УЧЕБНИКИ
НГТУ

Серия основана в 2001 году

А. М. ИВЛЕВА, А. Г. ПИНУС,
А. В. ЧЕХОНАДСКИХ

ОСНОВЫ АЛГЕБРЫ И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

2-е издание

НОВОСИБИРСК
2021

УДК 512:514.12(075.8)
И 255

Рецензенты:

д-р физ.-мат. наук, профессор *В.Г. Бардаков*
д-р физ.-мат. наук, профессор *Е.В. Семенко*

Учебник подготовлен
на кафедре алгебры и математической логики
для студентов I курса технических университетов

Ивлева А.М.

И 255 Основы алгебры и аналитической геометрии : учебник / А.М. Ивлева, А.Г. Пинус, А.В. Чехонадских. – 2-е издание. – Новосибирск : Изд-во НГТУ, 2021. – 286 с. – (Учебники НГТУ).

ISBN 978-5-7782-4430-6

Учебник содержит материал, соответствующий стандартной программе высшей алгебры, линейной алгебры и аналитической геометрии всех факультетов НГТУ и других технических университетов.

Благодаря простоте изложения книга может быть полезна для студентов-заочников, а также лекторов, читающих курс по этой тематике.

УДК 512:514.12(075.8)

ISBN 978-5-7782-4430-6

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Предлагаемый учебник представляет собой попытку связного изложения круга начальных понятий алгебры в целом и линейной алгебры с аналитической геометрией в частности. К сожалению, специальные книги по алгебре и геометрии при всей своей многочисленности страдают избыточной полнотой и недостаточной доступностью.

Авторы постарались избежать и того, и другого. Во-первых, материал отобран в точном соответствии с программой по алгебре и аналитической геометрии для технических специальностей и скомпонован так, чтобы книга могла изучаться последовательно, как и всякий учебник. Во-вторых, простота и подробность изложения ориентированы на читателя, располагающего только сведениями из школьной программы. Наконец, в учебник включено некоторое количество вопросов и задач, решение которых сделает знакомство читателя с материалом более близким.

То, что учебник рассчитан на первокурсника технического университета, позволяет надеяться, что некоторые наиболее сложные и трудоемкие доказательства могут быть опущены без ущерба для общего комплекса знаний, – скажем, приведена только формулировка теоремы Гаусса об алгебраической замкнутости поля комплексных чисел, понятие перестановки заменило собой даже поверхностное знакомство с группой подстановок и т. д.

Для преподавателей алгебраических разделов высшей математики может представлять интерес изложение конкретных вопросов, тем более, что главным ориентиром для авторов служил реальный объем семестрового лекционного курса. Для студентов-заочников чтение этой книги может в какой-то мере заменить слушание лекций.

Главы 1, 3, 4 и 5 написаны проф. А.Г. Пинусом; 2, 6 и 7 – доц. А.В. Чехонадских, главы 8–10 – доц. А.М. Ивлевой.

В тексте принята сквозная нумерация глав; для разделов она двойная, для пунктов – тройная. Новые термины выделяются курсивом, как и те многочисленные утверждения, которые допускают четкую формулировку, но не заслуживают того, чтобы называться теоремами.

Авторы надеются, что некоторое удовольствие от чтения этой книги искупит ее недостатки, и с признательностью примут любые замечания и пожелания для последующей доработки.

ГЛАВА 1
ОСНОВНЫЕ ПОНЯТИЯ: МНОЖЕСТВА, ЧИСЛА И МНОГОЧЛЕНЫ

1.1. ВВЕДЕНИЕ

Первые главы посвящены тем понятиям алгебры, без которых невозможно изложение ни линейной алгебры и ее приложений к аналитической геометрии, интересующих нас непосредственно, ни всей высшей математики. В современной науке под алгеброй понимают часть математики, занимающуюся изучением множеств с заданными на них операциями, т. е. отображениями, ставящими в соответствие различным наборам элементов взятого множества некоторые элементы этого же множества. Так, мы можем говорить о множестве натуральных чисел с заданными на этом множестве операциями сложения и умножения, т. е. операциями, ставящими в соответствие паре чисел n и m новое число $n + m$ или nm. Можно изучать множество векторов на плоскости с операцией их сложения: $\overline{a}, \overline{b} \to \overline{a} + \overline{b}$ и т. д.

Простейшие алгебраические операции – арифметические действия над натуральными и положительными рациональными числами – встречаются в самых ранних математических текстах. Термин же «Алгебра» происходит от названия сочинения Мухаммеда аль-Хорезми «Аль-джебр аль-мукабала», написанного в IX веке и излагающего общие приемы для решения задач, сводящихся к алгебраическим уравнениям первой и второй степени. В XV–XVI веках вместо словесных описаний алгебраических действий в математических сочинениях появляются общепринятые теперь знаки «+», «–», «$\sqrt{\ }$» и скобки, а также начинают применяться буквенные обозначения и складывается современная алгебраическая символика. Тем не менее вплоть до середины XIX века под алгеброй понималась наука о буквенных вычислениях: тождественные преобразования буквенных формул, решение алгебраических уравнений и т. д. И только со второй половины XIX века все большая часть алгебраических сочинений посвящается изучению произвольных, абстрактных ***множеств*** с заданными на них ***операциями***; а на первый план выдвигается ***аксиоматический метод***, т. е. изучение и сведение к уже известным, освоенным действиям абстрактных операций, определенных на множествах и удовлетворяющих тем или иным свойствам – так называемым ***аксиомам***.

В настоящее время алгебра представляет собой довольно сложное математическое образование, включающее в себя такие самостоятельные разделы, как теория групп, теория колец, теория полей, универсальная алгебра, теория топологических алгебр и другие. Одним из важных разделов алгебры является линейная алгебра, в которой изучаются линейные пространства, обобщающие известное из школьного курса математики понятие векторного пространства, а также матрицы. Результаты линейной алгебры нашли очень важное и эффективное применение в аналитической геометрии – науке об описании простейших геометрических объектов на алгебраическом языке – с помощью уравнений.

Роль алгебры в современной математике исключительно велика, и с очевидностью просматривается тенденция к дальнейшей «алгебраизации» математики. Типичный путь изучения многих математических объектов, порой далеких от алгебры, состоит в построении алгебраических систем, в той или иной мере отражающих поведение изучаемых объектов. Результативность подобного подхода связана с возможностью применить для решения задачи не только словесные рассуждения, но и мощный аппарат формальных алгебраических методов.

1.2. МНОЖЕСТВА И ИХ ЭЛЕМЕНТЫ

1.2.1. Под ***множеством*** мы будем понимать совокупность каких-либо объектов. Объекты, входящие в данное множество, называются ***элементами*** множества, и в дальнейшем формула $a \in A$ будет означать, что элемент a входит в совокупность элементов, образующих множество A, т. е. что элемент a ***принадлежит*** множеству A. Формула $a \notin A$ будет означать, что элемент a не входит в множество A.

Задание множеств, число элементов в которых не слишком велико, возможно непосредственным перечислением их элементов, при этом формула $A = \{ a_1,..., a_n \}$ будет означать, что элементами множества A являются элементы $a_1,..., a_n$, и только они.

Для задания множества достаточно задать критерий, позволяющий отобрать входящие в него элементы. Если подобным критерием является некоторое свойство $\Phi(x)$, которым могут обладать или не обладать конкретные объекты-элементы, то формула $A = \{ a \mid \Phi(a)\}$ означает, что множество A представляет собой совокупность всех тех элементов a, которые обладают свойством $\Phi(x)$. К примеру, если $\mathbf{R}$ – это множество всех действительных чисел и $f(x) = 0$ – некоторое уравнение, то формула $A = \{a \mid a \in \mathbf{R}$ и $f(a) = 0\}$ означает, что A есть множество всех действительных решений уравнения $f(x) = 0$. Так, в частности, получается равенство $\{-1;\ 1\} = \{a \mid a \in \mathbf{R}$ и $x^2 - 1 = 0\}$.

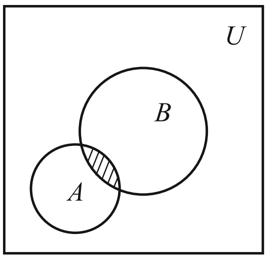
Если все элементы, входящие во множество A, являются в то же время элементами множества B, то A называется ***подмножеством*** множества B. Обозначать эту ситуацию мы будем следующим образом: $A \subseteq B$. ***Равенство*** двух множеств A и B означает, что в множества A и B входят одни и те же элементы. В частности, множества A и B равны тогда и только тогда, когда $A \subseteq B$ и $B \subseteq A$. Среди всех множеств особое место занимает множество, не содержащее ни одного элемента – ***пустое множество***, обозначаемое как $\varnothing$. Таким образом, для любого множества A имеют место утверждения: $\varnothing \subseteq A$ и $A \subseteq A$.

1.2.2. Операции над множествами. Помимо предложенного выше способа задания множества с помощью свойства (обладание им – это критерий для выделения элементов) другим способом могут служить конструкции, с помощью которых необходимые множества строятся из множеств, образованных ранее. Среди этих конструкций основными являются операции ***пересечения***, ***объединения*** и ***разности*** множеств.

Для любых множеств A и B через $A \cap B$ обозначим общую часть множеств A и B, т. е. элементами множества $A \cap B$ являются элементы, входящие во множества A и B одновременно. Если при этом

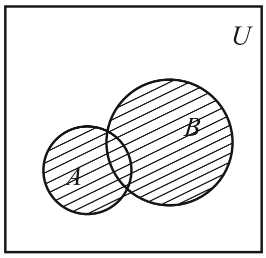
$$A = \{\, a \mid \Phi(a)\} \text{ и } B = \{\, a \mid \Psi(a)\},\text{ то } A \cap B = \{a \mid \Phi(a) \text{ и } \Psi(a)\}.$$

Одним из самых наглядных способов представления множеств являются их ***диаграммы*** (иногда их называют *диаграммами Венна*), когда множества изображаются как множества точек той или иной плоской геометрической фигуры. В этом случае иллюстрацией конструкции пересечения множеств служит диаграмма (рис. 1). Через $A_1 \cap \ldots \cap A_n$ обозначим пересечение нескольких множеств $A_1, \ldots, A_n$, т. е. $A_1 \cap \ldots \cap A_n = \{a \mid a \in A_1, \ldots,\ a \in A_n\}$.



$A \cap B$

Рис. 1



$A \cup B$

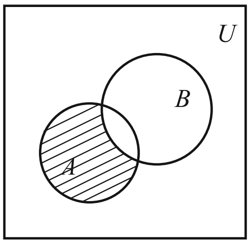
Рис. 2

Под объединением $A \cup B$ множеств A и B понимаем совокупность элементов, входящих хотя бы в одно из множеств A или B. Если $A = \{\, a \mid \Phi(a)\}$ и $B = \{\, a \mid \Psi(a)\}$, то $A \cup B = \{\, a \mid \Phi(a) \text{ или } \Psi(a)\}$ (рис. 2).

Через $A_1 \cup \ldots \cup A_n$ обозначим объединение нескольких множеств $A_1, ..., A_n$, т. е. $A_1 \cup \ldots \cup A_n = \{\, a \mid a \in A_1$, или $a \in A_2, ...,$ или $a \in A_n\}$.

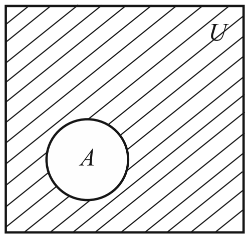
Разность $A \backslash B$ множеств A и B состоит из элементов множества A, не входящих во множество B (рис. 3).

Довольно часто все рассматриваемые в какой-то задаче элементы являются элементами некоторого «большого» множества U – ***универсального множества*** или ***универсума***, и тем самым встречающиеся в решении этой задачи множества являются подмножествами универсума U. При этом важную роль играет операция образования разности универсума U и рассматриваемого множества A, так называемое ***дополнение*** $\overline{A}$ к множеству A (рис. 4).



$A\backslash B$

Рис. 3



$\overline{A}$

Рис. 4

1.2.3. Теоретико-множественные тождества. Операции $A \cap B$, $A \cup B$, $A \setminus B$, $\overline{A}$ обладают рядом общих свойств, т. е. удовлетворяют некоторым равенствам, зависящим не от множеств, к которым они применяются, а только от последовательности выполнения этих операций. Например, для любого множества A имеет место равенство $\overline{\overline{A}} = A$. В справедливости этого равенства нетрудно убедиться с помощью приведенной диаграммы (рис. 4), иллюстрирующей понятие дополнения множества.

По-другому докажем следующее равенство:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Действительно, для доказательства мы должны показать, что любой элемент, входящий во множество $A \cap (B \cup C)$, является элементом множества $(A \cap B) \cup (A \cap C)$ и, наоборот, любой элемент множества $(A \cap B) \cup (A \cap C)$ входит во множество $A \cap (B \cup C)$. Пусть $a \in A \cap (B \cup C)$, тогда $a \in A$ и $a \in B \cup C$. Последнее означает, что $a \in B$ или $a \in C$. Наряду с отмеченным уже вхождением элемента a во множество A это влечет вхождение элемента a во множества $A \cap B$ или $A \cap C$ соответственно. А любой элемент как множества $A \cap B$, так и множества $A \cap C$ является элементом множества $(A \cap B) \cup (A \cap C)$, поэтому включение

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C) \qquad (1)$$

доказано. Обратное включение вытекает из следующей цепочки включений:

$$B, C \subseteq B \cup C; \quad A \cap B,\ A \cap C \subseteq A \cap (B \cup C);$$

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C). \tag{2}$$

Включения же (1) и (2) вместе означают доказываемое равенство $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ для любых множеств A, B, C.

1.2.4. Список основных тождеств. Только что доказанные тождества входят в число так называемых ***основных булевых тождеств***, имеющих место для произвольных множеств. Любое другое тождество, справедливое для произвольных множеств, может быть выведено из основных булевых тождеств чисто логическим путем, без анализа конкретного содержания понятий объединения, пересечения, дополнения множеств. Список этих булевых тождеств выглядит следующим образом.

1. Ассоциативность операций объединения $\cup$ и пересечения $\cap$:

$$A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C.$$

2. Коммутативность операций $\cup$ и $\cap$:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A.$$

3. Тождества идемпотентности:

$$A \cup A = A, \quad A \cap A = A.$$

4. Тождества дистрибутивности:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

5. Тождества поглощения:

$$A \cup (A \cap B) = A, \quad A \cap (A \cup B) = A.$$

6. Тождества двойственности:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

7. Тождества для пустого множества $\varnothing$ и универсума:

$$A \cup \varnothing = A,\ A \cap \varnothing = \varnothing,$$

$$A \cup U = U,\ A \cap U = A,\ A \cup \overline{A} = U,\ A \cap \overline{A} = \varnothing,$$

и эти тождества справедливы для любого множества A, являющегося подмножеством универсума U.

8. Тождество двойного дополнения:

$$\overline{\overline{A}} = A\,.$$

Читателю предоставляется возможность самому убедиться в справедливости этих тождеств, т. е. доказать приведенные равенства для любых множеств A, B, C. В этом можно убедиться и с помощью диаграмм.

1.2.5. Декартово произведение. Важную роль в построении новых множеств играет операция образования множества «двойных элементов» для двух данных множеств – их ***декартова произведения***. Упорядоченную пару элементов a, b будем обозначать как $<a, b>$. Таким образом, равенство двух пар $<a, b> = <c, d>$ имеет место тогда и только тогда, когда $a = c$ и $b = d$. Для любых множеств A и B под их декартовым произведением $A{\times}B$ понимают совокупность всех упорядоченных пар $<a, b>$, где первый элемент a взят из A, а второй – b из B, т. е.

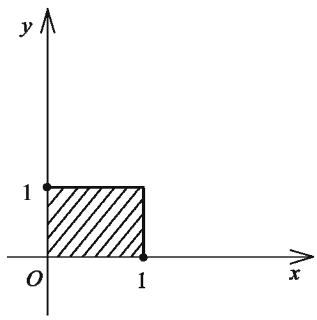


Рис. 5

$$A{\times}B = \{<a, b> \mid a \in A\,, b \in B\}.$$

К примеру, если $A = \{0,1\}$, $B = \{a, b, c\}$, то

$$A{\times}B = \{<0, a>, <0, b>, <0, c>, <1, a>, <1, b>, <1, c>\}.$$

Множество $[0,1]{\times}[0,1]$ есть множество $\{<\alpha,\ \beta> \mid \alpha, \beta \in \mathbf{R}$, и $0 \le \alpha \le 1,\ 0 \le \beta \le 1\}$, которому в декартовых координатах на плоскости соответствует квадрат точек, имеющих неотрицательные координаты, не превосходящие единицы (рис. 5).

1.3. ЧИСЛОВЫЕ МНОЖЕСТВА И ПОЛЯ

1.3.1. Понятие целого положительного или «натурального» числа связано с простым счетом предметов. Операции сложения и умножения натуральных чисел имеют ясно выраженное конкретное содержание. Множество всех натуральных чисел мы обозначим как **N**, таким образом, $\mathbf{N} = \{0, 1, 2,..., n, n + 1,...\}$. Рассмотрение простейших уравнений над натуральными числами, таких, к примеру, как $n + x = m$, приводит к тому, что часть из них, допустим $1 + x = 5$, имеет решения в натуральных числах, в то время как для других (таких, как $5 + x = 1$) никакое натуральное число решением не является. Идея ***вычитания*** как операции, обратной к сложению, привела к привлечению в качестве решений подобных уравнений нового понятия – ***отрицательного*** целого числа: как в приведенном выше примере $1 - 5 = 0 - 4$ или просто -4. Совокупность всех отрицательных целых чисел вместе с множеством натуральных чисел **N** дает в объединении множество всех ***целых чисел***, обозначаемое далее как $\mathbf{Z} = \{0, \pm 1, \pm 2,..., \pm n,...\}$. Операции сложения и умножения естественным образом распространяются с множества **N** на множество **Z**. Рассмотрение операции умножения вместо операции сложения в уравнениях, подобных приведенным выше, т. е. работа с уравнениями $ax = b$, где a, b – произвольные целые числа, вызвало к жизни понятие ***рационального числа*** $q = b/a$, частным случаем которого (при $a = 1$) являются целые числа. Множество всех рациональных чисел будем обозначать как **Q**. Естественным образом операции сложения и умножения продолжаются с множества **Z** на множество **Q**. Множество рациональных чисел, в отличие от **N** и **Z**, уже обладает некоторой «самодостаточностью» в решении простейших уравнений. Действительно, уравнения

$$r + x = q;$$

$$rx = q \text{ (при } r \neq 0)$$

имеют рациональные решения при любых $r, q \in \mathbf{Q}$. Однако рассмотрение уравнений большей степени, таких, к примеру, как

$$x^2 - 2 = 0 \quad \text{или} \quad x^2 + 2 = 0,$$

приводит нас к тем же проблемам, с какими мы сталкиваемся для уравнений вида $n + x = m$, имея в виду лишь натуральные числа.

Решение некоторых уравнений с рациональными коэффициентами, а также решение различных геометрических задач (отношение длины окружности к диаметру и т. д.) привело к возникновению понятия ***действительного числа***. Не вдаваясь в его сколько-нибудь детальный анализ, под действительным числом мы далее будем понимать выражение вида $a, \alpha_1\alpha_2 \ldots \alpha_n \ldots$, где a – некоторое целое число (положительное или отрицательное), так называемая ***целая часть*** действительного числа, а его ***дробная часть*** $\alpha_1\alpha_2 \ldots \alpha_n \ldots$ – конечная либо бесконечная последовательность цифр с известными ограничениями (мы не будем, к примеру, допускать выражения вида $a, \alpha_1\alpha_2 \ldots \alpha_n \ldots$, где, начиная с некоторого m, для всех $n \geq m$ имеет место равенство $\alpha_n = 9$).

Одно из фундаментальных свойств действительных чисел – это возможность установления взаимно однозначного соответствия между ними и точками числовой оси, т. е. прямой с фиксированной точкой – началом отсчета, направлением отсчета и заданным масштабом – единичным отрезком.

Обозначим совокупность всех действительных чисел как **R**. Операции сложения и умножения могут быть естественным образом продолжены (как ранее с **N** на **Z**, с **Z** на **Q**) с множества рациональных чисел на множество действительных чисел **R**. При этом основные свойства этих операций, имевшие место для рациональных чисел, оказываются справедливыми и для действительных чисел из **R**.

1.3.2. Комплексные числа. Однако, сталкиваясь даже с простейшими уравнениями вида

$$x^2 + 1 = 0,$$

$$x^6 + 5 = 0,$$

мы замечаем, что ни одно из действительных чисел не является решением этих уравнений. Потребность работать с такими числами, над которыми любое уравнение имело бы решение, привела к возникновению понятия комплексного числа. Оказалось, что от действительных чисел к подобной «самодостаточности» осталось сделать лишь один шаг: добавить ко множеству действительных чисел **R** корень одного-единственного уравнения

$$x^2 + 1 = 0.$$

Обозначим формально введенный нами корень этого уравнения как i и назовем число i ***мнимой единицей***. Потребность умножения мнимой единицы i на уже освоенные нами действительные числа $\beta \in \mathbf{R}$ приводит к появлению выражений вида βi – мнимых чисел, а необходимость складывать мнимые и действительные числа – к появлению выражений вида

$$\alpha + \beta i,$$

где $\alpha, \beta \in \mathbf{R}$, каковые мы и будем впредь называть ***комплексными числами***. Множество всех комплексных чисел обозначим через **C**, т. е.

$$\mathbf{C} = \{\alpha + \beta i \mid \alpha, \beta \in \mathbf{R}\}.$$

Произвольные комплексные числа будем обозначать далее как z, w, …, оставив обозначения x, y,... для произвольных действительных чисел. Сами действительные числа мы будем отождествлять с комплексными числами вида $\alpha + 0i$, т. е. далее будем считать равными числа α и $\alpha + 0i$, где $\alpha \in \mathbf{R}$. Естественные соображения, связанные со стандартными правилами раскрытия скобок и тем, что число i является решением уравнения $x^2 + 1 = 0$, т. е. $i^2 = -1$, приводят к следующему определению операций сложения и умножения на множестве комплексных чисел. Пусть

$$z_1 = \alpha_1 + \beta_1 i, \qquad z_2 = \alpha_2 + \beta_2 i \in \mathbf{C},$$

тогда положим

$$z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i,$$

$$z_1 \cdot z_2 = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i.$$

Нетрудно заметить, что привычные нам свойства операций сложения и умножения действительных чисел, такие, как перестановочность, правила раскрытия скобок и другие, имеют место и для операций над комплексными числами. Более подробно поговорим об этом ниже, при рассмотрении абстрактного понятия поля.

Помимо операций сложения и умножения на множестве комплексных чисел **C** вводятся и обратные к ним операции вычитания и деления:

$$z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)i;$$

$$\frac{z_1}{z_2} = \frac{(\alpha_1 + \beta_1 i)(\alpha_2 - \beta_2 i)}{(\alpha_2 + \beta_2 i)(\alpha_2 - \beta_2 i)} = \frac{\alpha_1\alpha_2 + \beta_1\beta_2}{\alpha_2^2 + \beta_2^2} + \frac{\alpha_2\beta_1 - \alpha_1\beta_2}{\alpha_2^2 + \beta_2^2} i,$$

если $z_2 \neq 0$. При этом $z_2 + (z_1 - z_2) = z_1$; $z_2(z_1 / z_2) = z_1$.

Легко видеть, что ограничение введенных нами операций до множества действительных чисел **R** (т. е. при $\beta = 0$) дает традиционные операции сложения, умножения, вычитания и деления действительных чисел.

1.3.3. В итоге нами построена цепочка все больших множеств

$$\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q} \subseteq \mathbf{R} \subseteq \mathbf{C},$$

на которых определены операции +, ·, –, /. Причиной расширения понятия числа в этой цепи каждый раз являлась потребность решения тех или иных уравнений. Однако появление комплексных чисел исчерпывает эту проблему. В конце XVIII века Гауссом была доказана так называемая *основная теорема алгебры*. ***Любое уравнение вида***

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

где $a_n, a_{n-1}, \ldots, a_0$ ***– произвольные фиксированные комплексные числа и*** $n \geq 1$, ***имеет решение во множестве*** **C** ***комплексных чисел, т. е. существует некоторое комплексное число*** $z_0 = \alpha_0 + \beta_0 i$ ***такое, что***

$$a_n(\alpha_0 + \beta_0 i)^n + a_{n-1}(\alpha_0 + \beta_0 i)^{n-1} + \ldots + a_1(\alpha_0 + \beta_0 i) + a_0 = 0.$$

Хотя эта теорема действительно является одной из важнейших, название «основная теорема алгебры» – это дань традиции, и мы надеемся, что читатель не закроет на этом месте книгу, решив, что он уже постиг все «премудрости» данной науки.

