагональной* матрицей называется матрица $\operatorname{diag}(d_1,\ldots,d_r,d_{r+1},\ldots,d_p)\in K^{m\times n}$, где $p=\min\{m,n\}$, в которой $d_j \,\vdots\, d_{j-1}$ $(j=2,3,\ldots,r)$, $d_j=0$ $(j=r+1,\ldots,p)$.

Теорема 13.1. *Для любой матрицы* $A \in K^{m\times n}$ *существует эквивалентная ей нормальная диагональная матрица* D, *называемая* канонической *или* нормальной диагональной формой *матрицы* A. *Диагональные элементы нормальной диагональной формы, называемые* инвариантными множителями, *определяются единственным образом с точностью до умножения на обратимые элементы из* K.

Пример 13.2. Найдем нормальную диагональную форму λ-матрицы

$$\begin{pmatrix} -\lambda+1 & 2\lambda^3-\lambda^2-2\lambda-1 & \lambda^3+2\lambda+3 & -\lambda-1 \\ -2\lambda+5 & -2\lambda^2+8\lambda+3 & -2\lambda-2 & \lambda+1 \\ -2\lambda^2+5\lambda+1 & -2\lambda^3+8\lambda^2+4\lambda & -2\lambda^2-2\lambda & \lambda^2+\lambda \\ \lambda-1 & -2\lambda^3+\lambda^2+5\lambda+4 & -\lambda^3-4\lambda-5 & 2\lambda+2 \end{pmatrix}$$

над полем комплексных чисел (или любым другим числовым полем). Из третьей строки вычтем вторую, умноженную на λ, и полученную строку переставим с первой:

$$\begin{pmatrix} 1 & \lambda & 0 & 0 \\ -2\lambda+5 & -2\lambda^2+8\lambda+3 & -2\lambda-2 & \lambda+1 \\ -\lambda+1 & 2\lambda^3-\lambda^2-2\lambda-1 & \lambda^3+2\lambda+3 & -\lambda-1 \\ \lambda-1 & -2\lambda^3+\lambda^2+5\lambda+4 & -\lambda^3-4\lambda-5 & 2\lambda+2 \end{pmatrix}.$$

Из второй, третьей и четвертой строки вычитаем первую, умноженную на $-2\lambda+5$, $-\lambda+1$ и $\lambda-1$ соответственно. Далее полученный первый столбец, умноженный на λ, вычитаем из четвертого столбца:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3\lambda+3 & -2\lambda-2 & \lambda+1 \\ 0 & 2\lambda^3-3\lambda-1 & \lambda^3+2\lambda+3 & -\lambda-1 \\ 0 & -2\lambda^3+6\lambda+4 & -\lambda^3-4\lambda-5 & 2\lambda+2 \end{pmatrix}.$$

Переставляем второй и четвертый столбцы. К третьей строке прибавляем вторую. Из четвертой вычитаем вторую, умноженную на 2. После этого к третьему столбцу прибавими второй, умноженный на 2, и из четвертого вычтем второй, умноженный на 3:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & \lambda^3+1 & 2\lambda^3+2 \\ 0 & 0 & -\lambda^3-1 & -2\lambda^3-2 \end{pmatrix}.$$

К четвертой строке прибавляем третью. Далее из четвертого столбца вычитаем третий, умноженный на 2:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda+1 & 0 & 0 \\ 0 & 0 & \lambda^3+1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Это нормальная диагональная матрица.