1.3.4. Комплексное число вида $\alpha - \beta i = \alpha + (-\beta)\, i$ мы будем называть сопряженным к числу $z = \alpha + \beta i$ и обозначать как $\overline{z}$. Для любого натурального $n \in \mathbf{N}$ $\quad (\overline{z})^n = \overline{(z^n)}$. Действительно, к примеру,

$$(\overline{z})^2 = (\alpha - \beta i)^2 = (\alpha^2 - \beta^2) - 2\alpha\beta i = \overline{(\alpha^2 - \beta^2) + 2\alpha\beta i} = \overline{(z^2)}.$$

Точно так же $\overline{z}_1 + \overline{z}_2 = \overline{z_1 + z_2}$ и если a – действительное число, то $\overline{az} = a\overline{z}$ и, конечно, $a = \overline{a}$. В силу таких замечаний, если $a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0$ – уравнение с действительными коэффициентами $a_n,\ a_{n-1}, \ldots,\ a_0$ и комплексное число $z_0 = \alpha_0 + \beta_0 i$ является корнем этого уравнения (т. е. $a_n(\alpha_0 + \beta_0 i)^n + \ldots + a_1(\alpha_0 + {}$ ${} + \beta_0 i) + a_0 = 0$), то

$$a_n\,(\overline{z}_0)^n + \ldots + a_1\overline{z}_0 + a_0 = a_n\overline{\left(z_0^n\right)} + \ldots + a_1\overline{z}_0 + a_0 =$$
$$= \overline{\left(a_n z_0^n\right)} + \ldots + \overline{a_1 z_0} + \overline{a}_0 = \overline{a_n z_0^n + \ldots + a_1 z_0 + a_0} = \overline{0} = 0.$$

Таким образом, ***любое уравнение с действительными коэффициентами, имеющее корнем комплексное число*** z_0***, имеет корнем и сопряженное число*** $\overline{z}_0$. Этот факт нам понадобится далее при анализе корней уравнений и разложении на множители многочленов в пункте 1.4.9.

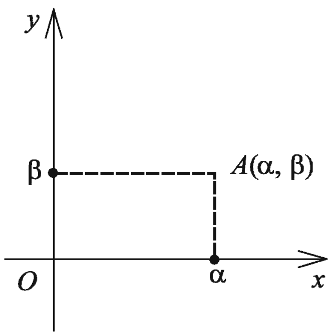


Рис. 6

1.3.5. Изображение комплексных чисел. Любое комплексное число $z = \alpha + {}$ ${} + \beta i$ однозначно определяется парой действительных чисел: $\alpha = \operatorname{Re} z$ – ***действительной частью*** числа z и $\beta = \operatorname{Im} z$ – коэффициентом при мнимой единице: ***мнимой частью*** числа z. Как мы уже вспоминали, числовая ось служит наглядным геометрическим представлением для действительных чисел. В то же время декартовы системы координат на плоскости, т. е. две взаимно перпендикулярные числовые оси, позволяют установить взаимно однозначное соответствие между парами действительных чисел <α, β> и точками плоскости (рис. 6).

А так как сопоставление $z = \alpha + \beta i \to \langle\alpha, \beta\rangle$ является взаимно однозначным соответствием между комплексными числами и парами действительных чисел, то мы приходим к взаимно однозначному соответствию между комплексными числами и точками плоскости, на которой задана декартова система координат (рис. 7).

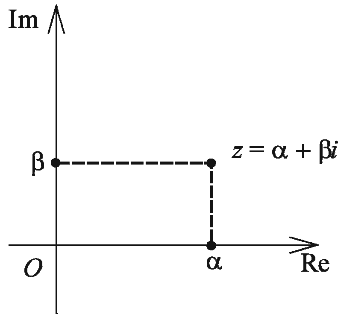


Рис. 7

При этом мы будем отождествлять число $z = \alpha + \beta i$ с изображающей это число точкой A с координатами <α, β>, а также с вектором $\overline{OA} = (\alpha; \beta)$, имеющим координаты α, β в рассматриваемой координатной системе (здесь мы чуть-чуть забегаем вперед, поскольку о векторах и их координатах речь пойдет в главе 3). Заметим, что действительные числа будут соответствовать точкам оси Ox, а мнимые – точкам оси Oy.

1.3.6. Модуль и аргумент. ***Модулем*** $|z|$ комплексного числа z назовем расстояние от этого числа до начала координат. Тем самым $|z| = \sqrt{\alpha^2 + \beta^2}$. Очевидно, что $|z| \geq 0$ для любого z и $|z| = 0$ тогда и только тогда, когда $z = 0$. ***Аргументом*** $\operatorname{Arg} z$ комплексного числа называется угол между положительным направлением оси Ox и вектором $\overline{OA}$, соответствующим числу z; если угол отсчитывается против часовой стрелки, он берется со знаком «плюс», если по часовой стрелке – со знаком «минус» (рис. 8).

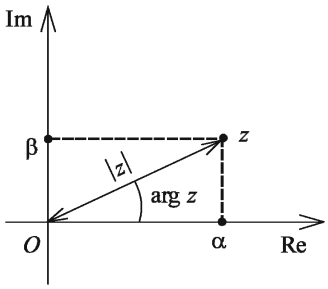


Рис. 8

Значение $\operatorname{Arg} z$ (как и любой угол) определено неоднозначно, а именно с точностью до слагаемых, кратных 2π. ***Главным значением*** $\arg z$ аргумента комплексного числа z назовем значение аргумента, заключеное в интервале $(-\pi, \pi]$. Очевидно, что

$$\arg z = \begin{cases} \operatorname{arctg}\dfrac{\beta}{\alpha}, \text{ если } \alpha > 0, \\ \operatorname{arctg}\dfrac{\beta}{\alpha} - \pi, \text{ если } \alpha, \ \beta < 0, \\ \operatorname{arctg}\dfrac{\beta}{\alpha} + \pi, \text{ если } \alpha < 0, \beta > 0, \\ \dfrac{\pi}{2}, \text{ если } \alpha = 0, \beta > 0, \\ -\dfrac{\pi}{2}, \text{ если } \alpha{=}0, \beta{<}0, \\ \text{не определен, если } \alpha = \beta = 0. \end{cases}$$

Стандартные соотношения в треугольнике приводят к равенствам $\alpha = |z| \cos \arg z$, $\beta = |z| \sin \arg z$, и поэтому число z может быть записано в виде $z = |z| (\cos \arg z + \sin \arg z)$. Последее представление числа z называют ***тригонометрической формой*** записи комплексного числа z (в то время как запись $z = \alpha + \beta i$ называют ***алгебраической формой*** числа z). Приведенные выше формулы позволяют легко переходить от алгебраической формы комплексного числа к тригонометрической и обратно. Заметим также, что для любого комплексного числа z, если

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

где r – действительное положительное число, то $|z| = r$ и $\arg z = \varphi$.

Заметим, что знаменитая формула Эйлера

$$e^{i\varphi} = \cos\varphi + i\sin\varphi$$

позволяет еще более упростить и запись комплексного числа, и дальнейшие операции с ним:

$$z = r(\cos\varphi + i\sin\varphi) = re^{i\varphi}.$$

Такая форма записи комплексного числа называется ***показательной***.

1.3.7. Умножение и деление в тригонометрической форме. Наибольшие проблемы в арифметике комплексных чисел вызывают операции умножения и деления, выполняемые в соответствии с формулами раздела 1.3.2. Значительно проще в этом отношении умножение и деление комплексных чисел в тригонометрической форме.

Пусть

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

тогда

$$z_1 z_2 = |z_1|(\cos\varphi_1 + i\sin\varphi_1)|z_2|(\cos\varphi_2 + i\sin\varphi_2) =$$

$$= |z_1||z_2|(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) =$$

$$= |z_1||z_2|((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\cos\varphi_1\sin\varphi_2 +$$

$$+ \sin\varphi_1\cos\varphi_2)) = |z_1||z_2|(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)).$$

Как замечено выше, это влечет равенство $|z_1 z_2| = |z_1||z_2|$ и $\arg(z_1 z_2) = \arg z_1 + \arg z_2$. Тем самым ***модуль произведения комплексных чисел равен произведению их модулей, а аргумент произведения равен сумме аргументов сомножителей.***

Вычислим теперь тригонометрическую форму числа $1/z_2$, обратного к ненулевому числу z_2. Так как для него

$$z_2 = |z_2|(\cos\varphi_2 + i\sin\varphi_2),$$

то

$$\frac{1}{z_2} = \frac{1}{|z_2|}\frac{1}{\cos\varphi_2 + i\sin\varphi_2},$$

а в соответствии с формулой п. 1.3.2

$$\frac{1}{\cos\varphi_2 + i\sin\varphi_2} = \frac{\cos\varphi_2 - i\sin\varphi_2}{\cos^2\varphi_2 + \sin^2\varphi_2} = \cos(-\varphi_2) + i\sin(-\varphi_2).$$

Таким образом,

$$\frac{1}{z_2} = \frac{1}{|z_2|}(\cos(-\varphi_2) + i\sin(-\varphi_2)),$$

т. е. модуль числа, обратного к z_2, равен $1/|z_2|$, а аргумент числа, обратного к z_2, противоположен $\arg z_2$.

Теперь используем это для вычисления тригонометрической формы частного z_1 / z_2:

$$\frac{z_1}{z_2} = z_1 \frac{1}{z_2} = |z_1| (\cos \varphi_1 + i \sin \varphi_1) \frac{1}{|z_2|} (\cos (-\varphi_2) + i \sin (-\varphi_2)) =$$

$$= \frac{|z_1|}{|z_2|} (\cos (\varphi_1 - \varphi_2) + i \sin (\varphi_1 - \varphi_2)).$$

Последнее равенство получено согласно сформулированному выше правилу вычисления произведения комплексных чисел в тригонометрической форме. Таким образом, ***модуль частного двух комплексных чисел равен частному модулей, а аргумент частного равен разности аргумента делимого и аргумента делителя***. Выпишем полученные нами формулы:

$$z_1 \cdot z_2 = |z_1| |z_2| (\cos (\arg z_1 + \arg z_2) + i \sin (\arg z_1 + \arg z_2)),$$

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|} (\cos (\arg z_1 - \arg z_2) + i \sin (\arg z_1 - \arg z_2)).$$

1.3.8. Возведение в степень. Предпоследняя формула п. 1.3.7 дает возможность получить так называемую ***формулу Муавра***, позволяющую легко вычислять натуральные степени комплексных чисел. Действительно, перемножая число $z = |z| (\cos \arg z + i \sin \arg z)$ само на себя n раз в соответствии формулой п. 1.3.7, получаем формулу Муавра:

$$z^n = |z|^n (\cos (n \arg z) + i \sin (n \arg z)).$$

1.3.9. Извлечение корней. В свою очередь формула Муавра позволяет получить формулу для вычисления корней n-й степени из комплексных чисел, где n – любое натуральное число. Действительно, корнем n-й степени из числа z называется любое число w такое, что $w^n = z$. Таким образом, в силу формулы Муавра равенство $w^n = z$ приводит к равенству

$$|w|^n (\cos (n \arg w) + i \sin (n \arg w)) = |z| (\cos \arg z + i \sin \arg z).$$

Отсюда получаем, что $|w|^n = |z|$, т. е. $|w| = \sqrt[n]{|z|}$ (причем здесь $\sqrt[n]{|z|}$ – арифметический корень n-й степени из положительного числа $|z|$), а также равенство

$$n \arg w = \arg z + 2\pi k,$$

где k – любое целое число. Таким образом,

$$\arg w = \frac{\arg z + 2\pi k}{n}$$

и любой корень w n-й степени из числа z имеет вид

$$w = \sqrt[n]{|z|}\left(\cos\frac{\arg z + 2\pi k}{n} + i\sin\frac{\arg z + 2\pi k}{n}\right).$$

В силу периодичности функций $\cos\varphi$ и $\sin\varphi$, период которых равен 2π, в действительности мы получаем n различных значений корней n-й степени из числа z, соответствующих значениям $k = 0, 1, 2, ..., n-1$, что согласуется с теоремой п. 1.3.3. Итак, окончательно: из любого комплексного числа z, отличного от нуля, существует ровно n различных корней n-й степени:

$$\left(\sqrt[n]{z}\right)_0 = \sqrt[n]{|z|}\left(\cos\frac{\arg z}{n} + i\sin\frac{\arg z}{n}\right);$$

$$\left(\sqrt[n]{z}\right)_1 = \sqrt[n]{|z|}\left(\cos\frac{\arg z + 2\pi}{n} + i\sin\frac{\arg z + 2\pi}{n}\right);$$

$$\left(\sqrt[n]{z}\right)_2 = \sqrt[n]{|z|}\left(\cos\frac{\arg z + 4\pi}{n} + i\sin\frac{\arg z + 4\pi}{n}\right);$$

....

$$\left(\sqrt[n]{z}\right)_{n-1} = \sqrt[n]{|z|}\left(\cos\frac{\arg z + 2\pi(n-1)}{n} + i\sin\frac{\arg z + 2\pi(n-1)}{n}\right).$$

Заметим, что все эти корни имеют один и тот же модуль, а значит, лежат на одной окружности с центром в начале координат.

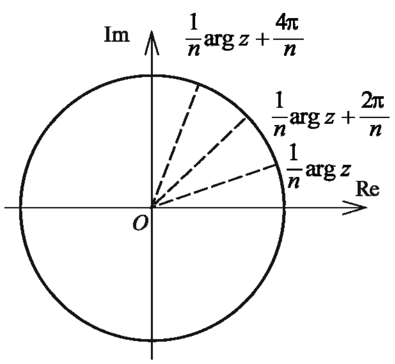


Рис. 9

Аргумент же очередного значения корня $\sqrt[n]{z}$ получается из аргумента предыдущего значения корня $\sqrt[n]{z}$ добавлением числа $(2\pi)/n$. Тем самым значения корней n-й степени из числа z делят окружность радиуса $\sqrt[n]{|z|}$ с центром в начале координат на n равных частей (рис. 9).

1.3.10. Числовые и абстрактные поля. Освоив таким образом арифметику комплексных чисел, попробуем отметить некоторые общие свойства операций сложения и умножения для множеств рациональных, действительных и комплексных чисел. Во всех трех случаях сложение и умножение обладают свойствами коммутативности и ассоциативности; для любого ненулевого числа x существует обратное к нему число $1/x$ и т. д. Можно выделить некоторый круг основных свойств операций сложения и умножения, которые позволяют довольно успешно производить арифметические вычисления как с рациональными, так и с действительными либо комплексными числами. Потребовав наличия этих свойств у произвольной пары операций с двумя аргументами на некотором множестве, мы можем надеяться на столь же успешное проведение арифметических вычислений с элементами рассматриваемого множества. Таким образом, мы приходим к понятию поля. Для краткости операция с одним аргументом именуется ***одноместной***, с двумя – ***двухместной***.

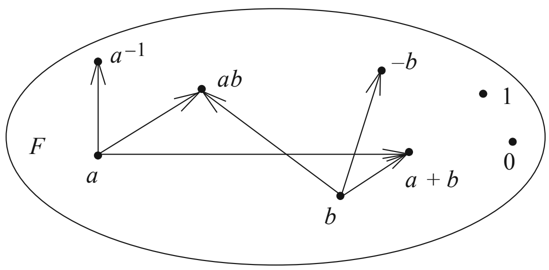


Рис. 10

Определение поля. Произвольное множество $\mathbf{F}$ с заданными на нем двумя двухместными операциями $+$, $\cdot$ и выделенными в этом множестве двумя элементами – нулем (0) и единицей (1) называется ***полем*** (рис. 10), если эти операции и выделенные элементы удовлетворяют следующим требованиям – аксиомам:

1) коммутативность сложения и умножения

$$a + b = b + a, \quad ab = ba;$$

2) ассоциативность сложения и умножения

$$a + (b + c) = (a + b) + c, \quad (ab)\, c = a\, (bc);$$

3) свойства нуля и единицы

$$a + 0 = a, \quad a \cdot 1 = a;$$

4) существование ***противоположного*** и ***обратного*** элементов:
для любого a существует b такой, что $a + b = 0$,
для любого $a \neq 0$ существует b такой, что $ab = 1$;

5) дистрибутивность умножения относительно сложения:

$$a\, (b + c) = ab + ac.$$

Говоря о выполнении этих аксиом, имеют в виду, что соответствующие равенства выполняются для любых элементов $a, b, c \in \mathbf{F}$.

Очевидным образом все пять аксиом выполняются для рациональных, действительных и комплексных чисел, т. е. $\langle \mathbf{Q}; +, \cdot, 0, 1 \rangle$, $\langle \mathbf{R}; +, \cdot, 0, 1 \rangle$, $\langle \mathbf{C}; +, \cdot, 0, 1 \rangle$ являются полями. Предоставляем читателю самому убедиться в том, что ни натуральные числа $\langle \mathbf{N}; +, \cdot, 0, 1 \rangle$, ни целые числа $\langle \mathbf{Z}; +, \cdot, 0, 1 \rangle$ не образуют поля (укажите, какие из аксиом 1–5 не выполняются на этих множествах).

1.3.11. В качестве нового примера поля приведем так называемое поле вычетов по модулю простого числа – ***простое поле Галуа***. Пусть p – некоторое ***простое*** натуральное число (т. е. делителями числа p являются только 1 и само p). Через $\mathbf{F}_p$ обозначим p-элементное множество $\{0, 1, \ldots, p-1\}$ и введем на $\mathbf{F}_p$ двухместные операции $+_p$ и $\cdot_p$ (т. е. ***сложение и умножение по модулю p***) следующим образом:

для $a, b, c \in \mathbf{F}_p$ имеет место равенство $a +_p b = c$ (соответственно $a \cdot_p b = c$), если c есть остаток от деления обычной, традиционной суммы (произведения) натуральных чисел a и b на число p. Например, если $p = 7$, то $3 +_7 4 = 0$, $5 +_7 5 = 3$, $1 \cdot_7 2 = 2$, $5 \cdot_7 5 = 4$ и т. д. Непосредственная проверка всех пяти аксиом из определения поля показывает, что $< \mathbf{F}_p ;\ +_p ,\ \cdot_p , 0, 1>$ является полем. В дополнительном комментарии нуждается, возможно, лишь аксиома 4. Хорошо известно следующее утверждение: для любых целых чисел a, b существуют целые числа c, d такие, что

$$ac + bd = (a, b),$$

где (a, b) – наибольший общий делитель чисел a и b (позже в п. 1.4.6 мы докажем аналогичное утверждение уже не для чисел, а для многочленов; доказательство обоих утверждений проводится по одной и той же схеме рассуждения, поэтому здесь мы его опускаем). В силу простоты числа p для любого числа $a \in \mathbf{F}_p$ окажется, что $(a, p) = 1$ и найдутся такие $c, d \in \mathbf{Z}$, что $ac + pd = 1$. Тогда если $c = pm + c_1$, где $c_1 \in \mathbf{F}_p$, то $ac_1 = p(-d - m) + 1$, т. е. остаток от деления произведения ac_1 на p равен 1. Значит, число $c_1 \in \mathbf{F}_p$ играет роль обратного к числу a, существование которого утверждает аксиома 4.

1.3.12. В заключение отметим некоторые простые свойства, вытекающие из аксиом 1 – 5, справедливые для любых полей.

Прежде всего, заметим, что ноль и единица в поле единственны. Предположим, что помимо отмеченных элементов 0, 1 поля $\mathbf{F}$ еще некоторые элементы $0'$, $1'$ поля играют ту же роль, что 0 и 1 соответственно, т. е. для любых $a \in \mathbf{F}$

$$a + 0' = a, \quad a \cdot 1' = a.$$

Рассмотрим элемент $0 + 0'$; по только что отмеченному свойству элемента $0'$ имеем равенство $0 + 0' = 0$. С другой стороны, по коммутативности сложения $0 + 0' = 0' + 0$, а $0' + 0 = 0'$ по свойству нуля. Тем самым $0 = 0'$, т. е. только один элемент 0 в поле обладает свойством нуля. Точно так же доказывается, что $1 = 1'$, т. е. единственный элемент в поле обладает свойством единицы.

Элементы, существование которых гарантирует аксиома 4 из определения поля в п. 1.3.10, названы ***противоположным*** (если $a + b = 0$) и ***обратным*** (если $ab = 1$) к элементу a. Они также единственны для каждого a. Покажем, к примеру, что из равенств $a + b_1 = 0,\ a + b_2 = 0$ вытекает совпадение элементов b_1 и b_2. Действительно, $b_1 + (a + b_2) = b_1 + 0 = b_1$. С другой стороны, ввиду коммутативности и ассоциативности сложения $b_1 + (a + b_2) = (a + b_2) + b_1 = (b_2 + a) + b_1 = b_2 + (a + b_1) = b_2 + 0 = b_2$. Тем самым $b_1 = b_2$. Точно так же доказывается единственность обратных элементов. Далее противоположный к a элемент поля будем обозначать как $-a$, а обратный – как a^{-1}. К примеру, в поле $\mathbf{F}_7$ выполняются равенства $-2 = 5, -4 = 3,\ 3^{-1} = 5,\ 4^{-1} = 2$ и т. д.

Заметим, что для любого элемента a поля $\mathbf{F}$ имеют место равенства

$$a \cdot 0 = 0;$$

$$(-1)\, a = -a.$$

Действительно,

$$a = a \cdot 1 = a\,(1 + 0) = a \cdot 1 + a \cdot 0 = a + a \cdot 0.$$

Прибавим к обеим частям равенства $a = a + a \cdot 0$ элемент $(-a)$, получим

$$0 = a + (-a) = (a + a \cdot 0) + (-a) =$$

$$= a + (a \cdot 0 + (-a)) = a + ((-a) + a \cdot 0)) = (a + (-a)) + a \cdot 0 =$$

$$= 0 + a \cdot 0 = a \cdot 0 + 0 = a \cdot 0.$$

Тем самым показано первое равенство из этой пары. Воспользуемся им для доказательства второго равенства:

$$0 = a \cdot 0 = a\,(1 + (-1)) = a \cdot 1 + a\,(-1) = a + (-1)\, a$$

и, значит, $(-1)\, a = -a$.

Важной особенностью, которой обладают все поля, является отсутствие в поле так называемых ***делителей нуля***, т. е. для любого поля $\mathbf{F}$ и любых его элементов a, b из равенства $ab = 0$ вытекает, что либо a, либо b равны нулю. Действительно, если $ab = 0$ и $a \neq 0$, то, умножив обе части равенства $ab = 0$ на элемент a^{-1}, получаем

$$0 = 0 \cdot a^{-1} = (a\,b)a^{-1} = a(b\,a^{-1}) = a(a^{-1}b) = (a\,a^{-1})b = 1 \cdot b = b \cdot 1 = b.$$

Результат n-кратного сложения единицы поля $\mathbf{F}$ – элемент $1 + ... + 1$ не зависит от способа выполнения этого сложения, т. е. от расстановки скобок (в силу ассоциативности сложения), и мы обозначим этот элемент как $n_{\mathbf{F}}$. Важнейшим свойством различных полей является их так называемая ***характеристика***. Будем говорить, что поле $\mathbf{F}$ имеет характеристику ноль, если для любого ненулевого натурального числа n элемент поля $n_{\mathbf{F}}$ отличен от нуля. В противном случае характеристикой поля назовем наименьшее ненулевое число n такое, что элемент поля $n_{\mathbf{F}}$ равен нулю. Например, характеристика полей $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ равна нулю, а поля $\mathbf{F}_7$ равна 7. Нетрудно показать (сделайте это сами!), что характеристика любого поля либо равна нулю, либо есть некоторое простое натуральное число p (характеристика поля $\mathbf{F}_p$ равна p). Любое поле $\mathbf{F}$ характеристики ноль содержит некоторое подмножество (точнее, ***подполе***), в точности такое же, как поле рациональных чисел. Любое поле $\mathbf{F}$ характеристики p содержит некоторое подполе, устроенное так же, как поле $\mathbf{F}_p$.

Легким упражнением для читателя будет доказательство того, что для любого элемента a поля $\mathbf{F}_p$ сумма p слагаемых $a + ... + a$ также равна нулю.

1.4. МНОГОЧЛЕНЫ НАД ПОЛЕМ

1.4.1. ***Многочленом*** от переменной x над полем $\mathbf{F}$ назовем выражение вида

$$a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0,$$

где n – натуральное число; $a_n, \ldots, a_0$ – элементы поля $\mathbf{F}$.

Элементы $a_n, \ldots, a_0$, входящие в многочлен, называются ***коэффициентами***, а слагаемые $a_n x^n$, $a_{n-1}x^{n-1}, \ldots, a_1 x, a_0$ – ***одночлена-***

ми многочлена. Если $a_n \neq 0$, то слагаемое $a_n x^n$ называется ***старшим членом*** многочлена, а число n – ***степенью*** многочлена. Заметим, что ненулевые числа поля **F** сами оказываются многочленами нулевой степени, а степень нуля можно было бы считать равной $-\infty$.

Два многочлена равны, если равны их степени и совпадают коэффициенты этих многочленов при одинаковых степенях переменной x, т. е. равенство $a_n x^n + \ldots + a_1 x + a_0 = b_k x^k + \ldots + b_1 x + b_0$ верно тогда и только тогда, когда $n = k$ и для любого $i = 1, \ldots, n$ верно равенство $a_i = b_i$.

1.4.2. Под ***суммой*** многочленов

$$f[x] = a_n x^n + \ldots + a_1 x + a_0\,;$$

$$g[x] = b_k x^k + \ldots + b_1 x + b_0$$

мы будем понимать многочлен

$$f[x] + g[x] = c_m x^m + \ldots + c_1 x + c_0\,,$$

где m есть наибольшее из чисел n и k, а $c_i = a_i + b_i$; если при этом $n > k$, то полагаем $b_{k+1} = \ldots = b_n = 0$, а если $n < k$, то полагаем $a_{n+1} = \ldots = a_k = 0$.

1.4.3. Под ***произведением*** многочленов $f[x]$ и $g[x]$ мы будем понимать многочлен

$$f[x]\,g\,[x] = \sum_{i=0}^{n+k}\left(\sum_{j=0}^{i} a_j b_{i-j}\right)x^i = a_n\,b_k x^{n+k} + \ldots +$$

$$+\,(a_0 b_i + a_1\,b_{i-1} + \ldots + a_i\,b_0)x^i + \ldots + (a_0 b_1 + a_1\,b_0)x + a_0 b_0$$

(здесь $a_i = 0$ при $i > n$ и $b_j = 0$ при $j > k$). Конечно, этот результат не отличается от простого раскрытия скобок с последующим приведением подобных.

1.4.4. Без труда проверяется, что для любых многочленов $f[x]$, $g[x]$, $h[x]$ над полем **F** имеют место следующие равенства:

– ассоциативность сложения и умножения

$$f[x] + (g[x] + h[x]) = (f[x] + g[x]) + h[x],$$

$$f[x]\ (g[x]\ h[x]) = (f[x]\ g[x])\ h[x];$$

– коммутативность сложения и умножения

$$f[x] + g[x] = g[x] + f[x], \qquad f[x]\ g[x] = g[x]\ f[x];$$

– существование нуля и единицы

$$f[x] + 0 = f[x], \quad f[x]\ 1 = f[x];$$

– дистрибутивность умножения относительно сложения

$$f[x]\ (g[x] + h[x]) = f[x]\ g[x] + f[x]\ h[x];$$

– существование многочлена $(-f[x])$, такого, что

$$f[x] + (-f[x]) = 0.$$

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

Легко убедиться, что для любых многочленов $f[x]$, $g[x]$ над полем **F** из равенства $f[x]\ g[x] = 0$ вытекает, что либо $f[x] = 0$, либо $g[x] = 0$. Более того, степень произведения многочленов $f[x]$ и $g[x]$ равна сумме их степеней.

1.4.5. Докажем теперь следующее утверждение, связанное с делимостью многочленов.

Теорема о делении с остатком. ***Для любых двух многочленов*** $f[x]$ ***и*** $g[x]$ ***над полем*** **F** ***существует пара многочленов*** $q[x]$ ***и*** $r[x]$ ***над полем*** **F** ***таких, что***

$$f[x] = g[x]\ q[x] + r[x],$$

причем степень многочлена $r[x]$ ***меньше степени многочлена*** $g[x]$. Многочлен $g[x]$ будем называть ***частным***, а $r[x]$ – ***остатком от деления*** $f[x]$ на $g[x]$.

Доказательство. Пусть $f[x] = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$; $g[x] = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0$ $(b_m \neq 0)$.

Если $f[x] = 0$ или если $n < m$, то доказываемое равенство будет справедливо при $q[x] = 0$ и $r[x] = f[x]$.

Если $n \geq m$, то вычтем из $f[x]$ многочлен $g[x]$, умноженный на $\frac{a_n}{b_m}x^{n-m}$: $f[x] - \frac{a_n}{b_m}x^{n-m}\, g[x] = f_1[x]$.

В результате старший член $a_n x^n$ многочлена $f[x]$ уничтожится и получится, что $f_1[x] = a'_{n-1}x^{n-1} + \ldots + a'_1 x + a'_0$ для некоторых элементов $a'_{n-1}, \ldots, a'_1, a'_0$ из поля $\mathbf{F}$.

Если $n - 1 \geq m$, то снова повторяют процесс понижения степени:

$$f_1[x] - \frac{a'_{n-1}}{b_m}x^{n-m-1}g[x] = f_2[x]$$

и т. д. Очевидно, что в конце концов, а именно на k-м шаге, где $k = n - m$, мы придем к многочлену $f_{k+1}[x] = r[x]$, равному нулю или имеющему степень ниже степени многочлена $g[x]$. Таким образом,

$$f[x] - \frac{a_n}{b_m}x^{n-m}g[x] = f_1[x];$$

$$f_1[x] - \frac{a'_{n-1}}{b_m}x^{n-m-1}g[x] = f_2[x];$$

....

$$f_k[x] - \frac{a_m^{(k)}}{b_m}\, g[x] = r[x] \ (k = n - m).$$

Складывая эти равенства почленно и вычеркивая одинаковые слагаемые, стоящие и в левой, и в правой части, получаем

$$f[x] - \left(\frac{a_n}{b_m}x^{n-m} + \frac{a'_{n-1}}{b_m}x^{n-m-1} + \ldots + \frac{a_m^{(k)}}{b_m}\right)g[x] = r[x]$$

или

$$f[x] = g[x]\, q[x] + r[x],$$

где

$$q[x] = \frac{a_n}{b_m} x^{n-m} + \frac{a'_{n-1}}{b_m} x^{n-m-1} + \ldots + \frac{a_m^{(k)}}{b_m}.$$

Теорема доказана.

Нетрудно показать, что многочлены $q[x]$ и $r[x]$, удовлетворяющие этой теореме, определяются единственным образом.

Если остаток от деления многочлена $f[x]$ на многочлен $g[x]$ равен нулю, то будем говорить, что многочлен $f[x]$ делится на $g[x]$.

1.4.6. Многочлен $d[x]$ над полем **F** назовем ***общим делителем*** двух произвольных многочленов $f[x]$ и $h[x]$, если и $f[x]$, и $h[x]$ делятся на $d[x]$. Многочлен $D[x]$ назовем ***наибольшим общим делителем (н. о. д.)*** многочленов $f[x]$ и $h[x]$, если $D[x]$ – общий делитель многочленов $f[x]$ и $h[x]$ и если $D[x]$ делится на любой другой общий делитель $d[x]$ многочленов $f[x]$ и $h[x]$.

Теорема о наибольшем общем делителе. ***Для любых двух многочленов*** $f[x]$ ***и*** $h[x]$ ***над полем*** **F** ***существует их наибольший общий делитель.***

Доказательство. В основе доказательства этой теоремы лежит так называемый ***алгоритм Евклида***.

Если $f[x] = h[x] = 0$, то н. о. д. многочленов $f[x]$ и $h[x]$ равен нулю.

Пусть $h[x] \neq 0$. По теореме о делении с остатком найдутся многочлены $q[x]$ и $r[x]$ (степень $r[x]$ меньше степени многочлена $h[x]$) такие, что $f[x] = h[x]\, q[x] + r[x]$.

Если $r[x] \neq 0$, то делим $h[x]$ на $r[x]$:

$$h[x] = r[x] q_1[x] + r_1[x].$$

Если $r_1[x] \neq 0$, то делим $r[x]$ на $r_1[x]$:

$$r[x] = r_1[x] q_2[x] + r_2[x].$$

Степень получающихся при таком процессе остатков $r[x]$, $r_1[x]$, $r_2[x]$, ... будет, очевидно, все время убывать. Но целые неотрицательные числа не могут убывать неограниченно. Следовательно, наш процесс оборвется, т. е. в конце концов мы придем к остатку $r_k[x]$, на который нацело разделится предыдущий остаток $r_{k-1}[x]$. Покажем, что этот последний остаток $r_k[x]$ и будет н. о. д. многочленов $f[x]$ и $h[x]$.

Запишем весь процесс деления:

$$f[x] = h[x]\,q[x] + r[x];$$

$$h[x] = r[x]q_1[x] + r_1[x];$$

....

$$r_{k-2}[x] = r_{k-1}[x]q_k[x] + r_k[x];$$

$$r_{k-1}[x] = r_k[x]q_{k+1}[x].$$

Так как многочлен $r_{k-1}[x]$ делится нацело на $r_k[x]$, то из предложенного равенства получаем, что $r_{k-2}[x]$ также делится нацело на $r_k[x]$.

$$r_{k-2}[x] = r_{k-1}[x]q_k[x] + r_k[x] = r_k[x]q_{k+1}[x]q_k[x] + r_k[x]\cdot 1 =$$

$$= r_k[x]\left(q_{k+1}[x]q_k[x] + 1\right).$$

Постепенно продвигаясь вверх таким образом, мы наконец дойдем до многочленов $f[x]$ и $h[x]$ и убедимся, что и $f[x]$ и $g[x]$ делятся на $r_k[x]$.

Остается показать, что $r_k[x]$ делится на любой другой общий делитель $d[x]$ многочленов $f[x]$ и $h[x]$. Так как $r[x] = f[x] - h[x]\,q[x]$, а $f[x]$ и $h[x]$ делятся на $d[x]$, то и $r[x]$ делится на $d[x]$. Так как $r_1[x] = h[x] - r[x]q[x]$, $h[x]$ и $r[x]$ делятся на $d[x]$, то и $r_1[x]$ делится на $d[x]$. Постепенно опускаясь в системе равенств вниз, мы доберемся до $r_k[x]$ и убедимся, что $r_k[x]$ делится на $d[x]$. Тем самым остаток $r_k[x]$ действительно является н. о. д. многочленов $f[x]$ и $h[x]$, существование которого и требовалось доказать.

Внимательно анализируя доказательство последней теоремы, можно получить

Следствие 1. ***Если*** $D[x]$ ***– это н. о. д. многочленов*** $f[x]$ ***и*** $h[x]$ ***над полем*** $\mathbf{F}$***, то найдутся многочлены*** $p[x]$ ***и*** $s[x]$ ***над полем*** $\mathbf{F}$ ***такие, что***

$$f[x]\,p[x] + h[x]\,s[x] = D[x].$$

Обратите внимание на то, что последнее утверждение совершенно аналогично тому утверждению о натуральных числах, которым мы пользовались в конце раздела 1.3.11 при рассмотрении простых полей Галуа, именно при доказательстве обратимости их элементов.

1.4.7. Пусть $f[x] = a_n x^n + \ldots + a_1 x + a_0$ – произвольный многочлен над полем $\mathbf{F}$. Заменив в нем переменную x каким-либо элементом $c \in \mathbf{F}$, мы получим элемент из того же поля $\mathbf{F}$ следующего вида:

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

называющийся значением многочлена $f[x]$ при значении переменного $x = c$ и обозначаемый как $f[c]$.

Заметим, что если $f[x] + g[x] = h[x]$ и $f[x]\,g[x] = p[x]$, то для любого $c \in \mathbf{F}$ справедливы равенства $f[c] + g[c] = h[c]$, $f[c]\,g[c] = p[c]$.

Элемент поля $\mathbf{F}$ называется ***корнем*** многочлена $f[x]$ над полем $\mathbf{F}$, если $f[c] = 0$.

В качестве следствия из теоремы о делении с остатком мы получим такое утверждение.

Теорема Безу. ***Для любого многочлена*** $f[x]$ ***над полем*** $\mathbf{F}$ ***и любого элемента*** c ***поля*** $\mathbf{F}$ ***остаток от деления*** $f[x]$ ***на многочлен*** $(x - c)$ ***равен*** $f[c]$.

Доказательство. Пусть $q[x]$ есть частное от деления $f[x]$ на многочлен $(x - c)$ и $r[x]$ – остаток от этого деления. Так как степень делителя $(x - c)$ равна единице, а степень остатка должна быть меньше степени делителя, то степень многочлена $r[x]$ нулевая, т. е. $r[x]$ – некоторый элемент d поля F. Подставляя в равенство

$$f[x] = (x - c)\,q[x] + d$$

вместо переменной x элемент c, получаем, что

$$f[c] = (c - c)\,q[c] + d,$$

т. е.

$$f[c] = 0 \cdot q[c] + d = d.$$

Теорема доказана.

Отсюда получаем

Следствие 2. ***Элемент*** c ***поля*** **F** ***является корнем многочлена*** $f[x]$ ***над этим полем тогда и только тогда, когда этот многочлен делится на*** $(x - c)$.

1.4.8. Элемент c поля **F** назовем ***корнем кратности k*** многочлена $f[x]$, если многочлен $f[x]$ делится на многочлен $(x-c)^k$ и не делится на многочлен $(x-c)^{k+1}$. Корни кратности 1 называются ***простыми корнями*** многочлена.

Если $c_1,..., c_s$ суть все корни многочлена $f[x]$, лежащие в поле **F** и имеющие кратности $k_1,..., k_s$ соответственно, то ***полагаем число корней многочлена*** $f[x]$ ***в поле*** **F** ***равным*** $k_1 + ... + k_s$. При этом многочлен $f[x]$ должен делиться на многочлен $(x - c_1)^{k_1}...$ $(x - c_s)^{k_s}$, имеющий степень $k_1 + ... + k_s$. С другой стороны, как уже замечалось выше, если многочлен $f[x]$ делится на многочлен $g[x]$, то степень многочлена $g[x]$ не больше степени многочлена $f[x]$. Тем самым имеет место

Следствие 3. ***Число корней любого многочлена*** $f[x]$ ***над полем*** **F, *лежащих в этом поле, не превышает степени многочлена*** $f[x]$.

На самом деле число корней, лежащих в этом поле, может быть строго меньше степени многочлена. Например, многочлен с действительными коэффициентами $f[x] = x^3 - 2x^2 + x - 2 = (x^2 + 1)(x - 2)$ имеет только один корень в поле действительных чисел. Однако если рассматривать корни этого многочлена в большем поле **C** комплексных чисел, то там этот многочлен будет иметь уже три корня: i, $-i$, 2.

Из сформулированной в пункте 1.3.2 ***основной теоремы алгебры*** вытекает следующее утверждение.

Следствие 4. ***Любой многочлен*** $f[x]$ ***степени*** n ***над полем*** **C *комплексных чисел имеет ровно*** n ***корней. В частности, любой многочлен над полем*** **C *комплексных чисел разложим в произведение линейных множителей, т. е. многочленов вида*** $(x - c)$, ***и константы из поля*** **C.**

Доказательство. По основной теореме алгебры многочлен $f[x]$ имеет корень c_1 в поле **C** и тем самым по следствию 3 найдется многочлен $f_1[x]$ над полем **C** такой, что

$$f[x] = (x - c_1) f_1[x].$$

Применяя, в свою очередь, основную теорему к многочлену $f_1[x]$, получаем его корень c_2, лежащий в поле **C**, и многочлен $f_2[x]$ такие, что

$$f_1[x] = (x - c_2) f_2[x].$$

Продолжая этот процесс n раз, где n – степень многочлена $f[x]$, получаем равенство

$$f[x] = (x - c_1)(x - c_2)...(x - c_n) f_n[x].$$

В силу того, что степени $f[x]$ и $(x - c_1)...(x - c_n)$ совпадают, степень многочлена $f_n[x]$ должна быть равна нулю, т. е. $f_n[x]$ является некоторым элементом d поля **C**. Таким образом,

$$f[x] = d(x - c_1)...(x - c_n),$$

что и требовалось доказать.

Некоторые из этих корней $c_1,..., c_n$ многочлена могут совпадать между собой. Таким образом, в общем случае

$$f[x] = d(x - c_1)^{k_1} ...(x - c_s)^{k_s},$$

где все комплексные числа $c_1,..., c_s$ попарно различны и $k_1 + ... + k_s = n$ – степень многочлена $f[x]$. Число c_1 – корень кратности $k_1,..., c_s$ – корень кратности k_s многочлена $f[x]$.

Обозначим теперь произведение $(x - c_2)^{k_2} ...(x - c_s)^{k_s}$ как $f_1[x]$. Тогда $f[x] = d(x - c_1)^{k_1} f_1[x]$.

Найдем производную $f'[x]$ многочлена $f[x]$:

$$f'[x] = \left(d(x - c_1)^{k_1} f_1[x]\right)' = d\left((x - c_1)^{k_1} f_1[x]\right)' =$$

$$= dk_1(x-c_1)^{k_1-1} f_1[x] + d(x-c_1)^{k_1} f_1'[x].$$

Рассмотрим два случая:

1) $k_1 = 1$. Тогда

$$f'[c_1] = dk_1 f_1[c_1] + d(c_1 - c_1) f_1'[c_1] = dk_1 f_1[c_1] \neq 0,$$

2) $k_1 > 1$. Тогда

$$f'[c_1] = dk_1(c_1 - c_1)^{k_1-1} f_1[c_1] + d(c_1 - c_1)^{k_1} f_1'[c_1] = 0,$$

так как $k_1 - 1 > 0$.

Итак, если c_1 – простой корень многочлена $f[x]$, то $f'[c_1] \neq 0$, а если c_1 – корень многочлена $f[x]$ кратности не меньше двух, то $f'[c_1] = 0$.

Нетрудно показать, что справедливо более общее утверждение.

Следствие 5. ***Число*** c ***является корнем кратности*** k ***многочлена*** $f[x]$ ***тогда и только тогда, когда*** $f[c] = 0,\ f'[c] = 0, \ldots,$ $f^{(k-1)}[c] = 0,\ f^{(k)}[c] \neq 0.$

1.4.9. Рассмотрим теперь проблему разложения на множители многочленов над полем **R** действительных чисел.

Следствие 6. ***Любой многочлен*** $f[x]$ ***над полем*** **R** ***действительных чисел разложим в произведение линейных множителей вида*** $(x - c)$, ***квадратных трехчленов вида*** $(x^2 + ax + b)$ ***и константы*** d, ***причем*** $c,\ a,\ b,\ d \in$ **R**, ***а трехчлены не имеют действительных корней.***

Доказательство. Пусть $f[x] = a_n x^n + \ldots + a_1 x + a_0$, где $a_n, \ldots,$ $a_1,\ a_0 \in$ **R**. Так как **R** $\subseteq$ **C**, т. е. действительные числа являются частным случаем комплексных чисел, то согласно следствию 4 многочлен $f[x]$ как многочлен над полем **C** может быть разложен в произведение линейных множителей

$$f[x] = d(x - c_1)\ldots(x - c_n),$$

где $d,\ c_1, \ldots,\ c_n \in$ **C**. Числа $c_1, \ldots,\ c_n$ являются корнями многочлена $f[x]$. Часть из них может быть действительными числами, пусть это $c_1,..., c_k$, где $k \le n$. Тогда корни $c_{k+1},..., c_n$ многочлена $f[x]$ с действительными коэффициентами не являются действительными числами. Согласно замечанию в п. 1.3.2 (вслед за формулировкой основной теоремы алгебры), для любого комплексного корня c_l $(k+1 \le l \le n)$ среди корней $c_{k+1},..., c_n$ найдется сопряженный к числу c_l, т. е. для некоторого j, такого, что $k + 1 \le j \le n$, окажется, что $c_l = \alpha_l + \beta_l i$, и $c_j = \alpha_l - \beta_l i$. Но тогда $(x - c_l)(x - c_j) = (x - \alpha_l - \beta_l i)(x - \alpha_l + \beta_l i) = x^2 - 2a_l x + \alpha_l^2 + \beta_l^2$.

Обозначим $a = -2\alpha_l$, $b = \alpha_l^2 + \beta_l^2$ и получим, что

$$(x - c_l)(x - c_j) = (x^2 + ax + b),$$

где a, b – некоторые действительные числа, причем дискриминант полученного квадратного трехчлена меньше нуля. Таким образом,

$$f[x] = d(x - c_1)...(x - c_k)\left(x^2 + a_1 x + b_1\right)...\left(x^2 + a_s x + b_s\right),$$

где $c_1,..., c_k$, a_1, $b_1,...,$ a_s, b_s – некоторые действительные числа. Нетрудно также заметить. что число d, равное частному от деления многочлена $f[x]$ на произведение

$$(x - c_1)...(x - c_k)\left(x^2 + a_1 x + b_1\right)...\left(x^2 + a_s x + b_s\right)$$

с действительными коэффициентами, и само будет действительным числом (конечно, это просто коэффициент при старшей степени многочлена $f[x]$). Следствие доказано.

Еще раз подчеркнем, что построенные в доказательстве следствия квадратные трехчлены $(x^2 + ax + b)$ не имеют действительных корней.

ГЛАВА 2
МАТРИЦЫ И ОПРЕДЕЛИТЕЛИ

2.1. МАТРИЦЫ И ДЕЙСТВИЯ С НИМИ

Матрицы, как и определители, представляют собой весьма удобные, а в большинстве случаев и просто незаменимые средства записи и обработки информации. Овладение этими, надо сказать, достаточно простыми средствами приходит с некоторым навыком обращения с ними. Поэтому мы включаем в текст вопросы, на которые читателю желательно ответить, и предложения что-то объяснить самостоятельно; они формулируются с красной строки.

2.1.1. Понятие матрицы и элемента матрицы. Прямоугольную таблицу, заполненную числами, будем называть ***числовой матрицей*** или просто ***матрицей***. Обозначать матрицы будем заглавными латинскими буквами. Вертикальные колонки таких таблиц называются **столбцами**, горизонтальные – ***строками***.

Говорят, что матрица имеет ***порядок*** $m \times n$, если у нее m строк и n столбцов. Когда нужно, порядок матрицы подписывается правее и ниже буквы, обозначающей матрицу: $A_{m\times n}$, $D_{3\times 3}$.

Линии, разграничивающие столбцы и строки, в матрицах обычно не проводятся. Зато сами они заключаются в круглые скобки:

$$(-2), \qquad (2 \quad 3 \quad 15 \quad -8),$$

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 4 & 5 \\ 6 & 1 \\ -1 & -11 \\ 3 & 4 \end{pmatrix}.$$

Кроме слова «порядок» в том же смысле используются слова ***размер*** или ***размерность***, а также ***формат*** матрицы.

Попробуйте назвать порядок пяти нарисованных выше матриц.

Если число строк и столбцов матрицы одинаково (или, что то же самое, она имеет порядок $n \times n$), то она называется квадратной.

Какие из приведенных выше матриц – квадратные?

Число, занимающее определенную клетку таблицы, с указанием строчного и столбцового номеров этой клетки, называется ***элементом*** матрицы. Вместо слов «строчный номер» или «столбцовый номер» часто говорят ***строчный*** или ***столбцовый индекс***. Обратите внимание, что при указании индексов строчный ставится первым (т. е. слева), столбцовый вторым (справа), нумерация строк идет сверху вниз, столбцов – слева направо:

$$A_{3\times 3} = (a_{ij})_{3\times 3} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Элементы, у которых столбцовые и строчные номера равны, вместе называются ***главной диагональю*** матрицы. Часто под главной диагональю подразумевают не столько элементы, сколько занимаемые ими места, благодаря чему возникают выражения типа «числа, стоящие на главной диагонали» и т. д.

Матрица называется ***верхнетреугольной***, если ниже главной диагонали у нее все элементы равны нулю: $a_{ij} = 0$ при $i > j$.

Унитреугольная матрица – это верхнетреугольная матрица, у которой на главной диагонали стоят единицы.

Попробуйте описать при помощи формул с участием элементов и индексов унитреугольные и ***нижнетреугольные*** матрицы (у которых элементы выше главной диагонали равны нулю).

Определение. Матрицы A и B равны, если у них одинаковый порядок и равны элементы, стоящие на одних и тех же местах:

$$a_{ij} = b_{ij}$$

для всех $i = 1, ..., m$ и $j = 1, ..., n$ (m – число строк у обеих матриц, n – число столбцов).

Таким образом, равны могут быть только полностью совпадающие матрицы, это следует иметь в виду во всех матричных равенствах и уравнениях.

Подумайте, что означает равенство $A = B$, где A – верхнетреугольная, а B – нижнетреугольная матрицы одинакового порядка?

Кстати, такие матрицы – верхне- и нижнетреугольные одновременно – называются ***диагональными***.