Пример 13.3. Найдем нормальную диагональную форму целочисленной матрицы

$$\begin{pmatrix} 84 & 24 & -72 & -72 & 72 \\ -28 & 8 & 60 & 60 & -76 \\ 488 & 224 & -240 & -240 & 164 \end{pmatrix}.$$

Из первого столбца вычтем второй, умноженный на 2. Из четвертого столбца вычтем третий. К пятому столбцу прибавим третий. Получим:

$$\begin{pmatrix} 36 & 24 & -72 & 0 & 0 \\ -44 & 8 & 60 & 0 & -16 \\ 40 & 224 & -240 & 0 & -76 \end{pmatrix}.$$

Из первой и третьей строк вычтем вторую, умноженную на 3 и 28 соответственно. Из первого, третьего и пятого столбцов вычтем второй, умноженный на -5, 7 и -2 соответственно. Получим:

$$\begin{pmatrix} 168 & 0 & -252 & 0 & 48 \\ -4 & 8 & 4 & 0 & 0 \\ 1272 & 0 & -1920 & 0 & 372 \end{pmatrix}.$$

К первой и третьей строкам прибавим вторую, умноженную на 63 и 480 соответственно. Прибавим к первому столбцу третий. Из второго отнимим третий, умноженный на 2. После этого надлежащим образом переставим строки и столбцы:

$$\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 504 & -84 & 48 & 0 \\ 0 & 3840 & -648 & 372 & 0 \end{pmatrix}.$$

Из третьей строки вычтем вторую, умноженную на 7. Из второго столбца вычтем четвертый, умноженный на 10. К третьему столбцу прибавим четвертый, умноженный на 2. Получим:

$$\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 24 & 12 & 48 & 0 \\ 0 & -48 & 12 & 36 & 0 \end{pmatrix}.$$

Из третьей строки вычтем вторую. Из второго и четвертого столбцов вычтем третий, умноженный на 2 и 4 соответственно. Получим:

$$\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \\ 0 & -72 & 0 & -12 & 0 \end{pmatrix}.$$

Вычтем из второго столбца четвертый, умноженный на 6. Далее умножим третью строку (или четвертый столбец) на -1 и после надлежащей перестановки столбцов получим нормальную диагональную матрицу:

$$\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \end{pmatrix}.$$

На матрицы из $K^{m\times n}$ можно распространить понятие минора и ранга. При элементарных преобразованиях ранг матрицы не изменяется.

Наибольший общий делитель всех миноров k-го порядка матрицы $A \in K^{m\times n}$ обозначим Δ_k $(k = 1, 2, \ldots, \min\{m, n\})$.

Утверждение 13.4. *При элементарных преобразованиях матрицы величины Δ_k не изменяются.*

Утверждение 13.5. *Для инвариантных множителей d_j матрицы A справедливы формулы*:

$$d_j = \begin{cases} \Delta_1 & \text{при } j = 1, \\ \dfrac{\Delta_j}{\Delta_{j-1}} & \text{при } j = 2, \ldots, r, \\ 0 & \text{при } j = r+1, \ldots, \min\{m, n\}, \end{cases}$$

где r — ранг матрицы A.

Пример 13.6.

$$\begin{pmatrix} \lambda+4 & -3 & -3 \\ 0 & \lambda-2 & 0 \\ 6 & -3 & \lambda-5 \end{pmatrix}.$$

Так как матрица содержит ненулевые константы, то $\Delta_1 = 1$. Все миноры второго порядка, содержащие вторую строку делятся на $\lambda - 2$. Вычислим остальные миноры второго порядка:

$$\begin{vmatrix} \lambda+4 & -3 \\ 6 & -3 \end{vmatrix} = -3\lambda + 6, \quad \begin{vmatrix} \lambda+4 & -3 \\ 6 & \lambda-5 \end{vmatrix} = \lambda^2 - \lambda - 2 = (\lambda+1)(\lambda-2),$$

$$\begin{vmatrix} -3 & -3 \\ -3 & \lambda-5 \end{vmatrix} = -3\lambda + 6.$$

Получаем $\Delta_2 = \lambda - 2$. Разлагая определитель исходной матрицы по второй строке, находим $\Delta_3 = (\lambda+1)(\lambda-2)^2$. Таким образом, инвариантные множители равны

$$d_1 = \Delta_1 = 1, \qquad d_2 = \frac{\Delta_2}{\Delta_1} = \lambda - 2, \qquad d_3 = \frac{\Delta_3}{\Delta_2} = (\lambda+1)(\lambda-2).$$