Наконец, квадратная матрица E называется ***единичной***, если на главной диагонали у нее стоят единицы, а все остальные элементы равны нулю: $E = (e_{ij})$, где $e_{ij} = \begin{cases} 1, i = j, \\ 0, i \neq j. \end{cases}$

2.1.2. Сложение матриц. Складывать можно только матрицы одинакового порядка, и ***суммой*** матриц $A = (a_{ij})$ и $B = (b_{ij})$ будет матрица $C = (c_{ij})$, элементы которой равны суммам соответственных элементов матриц A и B: $c_{ij} = a_{ij} + b_{ij}$ для всех i и j.

Поскольку сложение матриц сводится к сложению чисел – их элементов, все привычные свойства сложения сохраняются:

1) ассоциативность

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

2) коммутативность

$$A + B = B + A;$$

3) существование ***нуль-матриц*** $(0)_{m\times n}$ для всех порядков $m\times n$ таких, что для любой матрицы $A_{m\times n}$

$$(0) + A = A;$$

4) для всякой матрицы A существует ***противоположная матрица*** $(–A)$ того же порядка такая, что

$$A + (–A) = (0).$$

Каковы элементы матрицы $(0)_{m\times n}$? Как по элементам матрицы A указать элементы матрицы $(–A)$?

2.1.3. Умножение матриц на числа. При умножении матрицы любого порядка на число каждый ее элемент умножается на это число:

$$\alpha\, A_{m\times n} = \alpha\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & \ddots & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{pmatrix}.$$

Поскольку и эта операция сводится непосредственно к перемножению чисел, ее свойства вытекают прямо из свойств числового умножения:

5) ассоциативность: $\alpha(\beta A) = (\alpha\beta)A$;

6) дистрибутивность по матрицам: $\alpha(A + B) = \alpha A + \alpha B$;

7) дистрибутивность по числам: $(\alpha + \beta)\, A= \alpha A + \beta A$;

8) унитарность $1{\cdot}A = A$.

2.1.4. Позже, когда в главе 3 мы ознакомимся с определением ***линейного или векторного пространства***, окажется, что именно свойства 1–8 являются его аксиомами. Мы отложим пока осмысление того, что множество матриц одного и того же порядка с ***линейными операциями*** сложения и умножения на числа является линейным пространством, но предложим читателю убедиться самостоятельно (в общем виде или на числовых примерах), что, как и во всяком линейном пространстве, в пространстве матриц выполняются следующие свойства:

$$0 \cdot A = (0), \quad \alpha(0) = (0), \quad (-1)\, A = -A.$$

2.1.5. Умножение матриц. Определим сначала, как перемножается однострочная матрица, стоящая слева, и одностолбцовая – справа, с одинаковым числом элементов:

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 + a_2b_2 + \ldots + a_nb_n \end{pmatrix}.$$

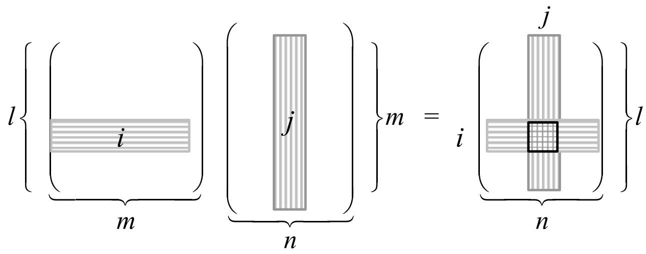
Итак, в результате получается матрица порядка 1×1, единственный элемент которой представляет собой сумму произведений пар чисел, стоящих в строке и столбце под одинаковыми номерами.

В общем случае матрицы можно перемножить, если длина строки у левого сомножителя равна высоте столбца у правого. Или, что то же самое, если число столбцов у матрицы, стоящей слева, равно числу строк у матрицы справа. ***Иначе перемножение неосуществимо*!**

А ***произведением матриц*** $A_{l\times m}$ и $B_{m\times n}$ будет матрица $C_{l\times n}$, унаследовавшая число строк у левого сомножителя A и число столбцов – у правого сомножителя B. Каждый элемент c_{ij} матрицы-произведения получается из умножения i-й строки матрицы A на j-й столбец матрицы B:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{im}b_{mj}\,.$$

Схематически это можно изобразить так:



Обратите внимание, что для того чтобы заполнить матрицу C, надо перемножить каждую из l строк матрицы A на каждый из n столбцов матрицы B.

Приведите примеры порядков матриц A и B, при которых:

– определено произведение AB, но не определено BA;

– определены, но имеют разные порядки оба произведения AB и BA;

– определены и имеют одинаковые порядки оба произведения AB и BA.

Укажите, в какой последовательности можно перемножить три матрицы $A_{2\times 5}$, $B_{2\times 2}$, $C_{5\times 3}$? Какой порядок будет иметь это произведение?

2.1.6. Ниже приведено несколько примеров перемножения матриц. Вам рекомендуется выполнить их самостоятельно и сравнить результаты:

1) $(1 \quad 4 \quad 5)\begin{pmatrix} 3 & 2 \\ 1 & -1 \\ 0 & 0 \end{pmatrix} = (7 \quad -2)$;

2) $\begin{pmatrix} 2 & 3 & 4 \\ -1 & -7 & -5 \end{pmatrix}\begin{pmatrix} 4 & 7 \\ 0 & 3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 27 \\ 6 & -33 \end{pmatrix}$;

3) $\begin{pmatrix} 4 & 7 \\ 0 & 3 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 \\ -1 & -7 & -5 \end{pmatrix} = \begin{pmatrix} 1 & -37 & -19 \\ -3 & -21 & -15 \\ -5 & -13 & -13 \end{pmatrix}$;

4) $\begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$;

5) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -2 & 11 \\ 0 & 1 & -7 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

2.1.7. Некоторые формы записи сумм и произведений. В пункте 2.1.8 возникнет ситуация, когда нужный нам результат будет суммой конечного, но заранее не определенного количества слагаемых. В разд. 2.3 появятся произведения с тем же свойством: число сомножителей конечно, но заранее не указано и потому может быть любым. Свойства у таких сумм и произведений те же, что и у привычных – с двумя или тремя величинами.

Вопрос заключается в том, как записывать сумму с таким (сразу не определенным) числом слагаемых, которые отличаются друг от друга индексом или парой индексов?

Обычно используются две формы записи:

– для сумм величин $x_1,..., x_n$

$$x_1 + \ldots + x_n = \sum_{i=1}^{n} x_i \; ;$$

– для их произведения

$$x_1 \ldots x_n = \prod_{i=1}^{n} x_i \; .$$

Буква i, которую можно заменить на любую другую, не занятую в формуле, называется ***индексом суммирования*** *(или* ***индексом перемножения***). А запись $\sum\limits_{i=m}^{n} x_i$ означает, что суммируются величины x_i с номерами i, принимающими все значения от m до n (обратите внимание на их расположение относительно знака суммы).

Нам придется сталкиваться с произведениями и суммами, участники которых имеют двойную нумерацию. В этом случае для каждого из двух номеров вводится свой индекс и указываются границы его изменения.

Так и следует понимать обозначения

$$\sum_{i,j=1}^{m} x_{ij} \quad \text{или} \quad \prod_{i=k}^{l}\prod_{j=m}^{n} x_{ij}\ .$$

Здесь в сумме номера i и j принимают всевозможные значения от 1 до m. А в произведении левый номер, которому соответствует индекс i, пробегает все значения от k до l, правый номер или индекс j соответственно от m до n.

Иногда более удобной, но и более громоздкой является такая запись того же выражения:

$$\prod_{\substack{i=k,\dots,l \\ j=m,\dots,n}} x_{ij}\ .$$

Существенно, что должны выполняться условия $m \le n$ и $k \le l$. Равенство означало бы, что для номера i допустимо единственное значение $i = k = l$.

Иногда необходимо рассматривать суммы или произведения, где первый и второй номера как-то связаны между собой. Это условие, записанное при помощи индексов, подписывается внизу под знаком суммы или произведения, либо вместе с границами для номеров. Так, записи $\sum\limits_{\substack{i,j=1 \\ i<j}}^{n} x_{ij}$ и $\sum\limits_{1\le i<j\le n} x_{ij}$ означают одно и то же: в сумме участвуют слагаемые, оба номера которых лежат среди чисел $1,\dots, n$ и первый обязательно меньше второго.

Во всяком случае, надо иметь в виду свойства однородности и дистрибутивности для таких записей: $k\sum\limits_{i=m}^{n} x_i = \sum\limits_{i=m}^{n} kx_i$ для любого числа k;

$$\sum_{i=m}^{n}(x_i + y_i)z_i = \sum_{i=m}^{n} x_i z_i + \sum_{i=m}^{n} y_i z_i\,.$$

Для доказательства этих формул приходится пользоваться не только дистрибутивностью, но и ассоциативностью и коммутативностью сложения участвующих в них величин. Впрочем, с нарушениями этих свойств нам сталкиваться не придется.

Важна также возможность перестановки операций суммирования по двум разным индексам:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} x_{ij}\,.$$

Обе эти суммы равны двойной сумме $\sum\limits_{i,j=1}^{m\ \ n} x_{ij}$ и представляют собой сумму всех элементов матрицы $(x_{ij})_{m\times n}$, различающуюся только порядком слагаемых: в левой части складываются элементы по столбцам, а затем складываются столбцовые суммы, в правой части сначала суммируются элементы строк, потом – построчные суммы.

2.1.8. Свойства матричного умножения. Примеры 2, 3 из п. 2.1.6 показывают, что результат умножения матриц зависит от расположения сомножителей: вообще говоря, $AB \neq BA$.

В этих примерах различны даже порядки результирующих матриц.

Но и для квадратных матриц нетрудно подобрать пример, показывающий, что ***умножение матриц некоммутативно***.

И это вполне понятно: ведь левая матрица участвует в умножении строками, а правая – столбцами; когда же матрицы переставляются, в суммах п. 2.1.5, занимающих прежнее положение, оказываются произведения совсем других элементов.

Разумеется, нельзя утверждать, что это *всегда* так. Матрицы A и B назовем перестановочными, если имеет место равенство $AB = BA$. Вот несколько типичных примеров перестановочных матриц:

1) $AE = EA = A$ (см. п. 2.1.1);

2) $A\,(0) = (0)\,A = (0)$, здесь (0) – нулевая матрица;

3) $A\,A^2 = A^2A = A^3$ – что еще предстоит доказать (здесь $A^2 = AA$);

4) рекомендуем проверить, что и в примере 5 из п. 2.1.6 перестановка матриц не изменяет результата, и заодно дадим следующее

Определение. Матрица A^{-1} называется **обратной** к квадратной матрице A, если $AA^{-1} = A^{-1}A = E$. (Конечно, у всех матриц здесь один и тот же порядок $n\times n$.)

Матричное умножение имеет следующие свойства (ниже A, B и C – произвольные матрицы таких порядков, что все умножения выполнимы):

1) ассоциативность $(AB)C = A(BC)$;

2) левая дистрибутивность $(A + B)C = AC + BC$; правая дистрибутивность $A(B + C) = AB + AC$;

3) однородность $\alpha(AB) = (\alpha A)B = A(\alpha B)$, где α – любое число;

4) существование единичной матрицы E_n для каждого порядка n:

$$AE_n = E_nA = A.$$

Мы приведем самую длинную проверку – свойства 1, а проверка свойств 2 и 3 остается как упражнение; свойство 4 уже знакомо.

Итак, пусть A, B, C – три матрицы порядков $l\times m$, $m\times n$, $n\times p$ соответственно. Тогда при любой расстановке скобок умножение выполнимо и дает матрицу порядка $l\times p$.

На месте $(i\,j)$ в результирующей матрице $D = (AB)C$ стоит число

$$d_{ij} = \sum_{q=1}^{n}\left(\sum_{k=1}^{m} a_{ik}b_{kq}\right)c_{qj},$$

так как если $F = AB$, то $f_{iq} = \sum\limits_{k=1}^{m} a_{ik}b_{kq}$ и $d_{ij} = \sum\limits_{q=1}^{n} f_{iq}c_{qj}$.

Для матрицы $D' = A(BC)$ получим, что

$$d'_{ij} = \sum_{k=1}^{m} a_{ik}\left(\sum_{q=1}^{n} b_{kq}c_{qj}\right),$$

здесь $G = BC$ и $g_{kj} = \sum_{q=1}^{n} b_{kq}c_{qj}$, а $d'_{ij} = \sum_{k=1}^{m} a_{ik}g_{kj}$.

Но, как мы видели в п. 2.1.7,

$$d_{ij} = \sum_{q=1}^{n}\left(\sum_{k=1}^{m} a_{ik}b_{kq}\right)c_{qj} = \sum_{q=1}^{n}\sum_{k=1}^{m} a_{ik}b_{kq}c_{qj} =$$

$$= \sum_{k=1}^{m}\sum_{q=1}^{n} a_{ik}b_{kq}c_{qj} = \sum_{k=1}^{m} a_{ik}\sum_{q=1}^{n} b_{kq}c_{qj} = d'_{ij}.$$

Это и требовалось доказать. Отсюда вытекает третий пример перестановочности матриц. Каких матриц? $(AB)C = A(BC)$. Выше есть два примера неперестановочности и три примера перестановочности.

Задача. Матрица A^{T} порядка $n\times m$ называется ***транспонированной*** для матрицы A порядка $m\times n$, если ее столбцы совпадают с вертикально развернутыми строками матрицы A, т. е. $a_{ij} = (a^{\mathrm{T}})_{ji}$. Что такое a^{T}? Докажите, что:

1) $(A^{\mathrm{T}})^{\mathrm{T}} = A$;

2) $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

2.1.9. Алгебра матриц. Теперь, ознакомившись с операциями над матрицами и их свойствами, мы можем дать одно из фундаментальных алгебраических понятий: ***алгебры над полем***.

Пусть дано некоторое числовое поле $\mathbf{F}$ и множество A, на котором заданы операции сложения, умножения на числа из $\mathbf{F}$ и перемножения между собой. Это множество называется ***ассоциативной алгеброй над полем*** $\mathbf{F}$, если в нем выполняются свойства, обнаруженные нами у матриц, а именно:

- сложение в алгебре A:

– ассоциативно: $(a + b) + c = a + (b + c)$ для всех $a, b, c \in A$;

– коммутативно: $a + b = b + a$ для всех $a, b \in A$;

– существует ***нулевой элемент*** 0 такой, что $a + 0 = a$ для всякого $a \in A$;

– для всякого $a \in A$ существует ему противоположный $(–a)$ такой, что $a + (–a) = 0$;

• умножение на числа из поля **F**:

– ассоциативно: $\alpha(\beta a) = (\alpha\beta)a$ для всех $\alpha, \beta \in \mathbf{F}$, $a \in A$;

– дистрибутивно по числам: $(\alpha + \beta)a = \alpha a + \beta a$ для всех $\alpha, \beta \in \mathbf{F}$, $a \in A$;

– дистрибутивно по элементам алгебры: $\alpha(a + b) = \alpha a + \alpha b$ для всех $\alpha \in \mathbf{F}$, $a, b \in A$;

– унитарно: $1\,a = a$ для $1 \in \mathbf{F}$ и любого $a \in A$;

• умножение элементов алгебры:

– ассоциативно: $(ab)c = a(bc)$ для всех $a, b, c \in A$;

– дистрибутивно слева: $(a + b)c = ac + bc$ и справа: $a(b + c) = ab + ac$ для всех $a, b, c \in A$;

– однородно: $\alpha(ab) = (\alpha a)b = a(\alpha\, b)$ для всех $\alpha \in \mathbf{F}$, $a, b \in A$.

И, как мы убедились, всем этим свойствам удовлетворяет множество квадратных матриц любого фиксированного порядка $n\times n$ с элементами из числового поля.

Заметим, что первые восемь свойств – сложения и умножения на числа полностью соответствуют аксиомам ***линейного пространства***, приводимым в п. 3.2.2 далее. Таким образом, можно было бы назвать алгеброй над полем такое линейное пространство, в котором элементы можно перемножать между собой, и это умножение ассоциативно, дистрибутивно со сложением (и справа, и слева) и однородно вместе с умножением на числа.

Можно добавить, что множество $\{A_{n\times n}\}$ таких матриц – это **унитарная алгебра** или **алгебра с единицей**: поскольку она содержит единичную матрицу E_n с тем свойством, что $E_n A = AE_n = A$ для любой матрицы $A_{n\times n}$. (В общем случае единицей алгебры A называется такой элемент $\boldsymbol{e} \in A$, что $a\boldsymbol{e} = \boldsymbol{e}a = a$ для всякого $a \in A$.)

Убедитесь, что многочлены (раздел 1.4) также образуют алгебру над полем.

2.2. ОПРЕДЕЛИТЕЛИ МАЛЫХ ПОРЯДКОВ

Определитель – это одна из важнейших характеристик квадратной матрицы. Проще всего сказать, что ***определитель***, или ***детерминант***, – это число, которое выражается определенным способом через элементы квадратной матрицы и характеризует матрицу в целом.

Иногда определителем называют и матрицу, стоящую под его значком. Реже определителем называют тот набор действий, который надо выполнить, чтобы получить из элементов матрицы само это число. Правда, никакой путаницы при этом не возникает, и в дальнейшем мы не будем уточнять, в каком именно смысле употребляется слово «определитель» – в каждом случае это будет ясно из контекста.

Хотя существует способ задать определитель матрицы для произвольного порядка n (с которым мы встретимся позже), сначала целесообразно ознакомиться с определителями матриц малых порядков или, короче говоря, ввести ***определители малых порядков***.

2.2.1. Определители малых порядков. Определитель $|A|$ матрицы A порядка 1×1 – это ее единственный элемент a:

$$|(a)| = a.$$

Определитель матрицы порядка 2×2 задается формулой

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

А для порядка 3×3 – следующей формулой:

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

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

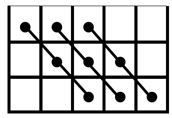
Кроме обозначения $|A|$ для определителя часто используются и другие: $\det A$ или $\Delta\,(A)$.

Чтобы легче запомнить слагаемые определителя третьего порядка, можно предложить следующий способ.

Пристроим к данной матрице четвертый и пятый столбцы, просто дублируя первый и второй:

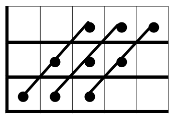
$$\left(\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right).$$

Теперь три диагональных произведения, спускающихся слева сверху – вправо вниз, включаются в сумму со своим знаком:



(+)

А три диагональных произведения, поднимающихся слева снизу – вправо вверх, – с противоположным знаком:



(–)

Сумма этих шести произведений и составляет выражение, выписанное выше.

Есть и другие возможности, одна из которых показана в следующем пункте.

2.2.2. Основные свойства определителей. Все восемь указанных здесь свойств можно проверить, исходя из приведенных в п. 2.2.1 формул или на числовых примерах, правда, первый способ дает доказательство свойства, второй – свидетельство в его пользу. Мы же приведем только логические доводы, позволяющие получить из одних свойств другие. Определитель первого порядка «подчиняется» свойствам 3, 5 и 7, для других он слишком мал, так что в целом можно иметь в виду только определители второго и третьего порядков. Естественно, впоследствии эти же правила подтвердятся и для определителей произвольных порядков (см. пп. 2.4.1 – 2.4.8). Обратите внимание, что определитель не меняется при транспонировании: $|A| = |A^{\mathrm{T}}|$, поэтому все свойства, сформулированные для строк, верны для столбцов и наоборот.

1. Если поменять местами две строки или два столбца определителя, его значение сменится на противоположное.

2. Для определителя третьего порядка выполняются правила ***разложения по первой строке или первому столбцу***:

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

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

Эти равенства проверяются просто раскрытием всех определителей по формулам п. 2.2.1. Заметим, что из-за равноправия всех строк между собой (свойство 1!) так же, как и всех столбцов, определитель третьего порядка можно разложить подобным образом по любой строке и любому столбцу, поменяв их местами с первой строкой и первым столбцом соответственно (со сменой знака).

В качестве упражнения читателю рекомендуется записать разложения определителя по второй строке и по третьему столбцу.

3. Определитель, имеющий нулевую строку или нулевой столбец (т. е. строку или столбец, заполненные нулями), равен нулю.

(Для третьего порядка это свойство легко следует из возможности переставить нулевую строку (столбец) на первое место и разложить определитель по первой строке (свойство 2). Для второго порядка достаточно воспользоваться определением.)

4. Если все элементы некоторой строки (столбца) умножить на одно и то же число, то и весь определитель умножится на это число.

(Это следует из того, что в каждое из произведений, сумма которых дает определитель, входит ровно один элемент каждой строки и каждого столбца – именно они и умножаются на число, а вместе с ними и вся сумма). Иными словами, общий множитель элементов строки или столбца матрицы можно выносить за знак определителя.

5. Определитель матрицы, две строки которой одинаковы, равен нулю. Это же верно для столбцов.

(Всё следует из свойства 1: если одинаковые строки переставить, то, с одной стороны, должен измениться знак определителя, с другой – не изменится ничего.)

6. Определитель матрицы, две строки которой пропорциональны, равен нулю. Это же верно для столбцов.

(Следует из свойств 4 и 5, ведь коэффициент пропорциональности можно вынести за знак определителя, после чего в нем останутся две одинаковые строки.)

7. Если какая-то строка определителя представлена в виде суммы двух строк, как, например, вторая строка определителя третьего порядка:

$$a_{21} = a'_{21} + a''_{21}, \quad a_{22} = a'_{22} + a''_{22}, \quad a_{23} = a'_{23} + a''_{23},$$

то и определитель представляется в виде суммы:

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

Это же верно и для столбцов.

Свойство вытекает из возможности разложить определитель по «суммарной» строке (свойство 2):

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

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

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

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

8. Определитель не изменится, если к одной его строке прибавить другую строку, домноженную на любое число.

(Это свойство следует из свойств 6 и 7. Определитель, возникающий после такого прибавления, раскладывается в сумму двух определителей (свойство 7), один из которых равен исходному, а другой содержит пропорциональные строки и потому равен нулю.) Например:

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

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \alpha \underbrace{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}_{\substack{\| \\ 0}}.$$

2.3. ОПРЕДЕЛИТЕЛЬ ПРОИЗВОЛЬНОГО ПОРЯДКА

Здесь и ниже мы увидим, как обобщается понятие определителя на случай квадратной матрицы любого порядка.

Формулу, выражающую определитель через элементы матрицы, можно указать уже сейчас.

Определение. Пусть A_n – квадратная матрица порядка n. Тогда

$$\det A = \sum_{п \in П} (-1)^{п} a_{1\ п(1)} \dots a_{n\ п(n)}.$$

Но чтобы раскрыть содержание всех обозначений этой формулы, нам потребуется несколько предварительных шагов.

2.3.1. Определение. Произведение n различных элементов матрицы A_n, взятых по одному из каждой ее строки и каждого ее столбца, назовем ***правильным***.

Как получить правильное произведение? Возьмем уже известный пример матрицы размера 3×3:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Выберем какой-то элемент из первой строки, скажем, a_{12}. Еще два элемента (всего их должно быть три!) должны быть взяты из других строк – в нашем случае второй и третьей и других столбцов – первого и третьего. Если мы возьмем из второй строки элемент a_{23}, то последним должен быть взят элемент a_{31}. Получилось правильное произведение $a_{12}a_{23}a_{31}$.

В общем случае, по-видимому, стоит идти этим же путем: брать элементы из разных строк поочередно, начиная с первой, следя, чтобы они не попадали в уже использованные столбцы.

Скажем, из первой строки взят элемент $a_{1\,\text{п}(1)}$ (как вы видите, номер попавшегося столбца мы обозначили п(1), конечно, это просто одно из чисел 1, ..., n). Из второй строки возьмем элемент $a_{2\,\text{п}(2)}$, из третей $a_{3\,\text{п}(3)}$ и так далее. Все числа п(1), п(2), ..., п(n) различны между собой, и все они находятся среди чисел 1, ..., n – ведь таковы номера столбцов. Таким образом, «п» как операция выбора оказывается просто перестановкой номеров 1, ..., n – перестановкой в самом обычном смысле, как мы говорим о перестановке стульев возле стола или занятий в расписании.

Во взятом выше примере у нас было п(1) = 2, п(2) = 3, п(3) = 1, перестановка п дала вместо исходного порядка номеров 1, 2, 3 результат 2, 3, 1.

2.3.2. Определение. Пусть п – некоторая перестановка. Условимся говорить, что пара $<i; j>$ где $i, j \in \{1, ..., n\}$ и $i < j$, ***повышающаяся***, если п(i) < п(j), и ***понижающаяся***, если п(i) > п(j) для взятой перестановки п.

Здесь все понятно: если хотя бы один из номеров сдвинулся с места, то их постепенное увеличение нарушилось, какой-то из меньших номеров попал на более высокое место, и наоборот. Тем самым возникли понижающиеся пары, обычно называемые ***инверсиями***.