Нормальная диагональная форма имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (\lambda+1)(\lambda-2) \end{pmatrix}.$$

Пример 13.7. Найдем нормальную диагональную форму целочисленной матрицы

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ -1 & 5 & 8 \end{pmatrix}$$

Имеем $\Delta_1 = 1$. Легко видеть, что каждый минор второго порядка, содержащий вторую строку или третий столбец, делится на 2. Отстается единственный минор второго порядка:

$$\begin{vmatrix} 1 & 1 \\ -1 & 5 \end{vmatrix} = 6,$$

поэтому $\Delta_2 = 2$. Вычисляем определитель всей матрицы:

$$\begin{vmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ -1 & 5 & 8 \end{vmatrix} = 8,$$

откуда $\Delta_3 = 8$. Получаем

$$d_1 = \Delta_1 = 1, \qquad d_2 = \frac{\Delta_2}{\Delta_1} = 2, \qquad d_3 = \frac{\Delta_3}{\Delta_2} = 4.$$

Нормальная диагональная форма имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Следствие 13.8. *Матрицы A и B из $K^{m\times n}$ эквивалентны тогда и только тогда, когда их нормальные диагональные формы совпадают (с точностью до умножения диагональных элементов на обратимые элементы из K).*

Матрица $P \in K^{n\times n}$ называется *унимодулярной*, если в $K^{n\times n}$ существует обратная к ней матрца.

Утверждение 13.9. *Пусть $P \in K^{n\times n}$. Для того, чтобы P была унимодулярной необходимо и достаточно, чтобы* $\det P \in K$ *и* $\det P \neq 0$.

Теорема 13.10. *Для того, чтобы матрицы A и B из $K^{m\times n}$ были эквивалентны необходимо и достаточно, чтобы существовали унимодулярные матрицы P из $K^{m\times m}$ и Q из $K^{n\times n}$, такие, что $B = PAQ$.*

Для нахождения унимодулярных матриц P и Q, таких, что $B = PAQ$ достаточно приписать снизу и справа от матрицы A по единичной матрице (соответствующего порядка) и элементарными преобразованиями привести A к B, при этом все преобразования со строками производить также с единичной матрицей, приписанной справа, а преобразования со столбцами — с единичной матрицей, приписанной снизу. Чаще удобнее приводить матрицу A не к виду B, а каждую из матриц A, B привести к нормальной диагональной форме D. Если $D = P_1AQ_1 = P_2AQ_2$, где P_1, P_2, Q_1, Q_2 — унимодулярные матрицы, то $B = (P_2^{-1}P_1)A(Q_1Q_2^{-1})$, причем матрицы $P_2^{-1}P_1$, $Q_1Q_2^{-1}$ — унимодулярные.

Пример 13.11. Приведем λ-матрицу
$$A = \begin{pmatrix} \lambda & \lambda+1 \\ 0 & \lambda \end{pmatrix}$$
к нормальной диагональной форме и найдем P, Q, такие, что $D = PAQ$.

Выпишем матрицу
$$\left(\begin{array}{cc|cc} \lambda & \lambda+1 & 1 & 0 \\ 0 & \lambda & 0 & 1 \\ \hline 1 & 0 & & \\ 0 & 1 & & \end{array}\right).$$
Из ее второго столбца вычтем первый:
$$\left(\begin{array}{cc|cc} \lambda & 1 & 1 & 0 \\ 0 & \lambda & 0 & 1 \\ \hline 1 & -1 & & \\ 0 & 1 & & \end{array}\right).$$
Теперь вычтем из первого столбца второй, умноженный на λ, и вычтем из второй строки первую, умноженную на λ. Получим:
$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ -\lambda^2 & 0 & -\lambda & 1 \\ \hline \lambda+1 & -1 & & \\ -\lambda & 1 & & \end{array}\right).$$
Умножаем первый столбец на -1 и переставляем его со вторым:
$$\left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & \lambda^2 & -\lambda & 1 \\ \hline -1 & -\lambda-1 & & \\ 1 & \lambda & & \end{array}\right).$$
Мы нашли
$$A = \begin{pmatrix} 1 & 0 \\ 0 & \lambda^2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} -1 & -\lambda-1 \\ 1 & \lambda \end{pmatrix}.$$
Матрицы P и Q определяются не однозначно.

13.2. Связь подобия числовых матриц с эквивалентностью λ-матриц

В данном разделе рассматриваются только λ-матрицы, т. е. элементы кольца $F[\lambda]^{m\times n}$, где F — поле.

Заметим, что кольцо $F[\lambda]^{m\times n}$ *многочленных матриц* (λ-матриц), изоморфно кольцу $F^{m\times n}[\lambda]$ *матричных многочленов*, т. е. многочленов, коэффициентами которых являются матрицы из $F^{m\times n}$. *Естественный изоморфизм* задается простым правилом, которое мы покажем на примере.

Пример 13.12. Многочленной матрице

$$\begin{pmatrix} \lambda^2+2\lambda-2 & -\lambda^2+5\lambda-1 & \lambda^2+3\lambda+2 \\ -\lambda^2+3\lambda+1 & 2\lambda^2-2\lambda+5 & 3\lambda^2+5\lambda+2 \\ \lambda^2-2\lambda-1 & \lambda^2-4\lambda-3 & 2\lambda^2-2\lambda-1 \end{pmatrix}$$

соответствует матричный многочлен

$$\begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}\cdot\lambda^2+\begin{pmatrix} 2 & 5 & 3 \\ 3 & -2 & 5 \\ -2 & -4 & -2 \end{pmatrix}\cdot\lambda+\begin{pmatrix} -2 & -1 & 2 \\ 1 & 5 & 2 \\ -1 & -3 & -1 \end{pmatrix}.$$

Естественный изоморфизм из $F[\lambda]^{m\times n}$ в $F^{m\times n}[\lambda]$ позволяет нам отождествлять λ-матрицы и соответствующие ей матричные многочлены.

Пусть A — матрица из $F^{n\times n}$, тогда $A-\lambda E$ называется ее *характеристической* матрицей.

Теорема 13.13. *Пусть A, B — матрицы из $F^{n\times n}$. Для того, чтобы A и B были подобны необходимо и достаточно, чтобы их характеристические матрицы $A-\lambda E$ и $B-\lambda E$ были эквивалентны.*

Более того, если $B-\lambda E = P(A-\lambda E)Q$, где P, Q — унимодулярные λ-матрицы, то $B = LAR$ и $L = R^{-1}$, где L — остаток от деления P на $B-\lambda E$ слева, а R — остаток от деления Q на $B-\lambda E$ справа. Таким образом, R — трансформирующая матрица.

Утверждение 13.14. *Имеет место эквивалентность*

$$J_n(\lambda_0)-\lambda E = \begin{pmatrix} \lambda_0-\lambda & 1 & \dots & 0 & 0 \\ 0 & \lambda_0-\lambda & \dots & 0 & 0 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \dots & \lambda_0-\lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda_0-\lambda \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & (\lambda-\lambda_0)^n \end{pmatrix}$$

Таким образом, все инвариантные множители характеристической матрицы жордановой клетки равны 1, *кроме последнего, равного* $(\lambda-\lambda_0)^n$.