Вернемся к тому же примеру. Из трех имеющихся пар пара $<1; 2>$ – повышающаяся, так как п(1) = 2 < 3 = п(2), пара $<1; 3>$ понижающаяся: п(1) = 2 > 1 = п(3). Понижающаяся и пара $<2; 3>$.

2.3.3. Определение. Перестановка п чисел 1, ..., n называется ***четной***, если $\prod\limits_{1\le i<j\le n} (\text{п}(j)-\text{п}(i)) > 0$, и ***нечетной***, если $\prod\limits_{1\le i<j\le n} (\text{п}(j)-\text{п}(i)) < 0$.

В этом определении надо разобраться. Прежде всего, произведение $\prod\limits_{1\le i<j\le n} (\text{п}(j)-\text{п}(i))$ не может равняться нулю: разным числам i и j соответствуют непременно различающиеся между собой п(i) и п(j).

Во-вторых, в произведении $\prod\limits_{1\le i<j\le n}(п(j)-п(i))$ любая пара номеров, скажем, k и m, должна встретиться в одной из скобок. Действительно, после перестановки п числа i и j должны занять какие-то места, пусть это и будут номера k и m (т. е. п(i) = k и п(j) = m). Но скобка (п(j) – п(i)) входит в произведение.

Если пара <i; j> повышающаяся, то (п(j) – п(i)) > 0. Если же это понижающаяся пара, то п(j) < п(i) и скобка (п(j) – п(i)) будет отрицательна. Порядок сомножителей значения не имеет, и теперь понятно, что произведение $\prod\limits_{1\le i<j\le n}(п(j)-п(i))$ будет положительно при четном числе понижающихся пар (четная перестановка!) и отрицательно – при нечетном.

Введем обозначение:

$$(-1)^{\pi}=\begin{cases}+1, \text{ если п четная;}\\ -1, \text{ если п нечетная.}\end{cases}$$

Это число мы будем называть знаком правильного произведения.

Последнее, на что необходимо обратить внимание, это то, что не только каждому правильному произведению соответствует некоторая перестановка п, но и каждой перестановке п номеров 1, ..., n соответствует единственное правильное произведение:

$$a_{1\,п(1)}a_{2\,п(2)}\dots a_{n\,п(n)}.$$

Если теперь обозначить П множество всех перестановок, то определение 2.3 станет вполне ясным: в сумме участвуют все правильные произведения элементов матрицы со знаками, отвечающими четности или нечетности соответствующих перестановок.

2.3.4. Пример. Попробуем применить определение 2.3 для случаев n = 1, 2, 3.

1. Если $n=1$, то никаких перестановок, кроме п(1) = 1, нет. Понижающихся пар, да и просто пар нет вообще. Поэтому $\det(a_{11})=a_{11}$.

2. Если $n=2$, то перестановок две: тождественная п(1) = 1, п(2) = 2 и «действительно» перестановка п(1) = 2, п(2) = 1.

В первом случае $\prod\limits_{1\le i<j\le n}(\text{п}(2)-\text{п}(1)) = 2 - 1 > 0$, так что тождественная перестановка четная.

Во втором случае $\prod\limits_{1\le i<j\le n}(\text{п}(2)-\text{п}(1)) = 1 - 2 < 0$, так что эта перестановка нечетная. Поэтому

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

3. В случае $n = 3$ правильных произведений шесть и перестановок шесть:

$$\begin{pmatrix} a_{11} & & \\ & a_{22} & \\ & & a_{33} \end{pmatrix} \quad \begin{pmatrix} a_{11} & & \\ & & a_{23} \\ & a_{32} & \end{pmatrix} \quad \begin{pmatrix} & a_{12} & \\ a_{21} & & \\ & & a_{33} \end{pmatrix}$$

$$\begin{pmatrix} & a_{12} & \\ & & a_{23} \\ a_{31} & & \end{pmatrix} \quad \begin{pmatrix} & & a_{13} \\ a_{21} & & \\ & a_{32} & \end{pmatrix} \quad \begin{pmatrix} & & a_{13} \\ & a_{22} & \\ a_{31} & & \end{pmatrix}.$$

Напоминаем, что если в правильном произведении участвует элемент a_{ij}, то в соответствующей перестановке $j = \text{п}(i)$. Поэтому перестановка, дающая в результате такие порядки:

а) 1, 2, 3 – четная перестановка;

б) 1, 3, 2 – нечетная, поскольку здесь

$$\prod_{1\le i<j\le 3}(\text{п}(j)-\text{п}(i)) = (3-1)(2-1)(2-3) < 0;$$

в) 2, 1, 3 – нечетная, поскольку здесь

$$\prod_{1\le i<j\le 3}(\text{п}(j)-\text{п}(i)) = (1-2)(3-2)(3-1) < 0$$

и так далее.

Результат известен нам из п. 2.2.1.

2.4. СВОЙСТВА ОПРЕДЕЛИТЕЛЕЙ

Они нам также уже известны для случая малых порядков (п. 2.2.2). Здесь мы доказываем упомянутые там и некоторые другие свойства в случае произвольного порядка.

2.4.1. Определитель транспонированной матрицы. Возможно, вы обратили внимание на то, что при построении определителя (а именно при возникновении перестановки из правильного произведения) строки получают некое преимущество перед столбцами.

Изменится ли результат, если мы начнем расставлять сомножители правильного произведения по номерам столбцов, а из номеров строк образуем перестановку?

Или, что то же самое, будет ли определитель транспонированной матрицы (где строки стали столбцами и наоборот, см. 2.1.8) равняться определителю исходной?

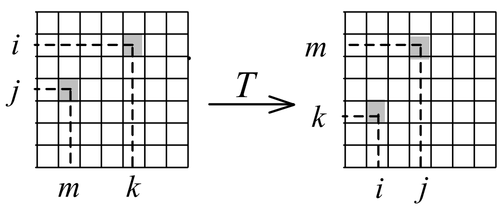
Ответ положительный: $\det A = \det A^{\mathrm{T}}$.

А для обоснования, во-первых, заметим, что произведение, правильное в исходной матрице, остается правильным и после транспонирования. Соответственно каждое правильное произведение транспонированной матрицы получилось из правильного произведения исходной, т. е. комплекты правильных произведений у обеих матриц одинаковы.

Во-вторых, вспомним, что четность или нечетность перестановки определяется количеством ее понижающихся пар. Но и понижающаяся и повышающаяся пара останутся таковыми после транспонирования!

В самом деле, пусть п – перестановка правильного произведения исходной матрицы, п′ – перестановка того же, но транспонированного произведения. Возьмем пару $\langle i, j\rangle$, тогда если k = п(i) и m = п(j), то i = п′(k) и j = п′(m).

Для случая понижающейся пары получается такая схема:



Из того, что п(i) = $k > m$ = п(j) получилось, что п′(m) = $j > i$ = п′(k) и пара <m; k> понижающаяся.

Аналогично разбирается случай повышающейся пары.

В результате и число понижающихся пар, и знак каждого правильного произведения сохранится при транспонировании.

Сохранится и их сумма, т. е. определитель.

Отсюда мы должны сделать важный вывод: ***все утверждения, верные для строк определителя, верны и для его столбцов, и наоборот.***

2.4.2. Перестановка строк в определителе. Сначала выясним, что произойдет с определителем матрицы при перестановке ее соседних строк: k-й и k + 1-й.

Прежде всего надо убедиться, что правильные произведения при этом останутся правильными. Но, когда строки поменялись местами, в каждой из них и остается по одному элементу взятого произведения. А столбцовые места элементов вообще не затрагиваются. Значит, правильность произведений сохраняется.

Выясним теперь, что произойдет со знаком правильного произведения. Для этого надо узнать, будут ли отличаться знаки выражений $\prod\limits_{1\le i<j\le n}(п(j)-п(i))$ и $\prod\limits_{1\le i<j\le n}(п'(j)-п'(i))$, где перестановка п соответствует исходному правильному произведению, а п′ – получившемуся после перемены мест.

Существенно, что для номера i, не совпадающего ни с k, ни с $k+1$, результаты перестановок п и п′ на i и j не различаются. А для k и k + 1 верны равенства п (k) = п′(k + 1) и п(k + 1) = п′(k).

(Это ясно из того, что элемент $a_{k\,п(k)}$ попал на место [k + 1; п(k)] и стал поэтому элементом $a_{k+1\,п(k)}$, также и элемент $a_{k+1\,п(k+1)}$.)

Все скобки выписанных выше произведений можно разбить на три группы:

1) скобки, в которых номера k и $k+1$ не участвуют, но из-за одинаковости действия перестановок п и п′ на прочие номера совпадают и соответствующие им скобки (п(j) – п(i)) и (п′(j) – п′(i));

2) скобки, в которых один из номеров – это k или $k+1$, а другой не равен ни k, ни $k+1$; в этом случае произойдет «парное» уравнивание скобок, пусть, например, ($i < k$, а $j = k$), (п(k) – п(i))

$(\text{п}(k+1) - \text{п}(i)) = (\text{п}'(k+1) - \text{п}'(i))(\text{п}'(k) - \text{п}'(i))$ (причем заметьте, случай $j = k + 1$ сюда уже вошел!), а случаи $i = k$ или $i = k + 1$ при $j > k$ рассматриваются аналогично;

3) единственная скобка с участием обоих номеров $i = k, j = k + 1$; здесь получится вот что:

$$\text{п}(k+1) - \text{п}(k) = -(\text{п}(k) - \text{п}(k+1)) = -(\text{п}'(k+1) - \text{п}'(k)),$$

таким образом, только одна эта скобка поменяет знак.

Следовательно, знак каждого правильного произведения изменится. Изменится и общий знак их суммы, т. е. определителя.

На этом рассмотрение случая, когда меняются местами соседние строки, закончено.

Если же такой обмен совершается между строками с номерами $i < j$, отличающимися больше чем на единицу, то результат может быть получен из предыдущего в несколько шагов.

Первый этап: i-я строка переставляется поочередно со строками, начиная с $(i + 1)$-й и кончая j-й, она при этом попадет на j-й этаж, а каждая из строк с исходными номерами $i + 1,..., j$ уменьшит номер своего этажа на единицу. Это требует $j - i$ перестановок соседних строк и столько же смен знака определителя.

Второй этап: j-я строка, стоящая теперь на $(j - 1)$-м этаже, поднимается наверх, переставляясь со строками, исходные номера которых $j - 1, j - 2,..., i + 1$. Это требует $j - i - 1$ перестановок соседних строк и столько же смен знака определителя.

В результате же i-я строка стоит на j-м месте, j-я строка попала на i-й этаж, а строки, начиная с $(i + 1)$-й и кончая $(j - 1)$-й, возвратились на свои начальные позиции.

Смен знака у определителя в общей сложности было $2(j - i) - 1$, а это число нечетное.

Тем самым доказано, что ***при перестановке строк*** (***столбцов***) ***матрицы ее определитель меняет знак***.

Что можно сказать о значении определителя, имеющего две одинаковые строки или два одинаковых столбца?

2.4.3. Разложение строки (столбца) определителя в сумму. Предположим, что каждый элемент a_{ij} i-й строки матрицы A разложен в сумму $a_{ij} = a'_{ij} + a''_{ij}$, где a'_{ij} и a''_{ij} – какие-то новые числа или выражения. Составим две новые матрицы A' и A'' так, что у первой из них в i-й строке стоят элементы $a'_{i1},\ldots, a'_{in}$, а у второй – элементы $a''_{i1},\ldots, a''_{in}$. Остальные строки обеих матриц такие же, как и у матрицы A. Как связаны между собой определители матриц A, A' и A''?

Для ответа заметим, что каждое правильное произведение матрицы A разлагается в сумму правильных произведений матриц A' и A'':

$$a_{1\,\text{п}(1)}\ldots a_{i\,\text{п}(i)}\ldots a_{n\,\text{п}(n)} = a_{1\,\text{п}(1)}\ldots(a'_{i\,\text{п}(i)} + a''_{i\,\text{п}(i)})\ldots a_{n\,\text{п}(n)} =$$
$$= a_{1\,\text{п}(1)}\ldots a'_{i\,\text{п}(i)}\ldots a_{n\,\text{п}(n)} + a_{1\,\text{п}(1)}\ldots a''_{i\,\text{п}(i)}\ldots a_{n\,\text{п}(n)}.$$

Существенно, что перестановка п при этом не меняется.

И если теперь воспользоваться определением, мы получим, что сумма правильных произведений исходного определителя разложится в две суммы: слагаемые первой содержат сомножители с одним штрихом, слагаемые второй – с двумя штрихами; соответствие слагаемых и перестановок п $\in$ П сохраняется.

Тем самым доказано, ***что если строка* (*столбец*) *определителя разложена в сумму двух новых строк* (*столбцов*), *то сам определитель равен сумме двух новых определителей, в каждом из которых вместо этой строки* (*столбца*) *стоит одна из двух новых.***

2.4.4. Общий множитель строки (столбца) определителя. Сформулируем только само свойство, а то, как его доказать, вполне ясно из рассмотрения п. 2.2.2.

Итак, предположим, что каждый элемент i-й строки определителя представлен в виде $a_{ij} = \alpha a'_{ij}$, где число α одно и то же для всех $j = 1,\ldots, n$. Если теперь мы образуем новую матрицу A', где все строки, кроме i-й, будут такими же, как в матрице A, а в i-й строке встанут элементы $a'_{i1},\ldots, a'_{in}$, тогда окажется, что $\det A = \alpha \det A'$, т. е. ***общий множитель строки или столбца определителя можно выносить за знак определителя.***

Или, что то же самое, ***если умножить строку* (*столбец*) *определителя на некоторое число, то значение определителя умножится на это же число.***

2.4.5. Пропорциональные строки или столбцы. Теперь мы предлагаем воспользоваться ответом на вопрос п. 2.4.2 и свойством п. 2.4.4 для решения следующей задачи.

Доказать, что ***определитель равен нулю, если две его строки* (*столбцы*) *пропорциональны*** (т. е. для некоторого числа α и некоторых номеров строк i и k и любых столбцовых номеров $j = 1,..., n$ выполняется $a_{ij} = \alpha a_{kj}$).

2.4.6. Разложение определителя по строке или столбцу. В пп. 2.4.3 – 2.4.5 мы видели, как влияют на определитель соотношения между элементами его строк или столбцов. После этого возникает вопрос: нельзя ли указать явно, как зависит значение определителя от самих элементов строки или столбца?

Следующая возможность бросается в глаза. Пусть нас интересует i-я строка. Давайте всю сумму, образующую определитель, разобьем на скобки так, чтобы в каждой из них были собраны все правильные произведения, содержащие какой-то один – но только один! – из элементов i-й строки. Тогда он станет общим множителем этой подсуммы и его можно будет вынести за знак суммирования (или за скобки). Схематично это выглядит так:

$$\sum_{\text{п}\in\Pi} (-1)^{\text{п}} a_{1\ \text{п}(1)} \ldots a_{n\ \text{п}(n)} = a_{i1}\left(\Sigma\ldots\right) + \ldots + a_{in}\left(\Sigma\ldots\right).$$

Если бы нам удалось найти явное выражение для сумм, стоящих в скобках, задача была бы решена.

2.4.7. Рассмотрим самый простой частный случай. Берется первая строка, к тому же все ее элементы, кроме a_{11}, равны нулю.

Матрица имеет такой вид:

$$A = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

Ясно, что все правильные произведения, содержащие иные, чем a_{11}, элементы первой строки, обращаются в нуль и вклада в значение определителя не вносят. А те, которые содержат a_{11}, не включают других элементов первого столбца.

Поэтому все реально задействованные в определителе элементы, кроме a_{11}, распределены в матрице:

$$\tilde{A} = \begin{pmatrix} a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Если обрезать по этому формату правильные произведения исходной матрицы, содержащие a_{11}, то получатся правильные произведения элементов этой новой матрицы: в них будет по $n-1$ элементу, и никакие два из них не попадут в одну строку или в один столбец. При этом образуются все правильные произведения усеченной матрицы, так как каждое из них, дополненное элементом a_{11}, станет правильным произведением исходной матрицы.

Кроме того, четность или нечетность перестановки, отвечающей правильному произведению, сохранится при переходе от исходной матрицы к усеченной: устранение из произведения элемента a_{11} не убавляет и не прибавляет ни одной понижающей пары! Действительно, для любого $k = 2,..., n$ пара $<1;\ k>$ повышающаяся, ибо $п(1) = 1$ и $п(k) \geq 2 > п(1)$.

Следовательно, сохраняются и знаки правильных произведений, откуда

$$\det A = \sum_{п \in \Pi} (-1)^{п} a_{1\ п(1)} \dots a_{n\ п(n)} = \sum_{\substack{п \in \Pi \\ п(1)=1}} (-1)^{п} a_{11} a_{2\ п(2)} \dots a_{n\ п(n)} =$$

$$= a_{11} \sum_{\tilde{п} \in \tilde{\Pi}} (-1)^{\tilde{п}} a_{2\ \tilde{п}(2)} \dots a_{n\ \tilde{п}(n)} = a_{11} \det \tilde{A},$$

где $\tilde{\Pi}$ – множество перестановок номеров $2,..., n$, а $\tilde{п}$ – элементы этого множества. И, как уже отмечалось, $(-1)^{п} = (-1)^{\tilde{п}}$, если $п(1) = 1$.

Вернемся теперь к общему случаю. Представим i-ю строку матрицы суммой n строк, в каждой из которых оставлен только один элемент a_{ij}, остальные же заменены на нули:

$$
\begin{array}{l}
+(a_{i1} \quad 0 \quad \dots \quad 0 \quad 0) \\
+(0 \quad a_{i2} \quad \dots \quad 0 \quad 0) \\
\vdots \\
+\underline{(0 \quad 0 \quad \dots \quad 0 \quad a_{in})} \\
(a_{i1} \quad a_{i2} \quad \dots \quad a_{in-1} \; a_{in}).
\end{array}
$$

Тогда исходный определитель $|A|$ тоже разлагается по свойству п. 2.4.3 в сумму определителей $\Delta_1, ..., \Delta_n$:

$$
\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{i1} & 0 & \dots & 0 \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \dots + \begin{vmatrix} a_{11} & \dots & a_{1n-1} & a_{1n} \\ \vdots & & & \vdots \\ 0 & \dots & 0 & a_{in} \\ \vdots & & & \vdots \\ a_{n1} & \dots & a_{nn-1} & a_{nn} \end{vmatrix}.
$$

Чтобы понять, чему равен определитель Δ_j, переведем элемент a_{ij} на место $(1;1)$ перестановками строк и столбцов, используя свойство 2.4.2:

$$
\Delta_j = \begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & a_{ij} & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} = (-1)^{i-1} \begin{vmatrix} 0 & \dots & a_{1j} & \dots & 0 \\ a_{11} & \dots & a_{ij} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} =
$$

$$
= (-1)^{i-1}(-1)^{j-1} \begin{vmatrix} a_{ij} & 0 & \dots & \dots & 0 \\ a_{1j} & a_{11} & \dots & \dots & a_{1n} \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ a_{nj} & a_{n1} & \dots & \dots & a_{nn} \end{vmatrix} = (-1)^{i+j} \det \tilde{A}_{ij} .
$$

Здесь $\tilde{A}_{ij}$ – это матрица, полученная из матрицы A изъятием i-й строки и j-го столбца с сохранением расположения и значений остальных элементов. Разумеется, $(-1)^{i-1}(-1)^{j-1}=(-1)^{i+j}$.

2.4.8. Опираясь на только что введенное обозначение, дадим следующее

Определение. ***Алгебраическим дополнением элемента*** a_{ij} матрицы $A_{n\times n}$ называется выражение $A_{ij}=(-1)^{i+j}\left|\tilde{A}_{ij}\right|$.

И теперь можно утверждать справедливость следующих ***формул разложения определителя по строке или по столбцу***:

$$\det A=\sum_{j=1}^{n}a_{ij}A_{ij}=\sum_{i=1}^{n}a_{ij}A_{ij}$$

(здесь в первой сумме фиксирован индекс i, что соответствует выбору i-й строки; во второй сумме – столбцовый индекс j).

Конечно, вторая формула вытекает из обоснованной нами первой ввиду свойства п. 2.4.1.

Или то же самое словами: ***определитель матрицы равен сумме произведений элементов любой своей строки или столбца на их алгебраические дополнения***.

2.5. ПРИЛОЖЕНИЯ АЛГЕБРАИЧЕСКИХ ДОПОЛНЕНИЙ К ВЫЧИСЛЕНИЮ ОБРАТНЫХ МАТРИЦ И РЕШЕНИЮ СИСТЕМ

В этой части нам предстоит ознакомиться с некоторыми применениями определителей и матриц; в частности, в формуле обратной матрицы проявится косвенная связь понятия определителя с матричным перемножением – точнее, вычислением обратной матрицы. Еще один интересный момент – это использование матриц и определителей для решения систем линейных алгебраических уравнений. Сначала вернемся к понятию алгебраического дополнения, определение п. 2.4.8 которого возникло из формул пп. 2.4.6 – 2.4.8.

2.5.1. Алгебраическое дополнение элемента квадратной матрицы. Предположим, что нам дана матрица A порядка $m\times m$.

Квадратную матрицу $\tilde{A}_{ij}$ порядка на единицу меньше мы назовем ***дополнительной*** к элементу a_{ij}, если она получается из матрицы A изъятием элемента a_{ij} вместе со всей i-й строкой и всем j-м столбцом и последующим сдвиганием получившихся фрагментов (мы уже имели дело с такими матрицами в п. 2.4.7). Схематически это будет выглядеть так:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1\,j-1} & a_{1j} & a_{1\,j+1} & \cdots & a_{1m} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1\,1} & \cdots & a_{i-1\,j-1} & a_{i-1\,j} & a_{i-1\,j+1} & \cdots & a_{i-1\,m} \\ a_{i1} & \cdots & a_{i\,j-1} & a_{i\,j} & a_{i\,j+1} & \cdots & a_{i\,m} \\ a_{i+1\,1} & \cdots & a_{i+1\,j-1} & a_{i+1\,j} & a_{i+1\,j+1} & \cdots & a_{i+1\,m} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & \cdots & a_{m\,j-1} & a_{mj} & a_{m\,j+1} & \cdots & a_{mm} \end{pmatrix};$$

$$\tilde{A}_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1\,j-1} & & a_{1\,j+1} & \cdots & a_{1m} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i-1\,1} & \cdots & a_{i-1\,j-1} & & a_{i-1\,j+1} & \cdots & a_{i-1\,m} \\ & & & & & & \\ a_{i+1\,1} & \cdots & a_{i+1\,j-1} & & a_{i+1\,j+1} & \cdots & a_{i+1\,m} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{m\,j-1} & & a_{m\,j+1} & \cdots & a_{mm} \end{pmatrix}.$$

Определитель $M_{ij} = \left|\tilde{A}_{ij}\right|$ назовем **дополнительным минором** элемента a_{ij}, поскольку слово «минор» означает «меньший», а матрица $\tilde{A}_{ij}$ у нас действительно уменьшенного формата по сравнению с исходной.

В соответствии с определением п. 2.4.8 алгебраическое дополнение A_{ij} элемента a_{ij} – это либо сам минор $\left|\tilde{A}_{ij}\right|$, если сумма $i + j$

номеров строки и столбца – четная, либо он же с противоположным знаком, если сумма $i+j$ – нечетная, т. е.

$$A_{ij}=(-1)^{i+j}\left|\tilde{A}_{ij}\right|=\begin{cases}\left|\tilde{A}_{ij}\right|, \text{ если } i+j \text{ четно,}\\ -\left|\tilde{A}_{ij}\right|, \text{ если } i+j \text{ нечетно.}\end{cases}$$

Можно сказать, что необходимость смены знака у дополнительного минора или, если угодно, знак «минус» перед минором при переходе к алгебраическому дополнению возникает в шашечном порядке (если знак менять не нужно, на соответствующей клетке мы поставим «плюс»):

$$\begin{pmatrix} + & - & + & - & \ldots \\ - & + & - & + & \\ + & - & + & & \\ - & + & & \ddots & \\ \vdots & & & & \end{pmatrix}.$$

2.5.2. Теперь проиллюстрируем вычисление определителя с помощью формул разложения по строке или столбцу. Пусть задана матрица

$$A=\begin{pmatrix} 1 & 0 & 2 & 0 & 5 \\ 1 & 0 & 0 & -1 & 3 \\ 0 & 1 & 0 & 0 & -1 \\ 4 & 2 & -4 & 0 & -3 \\ 2 & 1 & -1 & 2 & 1 \end{pmatrix}.$$

Чтобы найти ее определитель, разложим его по третьей строке: $|A|=a_{31}A_{31}+a_{32}A_{32}+a_{33}A_{33}+a_{34}A_{34}+a_{35}A_{35}$. Поскольку $a_{31}=a_{33}=$ $=a_{34}=0$, соответствующие алгебраические дополнения можно не вычислять. Фактически в формулу войдут только значения A_{32} и A_{35}, причем $|A|=A_{32}-A_{35}$, так как $a_{32}=-a_{35}=1$.