Утверждение 13.15. *Если многочлены* $f_1(\lambda), f_2(\lambda), \ldots, f_n(\lambda)$ *из* $F[\lambda]$ *попарно взаимно просты, то*

$$\begin{pmatrix} f_1(\lambda) & 0 & \ldots & 0 \\ 0 & f_2(\lambda) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & f_n(\lambda) \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & f_1(\lambda)\ldots f_n(\lambda) \end{pmatrix}.$$

Утверждение 13.16. *Пусть* $J = \operatorname{diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \ldots, J_{k_s}(\lambda_s))$, *тогда*

$$J-\lambda E \sim \begin{pmatrix} 1 &&&&&&& \\ & 1 &&&&&& \\ && \ddots &&&&& \\ &&& 1 &&&& \\ &&&& (\lambda-\lambda_1)^{k_1} &&& \\ &&&&& (\lambda-\lambda_2)^{k_2} && \\ &&&&&& \ddots & \\ &&&&&&& (\lambda-\lambda_s)^{k_s} \end{pmatrix}$$

Матрица из предыдущего утверждения хотя и диагональна, но в общем случае не является нормальной диагональной.

Рассмотрим теперь матрицу $J-\lambda E = \operatorname{diag}(J_1-\lambda E, J_2-\lambda E, \ldots, J_s-\lambda E)$, где $J_1, \ldots, J_s$ — жордановы клетки. Пусть $\lambda_1, \ldots, \lambda_s$ — все собственные значения матрицы J, причем $\lambda_i \neq \lambda_j$ при $i \neq j$. Пусть к собственному числу λ_i относится q_i жордановых клеток и пусть порядки этих клеток, расположенные в неубывающем порядке, суть $k_{i1} \le k_{i2} \le \ldots \le k_{iq_i}$. Обозначим

$$e_{n-j+1}(\lambda) = \prod_{i=1}^{k} (\lambda-\lambda_i)^{k_{ij}} \quad (j = 1, 2, \ldots, n),$$

полагая, что если для некоторого $q_i < j$, то соответствующий множитель $(\lambda-\lambda_i)^{k_{ij}}$ равен 1.

Утверждение 13.17. $J-\lambda E \sim \operatorname{diag}(e_1(\lambda_1), \ldots, e_n(\lambda_n))$, *причем матрица, стоящая справа, имеет нормальный диагональный вид.*

Пример 13.18. Пусть

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

Тогда

$$J - \lambda E \sim \mathrm{diag}(1, 1, 1, 1, \lambda - 2, \lambda - 2, (\lambda - 2)^3, (\lambda - 3)^2, (\lambda - 3)^2) \sim$$
$$\sim \mathrm{diag}(1, 1, 1, 1, 1, 1, \lambda - 2, (\lambda - 2)(\lambda - 3)^2, (\lambda - 2)^3(\lambda - 3)^2).$$

Последняя матрица является нормальной диагональной формой матрицы J.

Пусть

$$d_i = e_{i1}(\lambda)e_{i2}(\lambda)\dots e_{ip_i}(\lambda) \quad (i = 1, 2, \dots, n)$$

— разложение i-го инвариантного множителя на неприводимые над полем F множители, причем $e_{ij} \neq e_{ik}$ при $j \neq 1$. Система

$$e_{ij}(\lambda) \qquad (i = 1, 2, \dots, n;\ j = 1, 2, \dots p_i)$$

называется системой *элементарных делителей* матрицы.

Теорема 13.19. *Матрица A из $F^{n\times n}$ тогда и только тогда подобна жордановой матрице, когда система элементарных делителей матрицы $A - \lambda E$ состоит только из линейных многочленов*

$$(\lambda - \lambda_i)^{t_j} \qquad (i = 1, 2, \dots, n;\ j = 1, 2, \dots p_i).$$

В этом случае жорданова матрица составлена из жордановых клеток, которые по элементарным делителям можно найти следующим образом. Каждому элементарному делителю $(\lambda - \lambda_i)^{t_j}$ соответствует клетка t_j-го порядка с λ_j на диагонали. Жорданова форма определяется единственным с точностью до перестановки жордановых клеток образом.

Следствие 13.20. *Любая матрица над алгебраически замкнутым полем (в частности, над полем комплексных чисел) подобна жордановой.*

Пример 13.21. Найдем жорданову форму и соответствующую трансформирующую матрицу для матрицы

$$A = \begin{pmatrix} 2 & 17 & -68 \\ 0 & -2 & 16 \\ 0 & -1 & 6 \end{pmatrix}.$$

Составляем матрицу $A - \lambda E$ и приписываем к ней снизу единичную:

$$A - \lambda E = \left(\begin{array}{ccc} 2 - \lambda & 17 & -68 \\ 0 & -2 - \lambda & 16 \\ 0 & -1 & 6 - \lambda \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

С помощью преобразований строк и столбцов матрицу $A-\lambda E$ приведем к нормальной диагональной форме. Все преобразованиями со столбцами будем проделывать так же с матрицей, стоящей снизу. Умножим второй столбец на -1. Далее вычтем этот столбец, умноженный на $6-\lambda$, из третьего столбца:

$$\left(\begin{array}{ccc} 2-\lambda & -17 & 34-17\lambda \\ 0 & 2+\lambda & (\lambda-2)^2 \\ 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ 0 & -1 & 6-\lambda \\ 0 & 0 & 1 \end{array}\right)$$

Прибавляем третью строку, умноженную на 17, к первой. Вычитаем третью строку, умноженную на $\lambda+2$, из второй. В результате получим:

$$\left(\begin{array}{ccc} 2-\lambda & 0 & 34-17\lambda \\ 0 & 0 & (\lambda-2)^2 \\ 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ 0 & -1 & 6-\lambda \\ 0 & 0 & 1 \end{array}\right)$$

Переставляя строки и столбцы надлежащим образом, получим:

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 2-\lambda & 34-17\lambda \\ 0 & 0 & (\lambda-2)^2 \\ \hline 0 & 1 & 0 \\ -1 & 0 & 6-\lambda \\ 0 & 0 & 1 \end{array}\right)$$

Вычитаем из третьего столбца второй, умноженный на 17:

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & (\lambda-2)^2 \\ \hline 0 & 1 & -17 \\ -1 & 0 & 6-\lambda \\ 0 & 0 & 1 \end{array}\right)$$

Меняем знак у второго столбца:

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (\lambda-2)^2 \\ \hline 0 & -1 & -17 \\ -1 & 0 & 6-\lambda \\ 0 & 0 & 1 \end{array}\right)$$

Матрица $A-\lambda E$ приведена к нормальной диагональной форме. Элементарными делителями являются $\lambda-2$ и $(\lambda-2)^2$, поэтому A подобна жордановой матрице

$$J=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Чтобы найти трансформирующую матрицу Q (т. е. матрицу, для которой $J = Q^{-1}AQ$) элементарными преобразованиями приведем $J - \lambda E$ к нормальной диагональной форме. Составляем матрицу

$$\left(\begin{array}{ccc} 2-\lambda & 0 & 0 \\ 0 & 2-\lambda & 1 \\ 0 & 0 & 2-\lambda \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

Вычитаем из второго столбца третий, умноженный на $2 - \lambda$:

$$\left(\begin{array}{ccc} 2-\lambda & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -(2-\lambda)^2 & 2-\lambda \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda-2 & 1 \end{array}\right)$$

Вычитаем из третьей строки вторую, умноженную на $2 - \lambda$:

$$\left(\begin{array}{ccc} 2-\lambda & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -(2-\lambda)^2 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda-2 & 1 \end{array}\right)$$

Преставляя строки и столбцы надлежащим образом, получаем:

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (2-\lambda)^2 \\ \hline 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 2-\lambda \end{array}\right)$$