Дополнительный минор $\left|\tilde{A}_{32}\right|$ удобнее всего разложить по третьему столбцу, тогда получим

$$A_{32} = (-1)^{3+2}\left|\tilde{A}_{32}\right| = -\begin{vmatrix} 1 & 2 & 0 & 5 \\ 1 & 0 & -1 & 3 \\ 4 & -4 & 0 & -3 \\ 2 & -1 & 2 & 1 \end{vmatrix} =$$

$$= -\left(0 + (-1)(-1)^{2+3}\begin{vmatrix} 1 & 2 & 5 \\ 4 & -4 & -3 \\ 2 & -1 & 1 \end{vmatrix} + 0 + 2(-1)^{4+3}\begin{vmatrix} 1 & 2 & 5 \\ 1 & 0 & 3 \\ 4 & -4 & -3 \end{vmatrix} \right) =$$

$$= -(-7-44) = 51.$$

А дополнительный минор $\left|\tilde{A}_{35}\right|$ легче всего разложить по второй строке:

$$A_{35} = (-1)^{3+5}\left|\tilde{A}_{35}\right| = \begin{vmatrix} 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & -1 \\ 4 & 2 & -4 & 0 \\ 2 & 1 & -1 & 2 \end{vmatrix} =$$

$$= 1\,(-1)^{1+2}\begin{vmatrix} 0 & 2 & 0 \\ 2 & -4 & 0 \\ 1 & -1 & 2 \end{vmatrix} + 0 + 0 + (-1)(-1)^{2+4}\begin{vmatrix} 1 & 0 & 2 \\ 4 & 2 & -4 \\ 2 & 1 & -1 \end{vmatrix} =$$

$$= 8 - 2 = 6.$$

В результате получается $|A| = 51–6 = 45$.

2.5.3. Прежде чем обратиться к формуле обратной матрицы, установим еще одно свойство алгебраических дополнений. Непосредственно продолжая разд. 2.4, ответим на такой вопрос: ***чему равна сумма произведений элементов некоторой строки матри-***

цы A на алгебраические дополнения элементов другой ее строки с теми же столбцовыми номерами?

В соответствии с формулой разложения по строке результат получится таким же, как если бы мы вычисляли определитель

$$\begin{vmatrix} a_{11} & \ldots\ldots & a_{1n} \\ \vdots & & \vdots \\ a_{j-1\,1} & & a_{j-1\,n} \\ a_{i1} & & a_{in} \\ a_{j+1\,1} & & a_{j+1\,n} \\ \vdots & & \vdots \\ a_{n1} & \ldots\ldots & a_{nn} \end{vmatrix}$$

разложением по его j-й строке.

Но в этом определителе две строки одинаковы: i-я, стоящая на своем «законном» месте, и такая же строка, – но на j-м месте. Такой определитель равен нулю. Следовательно, справедливо равенство

$$\sum_{k=1}^{n} a_{ik}A_{jk} = 0 \quad (i \neq j).$$

Конечно, аналогичное утверждение верно и для столбцов.

2.5.4. Пусть дана квадратная матрица A. Составим матрицу A', элементами которой являются алгебраические дополнения A_{ij} элементов исходной матрицы, и транспонируем ее. Назовем полученную матрицу ***присоединенной матрицей*** и обозначим ее A^*. Теперь попробуем вычислить произведение матрицы A на присоединенную матрицу A^*. Пусть $B = A^* = A'^{\mathrm{T}}$, тогда каждый элемент b_{ij} матрицы B равен A_{ji}. Если $C = AB$, то по определению матричного умножения $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$, а это равно определителю $|A|$, если $i = j$, и нулю, если $i \neq j$, по пп. 2.4.8 и 2.5.3.

Таким образом, $C = |A|E$.

Если теперь предположить, что $|A| \neq 0$, то $|A|^{-1} AA^{*} = E$ и матрица

$$A^{-1} = |A|^{-1} A^{*} = \frac{1}{|A|}(A_{ji})$$

будет обратной к матрице А (см. **2.1.8**).

Замечание 1. Пока доказано лишь равенство $A\ A'^{\mathrm{T}} = |A| E$, а определение обратной матрицы требует выполнения и второго равенства с другим порядком сомножителей: $A'^{\mathrm{T}} A = |A| E$.

Попробуйте доказать это равенство как столбцовый вариант первого.

Замечание 2. А что будет, когда $|A| = 0$? Оказывается, в этом случае матрица A не имеет обратной. Это можно доказать, используя понятие ***ранга*** матрицы, которое будет рассмотрено нами в пункте **7.3.2** в связи с методом Гаусса.

Так мы приходим к теореме об обратной матрице:

если матрица А невырожденная: $|A| \neq 0$***, то обратная к ней матрица*** A^{-1} ***существует и определяется по формуле*** $A^{-1} = \frac{1}{|A|} A'^{\mathrm{T}}$***, а вырожденная матрица необратима, т. е. для нее не существует обратной матрицы*** A^{-1}.

Иными словами***, матрица обратима тогда и только тогда, когда ее определитель отличен от*** **0.** Заметим, что это отчасти оправдывает само название «определитель».

2.5.5. Формулы Крамера решения систем линейных уравнений. Определения системы линейных алгебраических уравнений и ее решения довольно просты, и можно посмотреть их уже сейчас в п. 6.2.1. Здесь мы ограничимся случаем, когда число уравнений равняется числу неизвестных: $m = n$. Тогда матрица $A = (a_{ij})$, составленная из коэффициентов левых частей уравнений, будет квадратной (она называется ***матрицей системы***).

Второе ограничение состоит в предположении, что $\det A$ отличен от нуля. Следовательно, матрица системы обратима.

Как будет сделано и позже в п. 6.2.2, мы образуем одностолбцовые матрицы неизвестных и свободных членов:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}; \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Тогда, основываясь на матричном умножении, мы сможем записать исходную систему уравнений в матричной форме:

$$\begin{pmatrix} a_{11} \dots a_{1n} \\ \vdots \\ a_{n1} \dots a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

или проще: $AX = B$.

Разумеется, раскрытие произведения AX даст в точности левые части уравнений так, как они записаны в определении системы.

Преимущество матричной записи в том, что сейчас мы можем домножить обе части системы слева на матрицу A^{-1}. Тогда в левой части мы получим $A^{-1}(AX) = (A^{-1}A)X = EX = X$, в правой – $A^{-1}B$, т. е. столбец неизвестных X оказался равен числовой одностолбцовой матрице $A^{-1}B$, что и равносильно нахождению решения системы, к тому же с гарантией его единственности.

Посмотрим внимательнее на i-й этаж равенства $X = A^{-1}B$, i-я строка матрицы $A^{-1}B$ получается из умножения на столбец B i-й строки матрицы A^{-1}. В свою очередь, эта строка состоит из алгебраических дополнений элементов i-го столбца матрицы A, поделенных на $|A|$.

Следовательно, при всех i выполняется равенство

$$x_i = \sum_{k=1}^{n} \frac{1}{|A|} A_{ki} b_k .$$

Осталось заметить, что правая часть здесь равна определителю матрицы:

$$\frac{1}{|A|}A_i = \frac{1}{|A|}\begin{pmatrix} a_{11}\dots & a_{1i-1} & b_1 & a_{1i+1}\dots & a_{1n} \\ \vdots & & & & \\ a_{n1}\dots & a_{ni-1} & b_n & a_{ni+1}\dots & a_{nn} \end{pmatrix}_{n\times n},$$

получающейся из матрицы системы A заменой ее i-го столбца на столбец свободных членов (конечно, мы раскрываем этот определитель по i-му столбцу). В результате мы получили классические ***формулы Крамера*** для решения квадратной системы линейных уравнений

$$x_i = \frac{1}{|A|}\,|A_i|, \qquad i = 1, \dots, n.$$

Или, в классических обозначениях $\Delta_i = |A_i|, \quad \Delta = A$

$$x_i = \frac{\Delta_i}{\Delta}, \qquad i = 1, \dots, n.$$

Как видите, второе предположение: $\Delta \neq 0$ – совершенно явно присутствует в итоговой формулировке.

2.5.6. Напоследок приведем без доказательства еще один результат, показывающий, насколько глубока связь между перемножением самих матриц и их определителей:

определитель произведения квадратных матриц равен произведению их определителей, т. е.

$$|AB| = |A|\,|B|.$$

Отсюда, в частности, следует, что вырожденная матрица A необратима (так как $|A| = 0$ и $|AB| = |A|\,|B| = 0$ для всех матриц B, но $|E| = 1$), что мы отмечали в п. **2.5.4**.

ГЛАВА 3
ВЕКТОРЫ И ЛИНЕЙНЫЕ ПРОСТРАНСТВА

3.1. ГЕОМЕТРИЧЕСКИЕ ВЕКТОРЫ. ЛИНЕЙНЫЕ ОПЕРАЦИИ НАД НИМИ И ИХ СВОЙСТВА

3.1.1. Мы начнем этот раздел с того, что напомним стандартное, хорошо известное из школьного курса математики описание точек пространства (плоскости) с помощью декартовых прямоугольных систем координат.

Под ***декартовой прямоугольной системой координат*** в пространстве понимается система из трех взаимно перпендикулярных числовых осей, пересекающихся в одной точке (начале координат). По традиции, будем обозначать их как Ox, Oy и Oz, здесь O – начало координат – общая точка осей, являющаяся началом отсчета на каждой из них.

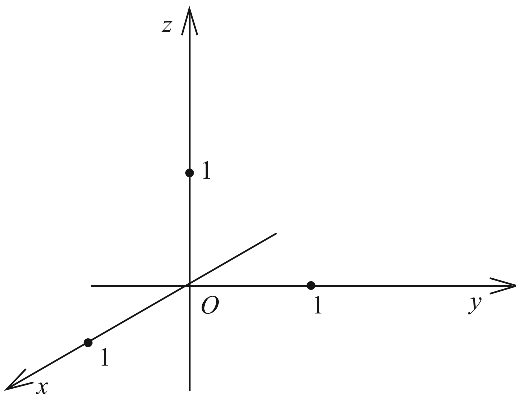


Рис. 11

Прежде всего, существует взаимно однозначное соответствие между точками числовой оси и действительными числами.

Стандартным образом поставив в соответствие произвольной точке A пространства числа, равные ее проекциям на оси Ox, Oy, Oz, мы получим взаимно однозначное соответствие между точками пространства и упорядоченными тройками действительных чисел $(x_0;\ y_0;\ z_0)$ – ***координатами точки в данной декартовой прямоугольной системе координат*** (рис. 12).

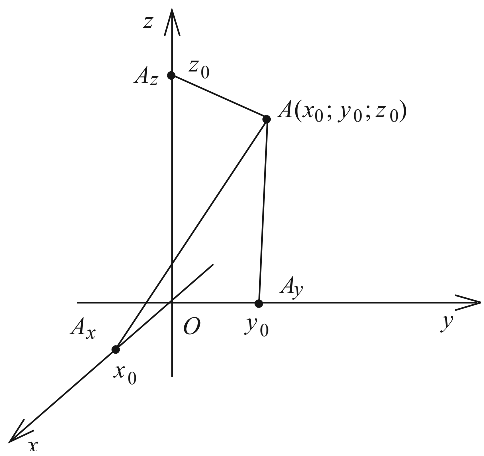


Рис. 12

3.1.2. Напомним также, что длина отрезка AB, где $(x_0;\ y_0;\ z_0)$, $(x_1;\ y_1;\ z_1)$ – координаты точек A и B соответственно, находится из теоремы Пифагора:

$$|AB| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}\ .$$

3.1.3. Традиционно под ***геометрическим вектором*** понимается направленный отрезок (на прямой, на плоскости, в пространстве), т. е. отрезок, один из концов которого определен как начало вектора, а другой – как его конец (рис. 13).

Для вектора с началом A и концом B будем пользоваться обозначением $\overline{AB}$; кроме того, геометрические векторы будем обозначать также символами $\overline{a},\ \overline{b},\ \overline{c},\ldots$

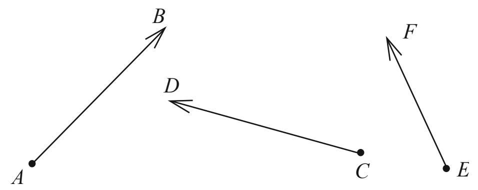


Рис. 13

Мы будем считать два вектора $\overline{a}$ и $\overline{b}$ ***равными***, если вектор $\overline{a}$ может быть совмещен с вектором $\overline{b}$ параллельным переносом, т. е. если:

а) равны длины векторов $\overline{a}$ и $\overline{b}$;

б) векторы $\overline{a}$ и $\overline{b}$ параллельны как отрезки;

в) совпадают направления векторов $\overline{a}$ и $\overline{b}$ (рис. 14).

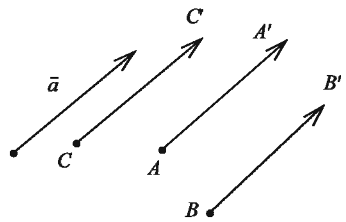


Рис. 14

Из этого определения вытекают два важнейших принципа:

а) возможность отождествления геометрического вектора с параллельным сдвигом пространства (каждой точке A, B,... пространства ставится в соответствие некоторая новая точка A', B',... такая, что $\overline{AA'} = \overline{a}$, $\overline{BB'} = \overline{a}$ и т. д.);

б) возможность считать данный вектор $\overline{a}$ выходящим из любой точки пространства, в частности из начала декартовой прямоугольной системы координат (рис. 15).

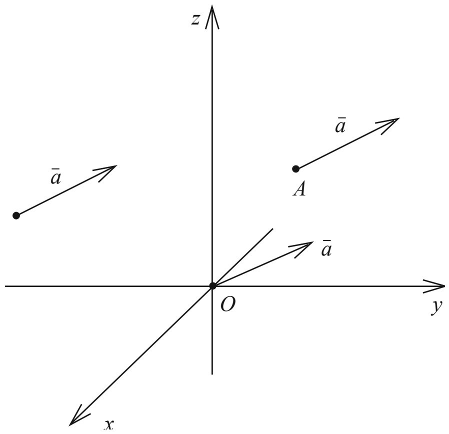


Рис. 15

Заметьте, что векторы могут быть разными, но равными.

3.1.4. Декартовы прямоугольные системы координат позволяют получить аналитическое (числовое) описание геометрических векторов.

Если l – некоторая числовая ось в пространстве и $\overline{AB}$ – произвольный вектор, то под ***проекцией*** $\text{пр}_l\overline{AB}$ ***вектора*** $\overline{AB}$ на ось l мы будем понимать длину отрезка $A'B'$, где A' и B' – проекции точек A и B на ось l (основания перпендикуляров, опущенных из точек A и B на ось l), взятую со знаком «плюс», если направление вектора $\overline{A'B'}$ совпадает с направлением оси l, и со знаком «минус» – в противном случае (рис. 16).

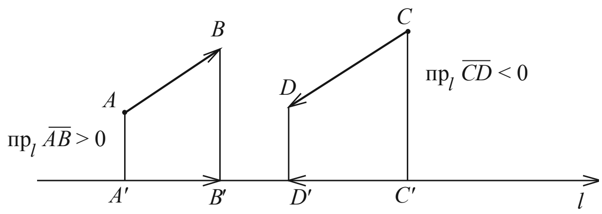


Рис. 16

Рассматривая все случаи, связанные со знаками проекций векторов на ось, для любой оси l и любых трех точек пространства можно установить формулу

$$\text{пр}_l\overline{AC} = \text{пр}_l\overline{AB} + \text{пр}_l\overline{BC}.$$

3.1.5. Под ***координатами вектора*** $\overline{a}$ в декартовой прямоугольной системе координат понимается упорядоченная тройка чисел $(\text{пр}_{Ox}\overline{a},\ \text{пр}_{Oy}\overline{a},\ \text{пр}_{Oz}\overline{a})$. Из этого определения следует, что если точки A и B в фиксированной декартовой прямоугольной системе координат имеют координаты $(x_0;\ y_0;\ z_0)$ и $(x_1;\ y_1;\ z_1)$ соответственно, то вектор $\overline{AB}$ в этой же системе координат имеет координаты $(x_1 - x_0;\ y_1 - y_0;\ z_1 - z_0)$. Отметим, в частности, что если начало этого вектора (точка A) совпадает с началом координат (точкой O), то координаты вектора $\overline{AB} = \overline{OB}$ совпадают с координатами его конца (точки B):

$$(x_1 - 0; y_1 - 0; z_1 - 0) = (x_1; y_1; z_1).$$

Точно так же мы получаем возможность решить вопрос о равенстве векторов $\overline{a}$ и $\overline{b}$, заданных своими координатами $(x_1;\ y_1;\ z_1)$ и $(x_2;\ y_2;\ z_2)$ соответственно:

$$\overline{a} = \overline{b} \Leftrightarrow x_1 = x_2,\ y_1 = y_2,\ z_1 = z_2.$$

Это свойство позволяет нам в дальнейшем писать равенство $\overline{a} = (x_1;\ y_1;\ z_1)$, когда числа $(x_1;\ y_1;\ z_1)$ являются координатами вектора $\overline{a}$ в фиксированной декартовой системе координат.

3.1.6. Благодаря теореме Пифагора координаты вектора $\overline{a} = (x_1;\ y_1;\ z_1)$ позволяют вычислить его длину:

$$|\overline{a}| = \sqrt{x_1^2 + y_1^2 + z_1^2},$$

в частности, для вектора $\overline{AB}$, если $A\ (x_1;\ y_1;\ z_1)$ – его начало, $B\ (x_2;\ y_2;\ z_2)$ – конец, то

$$\left|\overline{AB}\right| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Аналогично, по теореме Фалеса векторы $\overline{a}$ и $\overline{b}$ параллельны тогда и только тогда, когда их координаты пропорциональны: $\frac{x_1}{x_2} = \frac{y_1}{y_2} = \frac{z_1}{z_2}$, причем мы допускаем одновременное обращение в нуль и числителя, и знаменателя в одной или даже в двух из этих дробей.

3.1.7. Основную роль при работе с векторами играют так называемые ***линейные операции*** над ними: ***сложение векторов*** и ***умножение их на действительные числа***.

Под ***суммой*** $\overline{a} + \overline{b}$ ***векторов*** $\overline{a}$ и $\overline{b}$ понимается вектор, начало которого совпадает с началом вектора $\overline{a}$, а конец – с концом вектора $\overline{b}$ при условии, что конец вектора $\overline{a}$ совпадает с началом вектора $\overline{b}$ (правило треугольника, рис. 17).

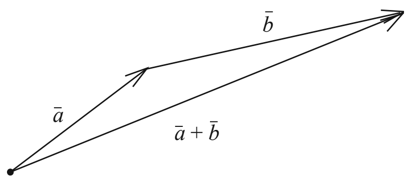


Рис. 17

Если в некоторой декартовой прямоугольной системе координат векторы $\overline{a}$ и $\overline{b}$ имеют координаты $(x_1;\, y_1;\, z_1)$ и $(x_2;\, y_2;\, z_2)$ соответственно, то в силу формулы в конце п. 3.1.4 вектор $\overline{a}$ + $\overline{b}$ будет иметь координаты $(x_1 + x_2;\, y_1 + y_2;\, z_1 + z_2)$, т. е.

$$(x_1;\, y_1;\, z_1) + (x_2;\, y_2;\, z_2) = (x_1 + x_2;\, y_1 + y_2;\, z_1 + z_2)\,.$$

3.1.8. Отметим некоторые простейшие свойства операции сложения векторов, которые потребуются нам в дальнейшем:

1) $\overline{a} + \overline{b} = \overline{b} + \overline{a}$ (коммутативность сложения);

2) $(\overline{a} + \overline{b}) + \overline{c} = \overline{a} + (\overline{b} + \overline{c})$ (ассоциативность сложения).

Докажем лишь первое из этих свойств, оставляя доказательство второго читателю. Пусть, как и выше, $(x_1;\, y_1;\, z_1)$ и $(x_2;\, y_2;\, z_2)$ – координаты векторов $\overline{a}$ и $\overline{b}$ соответственно, тогда

$$\overline{a} + \overline{b} = (x_1;\, y_1;\, z_1) + (x_2;\, y_2;\, z_2) = (x_1 + x_2;\, y_1 + y_2;\, z_1 + z_2) =$$

$$= (x_2 + x_1;\, y_2 + y_1;\, z_2 + z_1) = (x_2;\, y_2;\, z_2) + (x_1;\, y_1;\, z_1) = \overline{b} + \overline{a}\,.$$

Таким образом, коммутативность сложения векторов есть прямое следствие коммутативности сложения действительных чисел и определений пп. 3.1.5 и 3.1.7.

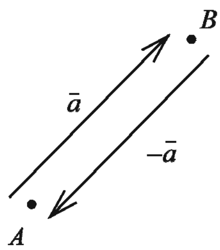


Рис. 18

Через $\overline{0}$ обозначим вектор, у которого начало и конец совпадают: $\overline{0} = \overline{AA} = \overline{BB} = \ldots$ В силу определения равенства векторов все нулевые векторы равны между собой, а их направление можно считать произвольным или неопределенным. Через $-\overline{a}$ обозначим вектор, начало которого совпадает с концом вектора $\overline{a}$, а конец – с началом $\overline{a}$ (если $\overline{a} = \overline{AB}$, то $-\overline{a} = \overline{BA}$, рис. 18).

Очевидно, что в любой декартовой прямоугольной системе координат вектор $\overline{0}$ имеет координаты (0; 0; 0) и, если $(x_1;\ y_1;\ z_1)$ – координаты вектора $\overline{a}$, то $(-x_1;\ -y_1;\ -z_1)$ – координаты вектора $-\overline{a}$.

Легко заметить справедливость следующих свойств:

3) $\overline{a}+\overline{0}=\overline{a}$ (свойство нуля);

4) $\overline{a}+(-\overline{a})=\overline{0}$ (свойство противоположного элемента).

3.1.9. Напомним теперь определение **умножения вектора на число**: любому вектору $\overline{a}$ и любому действительному числу λ ставится в соответствие новый вектор $\lambda\overline{a}$, определяемый следующими правилами:

а) длина вектора $\lambda\overline{a}$ в $|\lambda|$ раз больше длины вектора $\overline{a}$:

$$|\lambda\overline{a}|=|\lambda||\overline{a}|;$$

б) векторы $\lambda\overline{a}$ и $\overline{a}$ параллельны: $\lambda\overline{a}\ ||\ \overline{a}$;

в) если $\lambda>0$, то векторы $\lambda\overline{a}$ и $\overline{a}$ направлены одинаково, если же $\lambda<0$, то $\lambda\overline{a}$ и $\overline{a}$ направлены противоположно.

3.1.10. Путем разбора возможных случаев ($\lambda>0$, $\lambda=0$, $\lambda<0$) для любых вектора $\overline{a}$, числа λ и оси l из подобия треугольников устанавливается следующая формула:

$$\text{пр}_l(\lambda\overline{a})=\lambda\text{пр}_l\overline{a}.$$

В силу этой формулы, если $(x_1,\ y_1,\ z_1)$ – координаты вектора $\overline{a}$ в некоторой декартовой прямоугольной системе координат, то для любого числа λ координатами вектора $\lambda\overline{a}$ будет тройка чисел $(\lambda x_1;\ \lambda y_1;\ \lambda z_1)$, т. е.

$$\lambda(x_1;\ y_1;\ z_1)=(\lambda x_1;\ \lambda y_1;\ \lambda z_1).$$

На основе этой формулы и соответствующих свойств операций сложения и умножения чисел доказываются следующие свойства сложения векторов и умножения их на числа:

5) $\lambda(\mu\overline{a})=(\lambda\mu)\overline{a}$ (ассоциативность);

6) $1\overline{a}=\overline{a}$ (унитарность);

7) $\lambda(\overline{a}+\overline{b})=(\lambda\overline{a})+(\lambda\overline{b})$ (дистрибутивность сложения векторов относительно умножения на числа);

8) $(\lambda+\mu)\overline{a}=\lambda\overline{a}+\mu\overline{a}$ (дистрибутивность сложения чисел относительно умножения векторов на числа).