Итак,

$$Q_1(\lambda) = \begin{pmatrix} 0 & -1 & -17 \\ -1 & 0 & 6-\lambda \\ 0 & 0 & 1 \end{pmatrix}, \qquad Q_2(\lambda) = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 2-\lambda \end{pmatrix}.$$

Вычислим

$$Q_0(\lambda) = Q_1(\lambda)Q_2(\lambda)^{-1} = \begin{pmatrix} 1 & 17 & 0 \\ 0 & 2\lambda-8 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Найдем остаток Q при делении $Q_0(\lambda)$ на $J - \lambda E$ справа:

$$Q = Q_0(J) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot J + \begin{pmatrix} 1 & 17 & 0 \\ 0 & -8 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 17 & 0 \\ 0 & -4 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Матрица Q и есть искомая трансформирующая матрица и $J = Q^{-1}AQ$.

Матрицей *естественной формы*, или матрицей, имеющей *нормальную естественную форму*, называется клеточно-диагональная матрица $B = \mathrm{diag}(B_1, B_2, \ldots, B_s) \in F^{n\times n}$, где $B_1, \ldots, B_s$ являются матрицами, сопровождающими многочлены $f_1(\lambda), \ldots, f_s(\lambda)$ соответственно, причем $f_i(\lambda) \vdots f_{i-1}(\lambda)$ $(i = 2, 3, \ldots, s)$.

Утверждение 13.22. *Первые $n - s$ инвариантных множителя матрицы естественной формы являются 1. Остальные инвариантные множители равны $f_1(\lambda), \ldots, f_s(\lambda)$.*

Теорема 13.23. *Любая матрица A из $F^{n\times n}$ подобна над F единственной матрице B естественной формы, называемой* нормальной естественной формой *матрицы A, причем $B = \mathrm{diag}(B_1, B_2, \ldots, B_s) \in F^{n\times n}$, где $B_1, \ldots, B_s$ — матрицы, сопровождающие инвариантные множители $f_1(\lambda), \ldots, f_s(\lambda)$ матрицы. Две матрицы из $F^{n\times n}$ подобны над F тогда и только тогда, когда их естественные формы совпадают.*

Квазиестественной матрицей называется клеточно-диагональная матрица $B = \mathrm{diag}(B_1, B_2, \ldots, B_s) \in F^{n\times n}$, где $B_1, \ldots, B_s$ являются матрицами, сопровождающими многочлены вида

$$\Big(f_1(\lambda)\Big)^{k_1}, \Big(f_2(\lambda)\Big)^{k_2}, \ldots, \Big(f_s(\lambda)\Big)^{k_s},$$

где $f_1(\lambda), \ldots f_s(\lambda)$ неприводимы над полем F и имеют старшие коэффициенты, равные единице.

Утверждение 13.24. *Элементарные делители квазиестественной матрицы являются $\Big(f_1(\lambda)\Big)^{k_1}, \Big(f_2(\lambda)\Big)^{k_2}, \ldots, \Big(f_s(\lambda)\Big)^{k_s}$.*

Теорема 13.25. *Любая матрица A из $F^{n\times n}$ подобна над F квазиестественной матрице B, определяемой однозначно с точностью до порядка следования диагональных клеток, называемой* нормальной квазиестественной формой *матрицы A, причем $B = \mathrm{diag}(B_1, B_2, \ldots, B_s) \in F^{n\times n}$, где $B_1, \ldots, B_s$ — матрицы, сопровождающие элементарные делители $\Big(f_1(\lambda)\Big)^{k_1}, \Big(f_2(\lambda)\Big)^{k_2}, \ldots, \Big(f_s(\lambda)\Big)^{k_s}$ матрицы. Две матрицы из $F^{n\times n}$ подобны над F тогда и только тогда, когда их нормальные квазиестественные формы совпадают.*