3.1.11. Отметим еще несколько простейших свойств введенных линейных операций над векторами, которые могут быть доказаны и непосредственно исходя из определения этих операций, и как логические следствия свойств 1 – 8:

$$0\,\overline{a}=\overline{0},$$

$$\lambda\cdot\overline{0}=\overline{0},$$

$$(-1)\overline{a}=-\overline{a}.$$

3.2. АКСИОМАТИКА ЛИНЕЙНЫХ ПРОСТРАНСТВ И ПРОСТЕЙШИЕ СЛЕДСТВИЯ АКСИОМ

3.2.1. Указанные в пп. 3.1.8 и 3.1.10 свойства 1 – 8 операций сложения геометрических векторов и умножения их на действительные числа лежат в основе определения абстрактного линейного пространства L над произвольным полем **F**. Наличие таких свойств для линейных операций над объектами любого вида позволяет при работе с этими объектами использовать интуицию, развитую при работе с наглядными объектами – геометрическими векторами.

Пусть **F** – некоторое поле (числовое поле, см. 1.3.10) и L – произвольное множество, на котором заданы следующие операции:

а) двухместная операция сложения +, когда любой паре $\langle a, b\rangle$ элементов множества L ставится в соответствие некоторым образом определённый элемент этого же множества L, обозначаемый как $a+b$;

б) одноместные операции умножения элементов множества L на числа $\lambda\in\mathbf{F}$, когда любому элементу a из L и любому числу λ из **F** ставится в соответствие некоторый элемент из L, обозначаемый как λa (рис. 19).

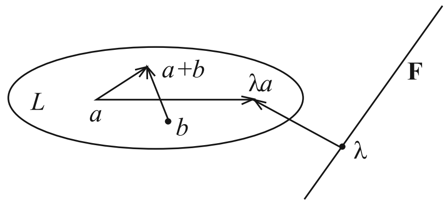


Рис. 19

3.2.2. Определение. Мы будем говорить, что множество L является **линейным пространством (векторным пространством) над полем F**, если операции сложения элементов и умножения их на число $\lambda \in \mathbf{F}$ обладают следующими свойствами. Для любых $a, b \in L$ и $\lambda, \mu \in \mathbf{F}$:

1) $a + b = b + a$;

2) $(a + b) + c = a + (b + c)$;

3) существует такой элемент в L (обозначим его как $\underline{0}$ – с подчёркиванием снизу), что для любого a из L

$$a + \underline{0} = a\,;$$

4) для любого элемента $a \in L$ существует такой элемент в L, обозначаемый как $-a$, что

$$a + (-a) = \underline{0}\,,$$

5) $\lambda(\mu a) = (\lambda\mu)a$;

6) $1\,a = a$;

7) $\lambda(a + b) = \lambda a + \lambda b$;

8) $(\lambda + \mu)a = \lambda a + \mu a$.

Элементы линейного (векторного) пространства иногда будем называть векторами, если их нельзя будет спутать с традиционными геометрическими векторами.

3.2.3. Приведенные здесь свойства 1 – 8 носят название аксиом линейного пространства. Отметим, что если аксиомы 1, 2, 5, 7 и 8 носят универсальный характер, т. е. содержат требования, относящиеся ко всем элементам линейного пространства L и всем элементам поля **F**, то аксиомы 3 и 4 включают в себя утверждения о существовании тех или иных элементов в L, обладающих некоторыми указанными свойствами. Причем аксиома 3 утверждает существование, по крайней мере, одного элемента $\underline{0}$ (называемого далее ***ну-***

левым элементом пространства L или просто нулем), в то время как аксиома 4 утверждает существование элемента $-a$, и такие элементы различны для различных элементов a из L. Элемент $-a$ будем далее называть ***противоположным*** элементу a.

3.2.4. Приведем ряд примеров линейных пространств.

1. Пусть L – совокупность геометрических векторов, рассматриваемых с точностью до равенства; $\mathbf{F} = \mathbf{R}$ – поле действительных чисел.

2. Пусть L – это совокупность $\mathbf{F}_{nm}$ всех матриц размера n×m над некоторым полем $\mathbf{F}$. Если рассматривать стандартные операции сложения матриц и умножения их на числа, то в силу свойств этих операций все аксиомы 1 – 8 линейного пространства оказываются выполненными, и тем самым в этом случае мы также имеем дело с линейным пространством над полем $\mathbf{F}$. Роль нулевого элемента $\underline{0}$ будет при этом играть матрица, состоящая из нулей поля $\mathbf{F}$, а роль противоположной $-A$ матрицы для матрицы A играет матрица, состоящая из чисел, противоположных числам, заполняющим матрицу A (подробнее см. п. 2.1.2).

Важную роль в дальнейшем будет играть линейное пространство $\mathbf{F}^m$ строк (т. е. матриц размера $1{\times}m$) или столбцов (т. е. матриц размера $m{\times}1$) длины m, составленных из элементов поля $\mathbf{F}$.

3. Обозначим через $\mathbf{R}^{\mathbf{R}}$ совокупность действительнозначных функций действительного аргумента, т. е. отображений $f\colon \mathbf{R} \to \mathbf{R}$ множества действительных чисел в себя самое. Обычным образом определим операции сложения этих функций и умножения их на числа:

- для $f,\ g \in \mathbf{R}^{\mathbf{R}}$ функция $(f+g) \in \mathbf{R}^{\mathbf{R}}$ определена следующим образом:

$$(f+g)(\alpha) = f(\alpha) + g(\alpha)$$

для любого $\alpha \in \mathbf{R}$;

- для функций $f \in \mathbf{R}^{\mathbf{R}}$ и $\beta \in \mathbf{R}$ функция $(\beta f) \in \mathbf{R}^{\mathbf{R}}$ определена следующим образом:

$$(\beta f)(\alpha) = \beta \cdot f(\alpha)$$

для любого $\alpha \in \mathbf{R}$.

Легко видеть справедливость для этих операций на $\mathbf{R}^{\mathbf{R}}$ всех аксиом 1–8 – все следует из соответствующих свойств чисел. При этом роль нулевого элемента играет нулевая функция, т. е. $\underline{0}(\alpha) = 0$ для любого $\alpha \in \mathbf{R}$, а противоположный элемент $-f$ для функции $f \in \mathbf{R}^{\mathbf{R}}$ определяется равенством $(-f)(\alpha) = -f(\alpha)$ для любого $\alpha \in \mathbf{R}$.

В случае $L = \mathbf{R}^{\mathbf{R}}$ будем в дальнейшем говорить о пространстве функций.

4. Заметим, наконец, что любое поле **F** может быть рассмотрено как линейное пространство над самим собой. Выполнение аксиом 1–8 для операций сложения и умножения на нем как на линейном пространстве автоматически следует из выполнения в **F** аксиом поля (см. п. 1.3.10).

Прервем здесь цепь примеров линейных пространств, хотя этим они далеко не исчерпываются.

3.2.5. В заключение этого раздела докажем ряд простейших следствий из аксиом линейного пространства, необходимых нам в дальнейшем.

Прежде всего покажем, что нулевой элемент пространства всегда единственный, т. е. что лишь один элемент $\underline{0}$ может удовлетворять требованиям аксиомы 3. Действительно, если $\underline{0}_1$ и $\underline{0}_2 \in L$ таковы, что для любого $a \in L$ имеет место

$$a + \underline{0}_1 = a \quad \text{и} \quad a + \underline{0}_2 = a,$$

то $\underline{0}_1 = \underline{0}_1 + \underline{0}_2 = \underline{0}_2 + \underline{0}_1 = \underline{0}_2$.

Покажем, что противоположный элемент $-a$ также однозначно определяется для каждого элемента a из L. Действительно, если $(-a)_1$, $(-a)_2$ таковы, что $a + (-a)_1 = \underline{0}$ и $a + (-a)_2 = \underline{0}$, то

$$(-a)_1 = (-a)_1 + \underline{0} = (-a)_1 + (a + (-a)_2) = ((-a)_1 + a) + (-a)_2 =$$

$$= (a + (-a)_1) + (-a)_2 = \underline{0} + (-a)_2 = (-a)_2 + \underline{0} = (-a)_2.$$

Далее вместо $a + (-b)$ будем использовать запись $a - b$.

Отметим, что в ходе последних доказательств многократно использовались свойства коммутативности и ассоциативности сложения элементов из L.

3.2.6. Докажем еще ряд простейших свойств линейных пространств, непосредственно вытекающих из аксиом 1–8.

Для любого элемента $a \in L$ и числа $\alpha \in \mathbf{F}$ имеют место равенства:

а) $0a = \underline{0}$,

б) $\alpha \underline{0} = \underline{0}$,

в) $(-\alpha)a = \alpha(-a) = -(\alpha a)$.

Действительно:

а) $0a = (0+0)a = 0a + 0a$.

Прибавив к обеим частям равенства $0a = 0a + 0a$ по элементу $-(0a)$, получим $0a + (-(0a)) = 0a$. А так как $0a + (-(0a)) = \underline{0}$, то и имеем искомое: $0a = \underline{0}$.

б) $\alpha \cdot \underline{0} = \alpha(\underline{0} + \underline{0}) = \alpha \cdot \underline{0} + \alpha \cdot \underline{0}$.

Прибавив к обеим частям равенства $\alpha \cdot \underline{0} = \alpha \cdot \underline{0} + \alpha \cdot \underline{0}$ по элементу $-(\alpha \cdot \underline{0})$, получим $\alpha \cdot \underline{0} + (-(\alpha \cdot \underline{0})) = \alpha \cdot \underline{0}$. А так как $\alpha \cdot \underline{0} + (-(\alpha \cdot \underline{0})) = \underline{0}$, то действительно $\alpha \cdot \underline{0} = \underline{0}$.

в) $\alpha a + (-\alpha)a = (\alpha + (-\alpha))a = 0a = \underline{0}$

и, значит, $(-\alpha)a$ есть единственный (в силу доказанного выше) противоположный элемент к элементу αa, т. е. $(-\alpha)a = -(\alpha a)$.

В силу только что доказанного $(-1)a = -(1a) = -a$ и тем самым $\alpha(-a) = \alpha((-1)a) = (\alpha(-1))a = (-\alpha)a$.

Аксиомы 1–8 линейных пространств гарантируют возможность таких действий:

- раскрытие скобок,
- приведение подобных,
- перенесение выражений из одной части равенства в другую с противоположным знаком и т. д., т. е. гарантируют в линейных пространствах тот «комфорт» при работе с символьными выражениями, к которому мы привыкли, работая с символами, соответствующими действительным числам.

3.3. ЛИНЕЙНАЯ НЕЗАВИСИМОСТЬ. БАЗИС И КООРДИНАТЫ. ИЗОМОРФИЗМ. ПОДПРОСТРАНСТВА

Основная цель настоящего раздела – создание механизма, позволяющего заменить работу с элементами линейного пространства, т. е. с объектами произвольного, неопределенного вида, на действия с числами и наборами чисел, т. е. с объектами, нам достаточно привычными. В итоге мы увидим, что изучение так называемых конечномерных линейных пространств сводится к изучению вполне наглядных пространств строк или столбцов, введенных в конце примера 2 разд. 3.2.4.

Главным в этом механизме будет понятие ***базиса пространства***. Однако предварительно нам понадобится еще несколько новых понятий.

3.3.1. ***Линейной комбинацией*** элементов $a_1,..., a_n$ линейного пространства L называется любое выражение вида $\alpha_1 a_1 + ... + \alpha_n a_n$, где $\alpha_1,..., \alpha_n$ – произвольные числа из того поля **F**, над которым рассматривается линейное пространство L. Числа α_i будем называть коэффициентами линейной комбинации.

3.3.2. Определение. Элементы $a_1,..., a_n$ пространства L назовем ***линейно независимыми***, если из равенства $\alpha_1 a_1 + ... + \alpha_n a_n = \underline{0}$ следуют равенства $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$ для произвольных чисел $\alpha_1,...,\alpha_n \in \mathbf{F}$. Или, что то же самое, равенство $\alpha_1 a_1 + ... + \alpha_n a_n = \underline{0}$ возможно в единственном случае, когда $\alpha_1 = ... = \alpha_n = 0$ – такая линейная комбинация называется тривиальной.

Прежде всего заметим, что ***любая линейно независимая система элементов пространства*** L ***не содержит нулевого элемента*** $\underline{0}$**.** Действительно, если $a_1 = \underline{0}$, $a_2,..., a_n$ – произвольная система элементов из L, содержащая $\underline{0}$, то равенство

$$1a_1 + 0a_2 + ... + 0a_n = \underline{0} + \underline{0} + ... + \underline{0} = \underline{0}$$

показывает, что система $a_1,..., a_n$ линейно независимой не будет: ведь $\alpha_1 = 1 \neq 0$.

3.3.3. Легко видеть, что элементы $a_1,..., a_n$ линейно независимы тогда и только тогда, когда любой элемент линейного пространства L, представимый в виде линейной комбинации элементов $a_1,..., a_n$, выразим в этом виде однозначно, т. е. когда для любых $\alpha_1,...,\alpha_n$, $\beta_1,...,\beta_n \in \mathbf{F}$ из равенства линейных комбинаций

$$\alpha_1 a_1 + ... + \alpha_n a_n = \beta_1 a_1 + ... + \beta_n a_n$$

следует равенство коэффициентов $\alpha_1 = \beta_1,..., \alpha_n = \beta_n$.

Действительно, если элементы $a_1,..., a_n$ линейно независимы, то из равенства линейных комбинаций получаем, что

$$(\alpha_1 a_1 + ... + \alpha_n a_n) - (\beta_1 a_1 + ... + \beta_n a_n) = \underline{0},$$

т. е. $(\alpha_1 - \beta_1)a_1 + ... + (\alpha_n - \beta_n)a_n = \underline{0}$, и в силу линейной независимости получаем равенства

$$\alpha_1 - \beta_1 = ... = \alpha_n - \beta_n = 0.$$

Таким образом, из равенства линейных комбинаций действительно вытекают равенства $\alpha_1 = \beta_1,..., \alpha_n = \beta_n$.

В свою очередь, свойство однозначности представления влечет линейную независимость элементов $a_1,..., a_n$, так как из равенства линейных комбинаций

$$\alpha_1 a_1 + ... + \alpha_n a_n = \underline{0} = 0\, a_1 + ... + 0\, a_n$$

непосредственно получаем равенства $\alpha_1 = 0,..., \alpha_n = 0$.

3.3.4. Мы будем говорить о ***линейной зависимости*** элементов $a_1,..., a_n$ в случае, если система элементов $a_1,..., a_n$ не является линейно независимой, т. е. когда существует набор чисел $\alpha_1,..., \alpha_n$, из которых хотя бы одно отлично от нуля, но таких, что

$$\alpha_1 a_1 + ... + \alpha_n a_n = \underline{0}.$$

3.3.5. Определение. Элементы $a_1,..., a_n$ пространства L назовем **базисом** пространства L, если $a_1,..., a_n$ максимальная (по включению) линейно независимая система элементов пространства L; иначе говоря, если элементы $a_1,..., a_n$ линейно независимы, но расширение $a_1,..., a_n, a_{n+1}$ этой системы путем добавления к ней любого элемента a_{n+1} оказывается линейно зависимым.

3.3.6. Теорема. ***Упорядоченная система элементов*** $a_1,..., a_n$ ***линейного пространства*** L ***является базисом тогда и только тогда, когда любой элемент пространства*** L ***однозначно представим в виде линейной комбинации элементов*** $a_1,..., a_n$. ***Иначе говоря, если для любого элемента*** $b \in L$ ***найдутся числа*** $\alpha_1,..., \alpha_n \in \mathbf{F}$ ***такие, что*** $b = \alpha_1 a_1 + ... + \alpha_n a_n$, ***и эти числа определены однозначно***.

Доказательство. Пусть система $a_1,..., a_n$ является базисом пространства L и b – произвольный элемент пространства L. В силу определения базиса расширенная система $a_1,..., a_n, b$ уже линейно зависима, т. е. найдутся такие числа $\alpha_1,...,\alpha_n, \beta \in \mathbf{F}$, не все равные нулю, что $\alpha_1 a_1 + ... + \alpha_n a_n + \beta b = \underline{0}$.

Заметим, что условие $\beta \neq 0$ заведомо выполнено, так как иначе $\beta b = \underline{0}$, и мы имеем равенство

$$\alpha_1 a_1 + ... + \alpha_n a_n + \underline{0} = \alpha_1 a_1 + ... + \alpha_n a_n = \underline{0},$$

причем хотя бы одно из оставшихся чисел $\alpha_1,..., \alpha_n$ отлично от нуля. Последнее же противоречит линейной независимости системы $a_1,..., a_n$. Итак, действительно, $\beta \neq 0$. Тогда $-\beta b = \alpha_1 a_1 + ... + \alpha_n a_n$ и, умножив обе части последнего равенства на число $(-\beta)^{-1}$, получим искомое представление элемента b в виде линейной комбинации элементов $a_1,..., a_n$:

$$b = (-\beta)^{-1}(-\beta b) = (-\beta)^{-1}(\alpha_1 a_1 + ... + \alpha_n a_n) =$$

$$= \left(\alpha_1(-\beta)^{-1}\right)a_1 + ... + \left(\alpha_n(-\beta)^{-1}\right)a_n.$$

В силу же п. 3.3.3 линейная независимость элементов $a_1,..., a_n$ влечет однозначность представления элемента $b \in L$ в виде линейной комбинации базисных элементов.

Покажем теперь обратное следование. Пусть любой элемент однозначно представим в виде линейной комбинации элементов некоторой системы $a_1,..., a_n$. Докажем, что эта система есть базис пространства L.

Линейная независимость $a_1,..., a_n$ следует из однозначности представления нулевого элемента $\underline{0}$ в виде линейной комбинации элементов $a_1,..., a_n$ – посмотрите окончание пункта 3.3.3.

Покажем теперь, что система $a_1,..., a_n$ не содержится ни в какой большей линейно независимой системе, т. е. что она максимальна по включению среди всех линейно независимых систем элементов пространства L. В самом деле, попробуем добавить в нее элемент $b \in L$. Тогда в силу представимости b для некоторых чисел $\beta_1,...,\beta_n \in \mathbf{F}$ имеет место равенство $b = \beta_1 a_1 + ... + \beta_n a_n$ и, значит,

$$\beta_1 a_1 + ... + \beta_n a_n - b = \beta_1 a_1 + ... + \beta_n a_n + (-1)b = \underline{0},$$

т. е. для любого $b \in L$ расширенная система $a_1,..., a_n, b$ линейно зависима. Тем самым теорема доказана полностью.

3.3.7. Итак, если $a_1,..., a_n$ – базис пространства L, то для любого элемента $b \in L$ найдутся, причем единственные, числа $\beta_1,...,\beta_n \in \mathbf{F}$ такие, что $b = \beta_1 a_1 + ... + \beta_n a_n$.

Будем впредь называть это равенство **разложением элемента b по базису** $a_1,..., a_n$, а числа $\beta_1,..., \beta_n$ – **координатами элемента** b в базисе $a_1,..., a_n$.

В силу теоремы 3.3.6, если $a_1,..., a_n$ – базис пространства L, то элемент пространства L однозначно определим своими координатами в этом базисе, т. е. соответствие

$$b \leftrightarrow (\beta_1;...;\ \beta_n),$$

где $\beta_1,..., \beta_n$ – координаты элемента b в базисе $a_1,..., a_n$, является взаимно однозначным соответствием между элементами пространства L и упорядоченными наборами из n чисел (числовыми строками длины n) – элементами пространства $\mathbf{F}^n$.

Это соответствие обладает дополнительными свойствами, позволяющими заменить вычисления, производимые с элементами пространства L, на вычисления с участием чисел – координат элементов пространства L в базисе $a_1,..., a_n$ (т. е. на вычисления в пространстве $\mathbf{F}^n$). Эти свойства таковы:

а) координаты суммы двух элементов равны сумме координат: если $\beta_1,..., \beta_n$ и $\gamma_1,..., \gamma_n$ – координаты в базисе $a_1,..., a_n$ элементов b и c соответственно, то $\beta_1+\gamma_1,..., \beta_n+\gamma_n$ – это координаты элемента $b + c$ в том же базисе;

б) при умножении элемента пространства на некоторое число из поля **F** все координаты элемента в базисе умножаются на это число: если $\beta_1,..., \beta_n$ – координаты элемента b в базисе $a_1,..., a_n$ и $\gamma \in \mathbf{F}$, то $\gamma\beta_1,..., \gamma\beta_n$ – координаты элемента γb в том же базисе.

Справедливость этих утверждений очевидна.

3.3.8. Обобщая отмеченные свойства соответствия

элемент пространства $\leftrightarrow$ *строка координат*,

приходим к следующему определению.

Определение. Два линейных пространства L_1 и L_2 над одним и тем же полем **F** называются **изоморфными**: $L_1 \cong L_2$, если существует взаимно однозначное отображение φ множества L_1 на L_2 такое, что для любых $a_1, a_2 \in L_1$ и любого $\alpha \in \mathbf{F}$

$$\varphi(a_1+a_2)=\varphi(a_1)+\varphi(a_2)\in L_2, \quad \varphi(\alpha a_1)=\alpha\varphi(a_1)\in L_2 .$$

Само же отображение φ называется при этом **изоморфизмом** пространств L_1 и L_2 (рис. 20).

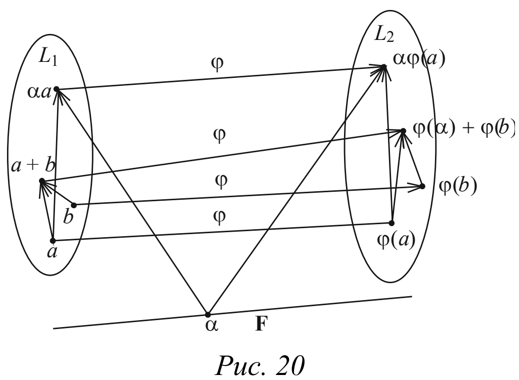


Рис. 20

Фактически изоморфизм линейных пространств означает одинаковое «устройство» этих пространств с точки зрения операций сложения и умножения на число, а также возможность с помощью отображений φ и φ^{-1} переходить при вычислениях из соображений удобства от одного пространства к другому. Теорема 3.3.6 и замечания 3.3.7 фактически утверждают, что отображение, задаваемое формулой $\varphi(b)=(\beta_1;...;\ \beta_n)$, где $\beta_1,..., \beta_n$ – координаты элемента b в фиксированном базисе $a_1,..., a_n$ пространства L, является изоморфизмом пространства L и пространства строк $\mathbf{F}^n$.

Оставим читателю для самостоятельного доказательства следующие простые факты:

а) если φ – изоморфизм пространства L_1 на L_2, то обратное отображение φ^{-1} является изоморфизмом пространства L_2 на L_1;

б) если φ_1 – изоморфизм пространства L_1 на L_2, а φ_2 – пространства L_2 на L_3, то суперпозиция $\varphi_2\varphi_1$ (т. е. последовательное выполнение отображений φ_1 и φ_2) является изоморфизмом пространства L_1 на L_3.

3.3.9. Определение. Число элементов базиса линейного пространства называется его ***размерностью***. Пространства, базис которых состоит из конечного числа элементов, называются ***конечномерными***, причем если этот базис состоит из n элементов, то пространство ***n-мерно***.

Таким образом, в силу отмеченных в п. 3.3.8 изоморфизма любого n-мерного пространства пространству $\mathbf{F}^n$ и свойств а), б) изоморфизмов, получаем, что все n-мерные пространства над некоторым фиксированным полем $\mathbf{F}$ изоморфны между собой, т. е. «устроены» одинаковым образом. При этом важно, чтобы рассматривались пространства над одним и тем же полем $\mathbf{F}$.

3.3.10. Естественным вопросом, возникающим в связи с определением 3.3.9, является вопрос об однозначности определения размерности линейного пространства. Ответ содержится в следующей теореме.

Теорема. ***Размерность линейного пространства определяется однозначно, т. е. если*** $a_1,..., a_n$ ***и*** $b_1,..., b_m$ ***– два базиса пространства*** L, ***то*** $n = m$.

Для доказательства теоремы нужно изучить метод Гаусса, поэтому оно будет приведено в пункте 7.3.1.

3.3.11. Рассмотрим некоторые примеры, связанные с введенными понятиями; при этом интерпретация этих понятий в линейном пространстве геометрических векторов будет рассматриваться в следующем разделе.