Пример 13.26. Инвариантные множители матрицы J из примера 13.18 равны

$$1,\ 1,\ 1,\ 1,\ 1,\ 1,\ \lambda - 2,\ (\lambda - 2)(\lambda - 3)^2 = \lambda^3 - 8\lambda^2 + 21\lambda - 18,$$

$$(\lambda - 2)^3(\lambda - 3)^2 = x^5 - 12\lambda^4 + 57\lambda^3 - 134\lambda^2 + 156\lambda - 72,$$

поэтому ее естественная форма равна

$$\begin{pmatrix}
2 &&&&&&&& \\
& 8 & -21 & 18 &&&&& \\
& 1 & 0 & 0 &&&&& \\
& 0 & 1 & 0 &&&&& \\
&&&& 12 & -57 & 134 & -156 & 72 \\
&&&& 1 & 0 & 0 & 0 & 0 \\
&&&& 0 & 1 & 0 & 0 & 0 \\
&&&& 0 & 0 & 1 & 0 & 0 \\
&&&& 0 & 0 & 0 & 1 & 0
\end{pmatrix}.$$

Элементарные делители матрицы J равны

$$\lambda - 2,\ \lambda - 2,\ (\lambda - 2)^3 = \lambda^3 - 6\lambda^2 + 12\lambda - 8,\ (\lambda - 3)^2 = \lambda^2 - 6\lambda + 9,\ (\lambda - 3)^2 = \lambda^2 - 6\lambda + 9,$$

поэтому квазиестественная форма равна

$$\begin{pmatrix} 2 &&&&&&&& \\ & 2 &&&&&&& \\ && 6 & -12 & 8 &&&& \\ && 1 & 0 & 0 &&&& \\ && 0 & 1 & 0 &&&& \\ &&&&& 6 & -9 && \\ &&&&& 1 & 0 && \\ &&&&&&& 6 & -9 \\ &&&&&&& 1 & 0 \end{pmatrix}.$$

Пример 13.27. Инвариантные множители матрицы A из примера 13.21 равны

$$1, \quad \lambda-2, \quad (\lambda-2)^2 = \lambda^2 - 4\lambda + 4,$$

поэтому ее естественная форма равна

$$B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & -4 \\ 0 & 1 & 0 \end{pmatrix}.$$

В данном примере квазиестественная форма совпадает с естественной, так как элементарные делители ненулевой степени совпадают с инвариантными множителями.

Найдем матрицу Q, такую, что $B = Q^{-1}AQ$. Для этого приведем $B - \lambda E$ к нормальной диагональной форме. Запишем матрицу

$$\left(\begin{array}{ccc} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & -4 \\ 0 & 1 & -\lambda \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right).$$

Прибавим к третьему столбцу второй, умноженный на λ. Вычтем из второй строки третью, умноженную на $4-\lambda$. Получим:

$$\left(\begin{array}{ccc} 2-\lambda & 0 & 0 \\ 0 & 0 & -\lambda^2+4\lambda-4 \\ 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{array}\right).$$

Переставляя строки и столбцы и меняя знак у двух строк, получаем:

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \lambda-2 & 0 \\ 0 & 0 & (\lambda-2)^4 \\ \hline 0 & 1 & 0 \\ 1 & 0 & \lambda \\ 0 & 0 & 1 \end{array}\right).$$

Итак,

$$Q_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \lambda \\ 0 & 0 & 1 \end{pmatrix}, \quad Q_0 = Q_1 Q_3^{-1} = \begin{pmatrix} -1 & 0 & -17 \\ 0 & -1 & 6 \\ 0 & 0 & 1 \end{pmatrix}$$

(матрица Q_1 вычислена в примере 13.21). Далее необходимо найти остаток Q при делении Q_0 на $J - \lambda E$, но так как матрица Q_0 — скалярная, то $Q = Q_0$. Итак, Q_0 — искомая трансформирующая матрица.

Теорема 13.28. *Последний инвариантный множитель матрицы $A \in F^{n\times n}$ является ее минимальным многочленом.*

Предметный указатель