1. $L = \mathbf{F}_{nm}$ – совокупность всех матриц порядка $n \times m$ над полем $\mathbf{F}$. Через E_{ij} обозначим матрицу из $\mathbf{F}_{nm}$, состоящую сплошь из нулей, кроме единственного элемента, стоящего в i-й строке, j-м столбце и равного 1 (единице поля $\mathbf{F}$). Проверим линейную независимость совокупности элементов E_{ij} $(1 \le i \le n,\ 1 \le j \le m)$ пространства $\mathbf{F}_{nm}$:

пусть для некоторых чисел $\alpha_{ij} \in \mathbf{F}$

$$\sum_{1\le i\le n,}\sum_{1\le j\le m}\alpha_{ij}E_{ij} = \underline{0} = \begin{pmatrix} 00...0 \\ 00...0 \\ ...... \\ 00...0 \end{pmatrix}.$$

Но по определению матричных операций сложения и умножения на число

$$\sum_{1\le i\le n, 1\le j\le m}\alpha_{ij}E_{ij} = \begin{pmatrix} \alpha_{11}\alpha_{12}...\alpha_{1m} \\ \alpha_{21}\alpha_{22}...\alpha_{2m} \\ ...... \\ \alpha_{n1}\alpha_{n2}...\alpha_{nm} \end{pmatrix}$$

и поскольку у равных матриц на одинаковых местах стоят равные элементы, отсюда следуют равенства $\alpha_{ij} = 0$ для любых $1 \le i \le n$, $1 \le j \le m$. Тем самым матрицы E_{ij} $(1 \le i \le n,\ 1 \le j \le m)$ действительно линейно независимы. С другой стороны, если

$$A = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ & ...... & & \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nm} \end{pmatrix} \in \mathbf{F}_{nm},$$

то очевидно, что $\sum\limits_{1\le i\le n, 1\le j\le m} \beta_{ij} E_{ij} = A$, т. е. любой элемент пространства $\mathbf{F}_{nm}$ представим в виде линейных комбинаций элементов E_{ij} $(1 \le i \le n,\ 1 \le j \le m)$.

Таким образом, в силу теоремы 3.3.6 совокупность E_{ij} $(1 \le i \le n,\ 1 \le j \le m)$ является базисом пространства $\mathbf{F}_{nm}$, а размерность этого пространства равна nm (числу матриц вида E_{ij}).

2. $L = \mathbf{R}^{\mathbf{R}}$ – совокупность всех действительнозначных функций одного действительного аргумента (см. 3.2.4 (3)). Функции x^n $(n \in \mathbf{N})$ входят в $\mathbf{R}^{\mathbf{R}}$. Покажем линейную независимость любой совокупности $\{1,\ x,\ x^2,\ ...,\ x^n\}$.

В самом деле, для любых действительных чисел $\alpha_0,\ \alpha_1, ...,\ \alpha_n$ равенство

$$\alpha_n x^n + \alpha_{n-1} x^{n-1} + ... + \alpha_1 x + \alpha_0 \cdot 1 = \underline{0},$$

где $\underline{0}$ – тождественно нулевая функция, влечет то, что корнями многочлена

$$\alpha_n x^n + \alpha_{n-1} x^{n-1} + ... + \alpha_1 x + \alpha_0$$

являются любые числа. Но последнее возможно лишь в случае, когда все коэффициенты $\alpha_n,\ \alpha_{n-1}, ...,\ \alpha_1,\ \alpha_0$ равны 0 (см. п. 1.4.8).

Итак, система $\{1,\ x,\ x^2,\ ...,\ x^n\}$ действительно линейно независима, а, значит, в силу теоремы 3.3.10 в пространстве $\mathbf{R}^{\mathbf{R}}$ не существует конечной максимальной линейно независимой системы элементов – конечного базиса, т. е. это пространство ***бесконечномерно***.

3.3.12. В дальнейшем при работе с элементами линейного пространства нам будет удобнее пользоваться матричной записью. В частности, если $a_1, ...,\ a_n$ – некоторая система элементов пространства L, то через $\overline{a}$ обозначим строку $(a_1, \dots,\ a_n)$. Если $\alpha_1, ...,\ \alpha_n$ – набор чисел из поля $\mathbf{F}$, то через $\overset{|}{\alpha}$ обозначим столбец

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}.$$

В этом случае линейная комбинация $\alpha_1 a_1 + \dots + \alpha_n a_n$ может быть записана как результат матричного произведения строки $\overline{a}$ и столбца $\overset{|}{\alpha}$, т. е. $\alpha_1 a_1 + \dots + \alpha_n a_n = \overline{a}\,\overset{|}{\alpha}$.

Заметим, что линейная независимость системы элементов $a_1, \dots, a_n$ пространства L влечет то, что для любых числовых столбцов

$$\overset{|}{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}, \quad \overset{|}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix}$$

из равенства $\overline{a}\,\overset{|}{\alpha} = \overline{a}\,\overset{|}{\beta}$ вытекает равенство $\overset{|}{\alpha} = \overset{|}{\beta}$ (т. е. $\alpha_1 = \beta_1, \dots, \alpha_n = \beta_n$), а для любых числовых матриц A, B порядка $n \times m$ из равенства $\overline{a}A = \overline{a}B$ вытекает равенство $A = B$.

3.3.13. Пусть $a_1, \dots, a_n$ и $b_1, \dots, b_n$ – два базиса линейного пространства L. Для произвольного элемента c пространства L найдутся наборы чисел $\alpha_1, \dots, \alpha_n$ и $\beta_1, \dots, \beta_n$ – координаты элемента c в базисах $a_1, \dots, a_n$ и $b_1, \dots, b_n$ соответственно.

Возникает естественный вопрос: как, зная координаты элемента c в базисе $a_1, \dots, a_n$ («старом базисе»), найти координаты элемента c в базисе $b_1, \dots, b_n$ («новом базисе»)? Для решения этого вопроса нам потребуется следующее определение (оно пригодится позже в п. 8.2.1 и далее).

Определение. ***Матрицей перехода*** от базиса $a_1, \dots, a_n$ к базису $b_1, \dots, b_n$ назовем матрицу

$$A=\begin{pmatrix}\alpha_{11}\alpha_{12}...\alpha_{1n}\\ \alpha_{21}\alpha_{22}...\alpha_{2n}\\ ....\\ \alpha_{n1}\alpha_{n2}...\alpha_{nn}\end{pmatrix},$$

где числа α_{ij} являются координатами элемента a_j в базисе $b_1,..., b_n$, т. е. столбцы матрицы A есть координаты векторов «старого» базиса в «новом»:

$$\begin{gathered}a_1=\alpha_{11}b_1+\alpha_{21}b_2+...+\alpha_{n1}b_n,\\ a_2=\alpha_{12}b_1+\alpha_{22}b_2+...+\alpha_{n2}b_n,\\ ....\\ a_n=\alpha_{1n}\, b_1+\alpha_{2n}b_2+...+\alpha_{nn}b_n,\end{gathered}$$

т. е. $(a_1,..., a_n)=(b_1,..., b_n)A$, или в матричной форме, $\overline{a}=\overline{b}A$.

3.3.14. **Теорема**. 1) ***Если*** $\alpha_1,..., \alpha_n$ и $\beta_1,..., \beta_n$ ***– координаты элемента*** c ***пространства*** L ***в базисах*** $a_1,..., a_n$ ***и*** $b_1,..., b_n$ ***соответственно и*** A ***– матрица перехода от базиса*** $a_1,..., a_n$ ***к базису*** $b_1,..., b_n$***, то*** $\overset{|}{\beta}=A\overset{|}{\alpha}$.

2) ***Если*** A ***– матрица перехода от базиса*** $a_1,..., a_n$ ***к базису*** $b_1,..., b_n$ ***(т. е.*** $\overline{a}=\overline{b}A$), B ***– матрица перехода от базиса*** $b_1,..., b_n$ ***к базису*** $c_1,..., c_n$ ***(т. е.*** $\overline{b}=\overline{c}B$), ***а*** C ***– матрица перехода от базиса*** $a_1,..., a_n$ ***к базису*** $c_1,..., c_n$ ***(т. е.*** $\overline{a}=\overline{c}C$ ***), то*** $C=BA$.

Доказательство. 1) В силу условий теоремы имеют место равенства $c=\overline{a}\overset{|}{\alpha}$, $c=\overline{b}\overset{|}{\beta}$ и $\overline{a}=\overline{b}A$.

Подставляя последнее из этих равенств в первое, получаем, что $c=(\overline{b}A)\overset{|}{\alpha}=\overline{b}(A\overset{|}{\alpha})$ (в силу ассоциативности умножения матриц) и, значит, $\overline{b}\overset{|}{\beta}=\overline{b}(A\overset{|}{\alpha})$. Как замечено в п. 3.3.12, в силу линейной независимости системы $b_1,..., b_n$ последнее равенство влечет $\beta = A\alpha$ и утверждение (1) доказано.

2) Действительно, $\overline{a} = \overline{c}C$; с другой стороны, подставляя в равенство $\overline{a} = \overline{b}A$ вместо строки $\overline{b}$ произведение $\overline{c}B$, получаем $\overline{a} = (\overline{c}B)A = \overline{c}(BA)$. Линейная независимость системы $c_1,..., c_n$ наряду с равенствами $\overline{a} = \overline{c}C$ и $\overline{a} = \overline{c}(BA)$ влечет равенство $C = BA$, что и требовалось доказать.

3.3.15. Следствие. ***Если*** A ***– матрица перехода от базиса*** $a_1,..., a_n$ ***к базису*** $b_1,..., b_n$, ***то определитель*** $|A|$ ***не равен нулю и матрицей перехода от базиса*** $b_1,..., b_n$ ***к базису*** $a_1,..., a_n$ ***будет матрица*** A^{-1}, ***обратная к матрице*** A.

Доказательство. Пусть B – матрица перехода от базиса $b_1,..., b_n$ к базису $a_1,..., a_n$. Так как

$$
\begin{aligned}
a_1 &= 1\,a_1 + 0\,a_2 + ... + 0\,a_n,\\
a_2 &= 0\,a_1 + 1\,a_2 + ... + 0\,a_n,\\
&.................\\
a_n &= 0\,a_1 + 0\,a_2 + ... + 1\,a_n,
\end{aligned}
$$

то единичная матрица

$$
E = \begin{pmatrix} 10...0 \\ 01...0 \\ ...... \\ 00...1 \end{pmatrix}
$$

играет роль матрицы перехода от базиса $a_1,..., a_n$ к самому себе. В силу теоремы 3.3.14 получаем равенство $E = BA$, т. е. B – матрица, обратная к матрице A: $B = A^{-1}$. Вспоминая условие существования обратной матрицы из главы 2, получаем, что $|A| \neq 0$.

3.3.16. Терема 3.3.14 «развязывает нам руки» при проведении вычислений с координатами элементов в разных базисах линейного пространства, т. е. имея координаты элемента в одном базисе и выяснив, что при решении какой-либо задачи удобнее работать в другом базисе, с помощью матрицы перехода мы можем пересчитать координаты нашего элемента в новом, более удобном базисе.

3.3.17. Важнейшую роль в изучении строения линейных пространств и при работе с ними играет понятие подпространства.

Определение. Непустое подмножество M линейного пространства L называется ***подпространством***, если M замкнуто относительно сложения и умножения на числа, т. е. если:

1) для любых a, $b \in M$ элемент $a + b$ также лежит в M;

2) для любого $a \in M$ и любого $\alpha \in \mathbf{F}$ элемент αa лежит в M (рис. 21).

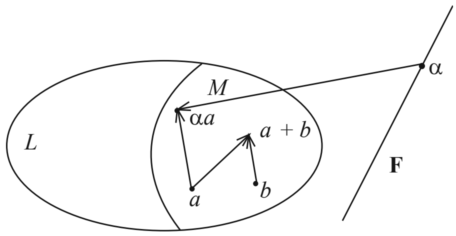


Рис. 21

Прежде всего заметим, что любое подпространство M линейного пространства над полем $\mathbf{F}$ и само является линейным пространством над этим же полем $\mathbf{F}$. Действительно, ограничение операций сложения и умножения на числа из $\mathbf{F}$ с пространства L до множества M в силу свойств 1, 2 определения делает их операциями на множестве M. В п. 3.2.3 отмечалось, что аксиомы 1, 2, 5, 7 и 8 линейных пространств носят универсальный характер, т. е. являются утверждениями, относящимися ко всем элементам линейного пространства L, а значит, и ко всем элементам множества M. Тем самым эти аксиомы выполняются для M с операциями сложения и умножения на числа из $\mathbf{F}$, заимствованными из пространства L. Пусть теперь a – некоторый элемент из M; тогда $\underline{0} = 0a \in M$ и $-a = (-1)a \in M$. Нетрудно заметить, что нулевой элемент $\underline{0}$ пространства L играет роль нулевого элемента и во множестве M, так же как элемент $-a$, противоположный элементу a в пространстве L, играет роль противоположного элемента к элементу a и во множестве M. В силу этого для M оказываются справедливыми и аксиомы 3, 4. Выполнение же аксиом 1–8 для множества M, наделенного операциями сложения и умножения на числа из **F**, и означает, что M является линейным пространством над полем **F**.

3.3.18. Заметим, что в любом пространстве L множество $\{\underline{0}\}$ является подпространством, и так как любое подпространство пространства L содержит элемент $\underline{0}$, то $\{\underline{0}\}$ оказывается наименьшим подпространством пространства L. Точно так же само множество L является своим подпространством. Подпространства пространства L, отличные от L и $\{\underline{0}\}$, будем называть далее ***собственными подпространствами*** пространства L.

Рассмотрим ряд примеров подпространств линейных пространств. Пусть π и l – некоторые фиксированные плоскость и прямая нашего трехмерного пространства. Пусть M_π и M_l – совокупности геометрических векторов, параллельных плоскости π и прямой l соответственно. Совершенно очевидно, что множества M_π и M_l замкнуты относительно сложения и умножения на числа из **R** и тем самым M_π и M_l являются подпространствами пространства геометрических векторов. Далее в п. 3.4.2 мы покажем, что любое собственное подпространство геометрических векторов имеет вид либо M_π, либо M_l для подходящих плоскости π или прямой l.

Пусть M – подмножество пространства $\mathbf{F}_n$ матриц порядка $n \times n$ над полем **F**, состоящее из тех матриц, в которых ниже главной диагонали стоят нули:

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \ldots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \ldots & \alpha_{2n} \\ 0 & 0 & \alpha_{33} & \ldots & \alpha_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \alpha_{nn} \end{pmatrix}.$$

Замкнутость множества M относительно сложения и умножения на числа из **F** очевидна. Тем самым M является подпространством пространства $\mathbf{F}_n$.

Через $\mathbf{C}_0$ обозначим совокупность всех непрерывных функций из множества всех функций $\mathbf{R}^{\mathbf{R}}$. Из хорошо известных теорем анализа (сумма непрерывных функций есть функция непрерывная; произведение непрерывной функции на любое действительное число также является непрерывной функцией) вытекает замкнутость множества $\mathbf{C_0}$ относительно сложения и умножения на числа из $\mathbf{R}$. Тем самым $\mathbf{C_0}$ является подпространством пространства $\mathbf{R}^{\mathbf{R}}$.

3.3.19. Пусть M – подпространство линейного пространства L и $a_1,..., a_m$ – базис пространства M. Система элементов $a_1,..., a_m$, рассматриваемая как система элементов пространства L, является линейно независимой (так как она линейно независима в M).

Если система $a_1,..., a_m$ является максимальной линейно независимой системой в L (т. е. базисом в L), то $L = M$. Действительно, в этом случае для любого $b \in L$ по теореме 3.3.6 найдутся числа $\alpha_1,...,\alpha_m \in \mathbf{F}$ такие, что $b = \alpha_1 a_1 + ... + \alpha_m a_m$, и так как $a_i \in M$ и M является подпространством, то и $\alpha_1 a_1 + ... + \alpha_m a_m = b \in M$, т. е. в этом случае любой элемент из пространства L входит в подпространство M: $L = M$.

Если множество $\{a_1,..., a_m\}$ не является максимальной линейно независимой системой в L, то ее можно дополнить до такой максимальной системы, т. е. до базиса L, за счет добавления некоторых элементов $a_{m+1},..., a_n$. Так как элемент a_{m+1} не принадлежит подпространству M в этом случае, то $L \neq M$ и $m < n$. Резюмируя эти рассуждения и теорему 3.3.10, получаем следующее утверждение.

Теорема. ***Размерность любого подпространства линейного пространства не превышает размерности самого пространства и совпадает с этой размерностью тогда и только тогда, когда подпространство совпадает с самим пространством.***

С другой стороны, если L – n-мерное линейное пространство, то в L существуют подпространства размерности m для любого $0 \leq m \leq n$.

Действительно, пусть $a_1,..., a_n$ – некоторый базис пространства L и $m \leq n$. Пусть M_m – это множество всевозможных линейных комбинаций начальных m элементов этого базиса:

$$M_m = \{\alpha_1 a_1 + ... + \alpha_m a_m \mid \alpha_i \in \mathbf{F}\}.$$

Предоставляем читателю самостоятельно убедиться в том, что M_m является m-мерным подпространством пространства L.

3.3.20. Определение. Для любого множества A элементов линейного пространства L под ***линейной оболочкой*** $L(A)$ множества A понимается совокупность элементов L, представимых в виде линейной комбинации элементов из A, т. е. $L(A) = \{\alpha_1 a_1 + \ldots + \alpha_n a_n \mid \alpha_1, \ldots, \alpha_n \in \mathbf{F},\ a_1, \ldots, a_n \in A$ и n – произвольное натуральное число$\}$

Читателя не затруднит самостоятельно убедиться в том, что для любого множества A его линейная оболочка $L(A)$ является наименьшим подпространством пространства L, содержащим множество A или, как иначе говорят, ***подпространством пространства L, порожденным множеством A***.

3.3.21. Определение. Будем говорить, что линейное подпространство L является ***суммой подпространств*** M_1 и M_2, и писать $L = M_1 + M_2$, если для любого элемента $a \in L$ найдутся $b_1 \in M_1$ и $b_2 \in M_2$ такие, что $a = b_1 + b_2$. Если при этом $M_1 \cap M_2 = \{\underline{0}\}$, то сумма называется ***прямой суммой*** и обозначается $L = M_1 \oplus M_2$.

Заметим, что элементы b_1 и b_2 в случае прямой суммы находятся однозначно. Действительно, если $b_1, b_1' \in M_1$, $b_2, b_2' \in M_2$ и $a = b_1 + b_2 = b_1' + b_2'$, то $b_1 + (-1)b_1' = b_2' + (-1)b_2$. Но так как $b_1 + (-1)b_1' \in M_1$ и $b_2' + (-1)b_2 \in M_2$, а также $M_1 \cap M_2 = \{\underline{0}\}$, то $b_1 + (-1)b_1' = \underline{0}$ и $b_2 + (-1)b_2' = \underline{0}$, т. е. $b_1 = b_1'$ и $b_2 = b_2'$.

Итак, ***если пространство L является прямой суммой подпространств*** M_1 ***и*** M_2***, то любой элемент из L представим в виде суммы элементов подпространств*** M_1 ***и*** M_2***, и это представление однозначно***.

3.3.22. Если $a_1, \ldots, a_m$ – базис пространства M_1 и $b_1, \ldots, b_s$ – базис пространства M_2, а пространство L является прямой суммой подпространств M_1 и M_2, то $a_1, \ldots, a_m$, $b_1, \ldots, b_s$ будет базисом пространства L. Действительно, если для некоторых $\alpha_1, \ldots, \alpha_m$, $\beta_1, \ldots, \beta_s \in \mathbf{F}$

$$\alpha_1 a_1 + \ldots + \alpha_m a_m + \beta_1 b_1 + \ldots + \beta_s b_s = \underline{0},$$

то $$\alpha_1 a_1 + ... + \alpha_m a_m = (-\beta_1) b_1 + ... + (-\beta_s) b_s .$$

Но $\alpha_1 a_1 + ... + \alpha_m a_m \in M_1$ и $(-\beta_1) b_1 + ... + (-\beta_s) b_s \in M_2$, а так как L есть прямая сумма подпространств M_1 и M_2, то $M_1 \cap M_2 = \{\underline{0}\}$, и значит,

$$\alpha_1 a_1 + ... + \alpha_m a_m = \underline{0},$$

$$\beta_1 b_1 + ... + \beta_s b_s = \underline{0}.$$

В силу же того, что $a_1,..., a_m$ образуют базис M_1 и $b_1,..., b_s$ – базис M_2, получаем, что

$$\alpha_1 = 0,..., \ \alpha_m = 0, \quad \beta_1 = 0,..., \ \beta_s = 0,$$

чем и доказана линейная независимость системы $a_1,..., a_m$, $b_1,..., b_s$.

Для доказательства того, что $a_1,..., a_m$, $b_1,..., b_s$ – базис пространства L, остается показать, что любой элемент пространства L представим в виде линейной комбинации системы элементов $a_1,..., a_m$, $b_1,..., b_s$.

Пусть $c \in L$; так как L – прямая сумма M_1 и M_2, то найдутся $d_1 \in M_1$ и $d_2 \in M_2$ такие, что $c = d_2 + d_2$. В силу же того, что $a_1,..., a_m$ – базис M_1 и $b_1,..., b_s$ – базис M_2, найдутся числа $\alpha_1,..., \alpha_m$, $\beta_1,..., \beta_s$, такие, что

$$d_1 = \alpha_1 a_1 + ... + \alpha_m a_m, \quad d_2 = \beta_1 b_1 + ... + \beta_s b_s .$$

Тем самым

$$c = \alpha_1 a_1 + ... + \alpha_m a_m + \beta_1 b_1 + ... + \beta_s b_s,$$

чем и завершается доказательство того, что $a_1,..., a_m$, $b_1,..., b_s$ – базис пространства L – прямой суммы пространств M_1 и M_2.

3.3.23. При решении тех или иных задач в линейных пространствах сложность решения или объем вычислений, как правило, существенно зависит от размерности пространства: уменьшение размерности ведет к упрощению решения и сокращению объема вычислений. При этом довольно часто используется следующий прием: исходя из содержания задачи удается подобрать пару подпространств M_1 и M_2 данного пространства L таким образом, что L есть прямая сумма подпространств M_1 и M_2 и возможно решение той же задачи в этих подпространствах (пространствах меньшей размерности), окончательное же решение исходной задачи в пространстве L при этом представляется как сумма решений задач в пространствах M_1 и M_2.

3.3.24. Полезным является также следующий факт: ***если*** $a_1,..., a_n$ ***– базис пространства*** L***, а число*** m ***таково, что***

$$1 \le m < n \quad \text{и} \quad M_1 = \{\alpha_1 a_1 + ... + \alpha_m a_m \mid \alpha_1,..., \alpha_m \in \mathbf{F}\},$$

а

$$M_2 = \{\alpha_{m+1} a_{m+1} + ... + \alpha_n a_n \mid \alpha_1,..., \alpha_n \in \mathbf{F}\},$$

то M_1 ***и*** M_2 ***– подпространства пространства*** L, ***и*** L ***является прямой суммой подпространств*** M_1 ***и*** M_2. Доказательство этого факта оставляем читателю.

3.3.25. Пусть теперь M_1 – некоторое подпространство пространства L и $a_1,..., a_m$ – его базис. В п. 3.3.19 замечалось, что базис подпространства может быть дополнен до базиса пространства, т. е. можно найти некоторые элементы $a_{m+1},..., a_n \in L$ такие, что $a_1,..., a_m$, $a_{m+1},..., a_n$ образуют базис пространства L. Пусть

$$M_2 = \{\alpha_{m+1} a_{m+1} + ... + \alpha_n a_n \mid \alpha_{m+1},..., \alpha_n \in \mathbf{F}\}.$$

Как замечено в 3.3.24, пространство L будет прямой суммой пространств M_1 и M_2.

3.3.26. Суммируя утверждения 3.3.22 – 3.3.25, сформулируем следующую теорему.

Теорема. 1) ***Если пространство*** L ***есть прямая сумма пространств*** M_1 ***и*** M_2***, то размерность его есть сумма размерностей пространств*** M_1 ***и*** M_2***, при этом объединение базисов пространств*** M_1 ***и*** M_2 ***дает базис пространства*** L.

2) ***Для любого подпространства*** M_1 ***пространства*** L ***найдется (и не одно) подпространство*** M_2 ***пространства*** L ***такое, что*** L ***будет прямой суммой подпространств*** M_1 ***и*** M_2.

3.4. БАЗИСЫ И ПОДПРОСТРАНСТВА В ПРОСТРАНСТВЕ ГЕОМЕТРИЧЕСКИХ ВЕКТОРОВ

Цель настоящего раздела – дать геометрическую интерпретацию введенных в предыдущем разделе понятий для случая пространства геометрических векторов, и все дальнейшие рассмотрения этого раздела будут проводиться в этом пространстве. Первым делом остановимся на вопросах линейной зависимости геометрических векторов.

3.4.1. Теорема. 1) ***Два геометрических вектора*** $\overline{a}$ ***и*** $\overline{b}$ ***линейно зависимы тогда и только тогда, когда они параллельны – или, что то же самое, коллинеарны.***

2) ***Три геометрических вектора*** $\overline{a}$, $\overline{b}$ ***и*** $\overline{c}$ ***линейно зависимы тогда и только тогда, когда они все вместе параллельны некоторой плоскости, т. е. компланарны.***

3) ***Любые четыре геометрических вектора линейно зависимы.***

Доказательство. Прежде всего, заметим, что векторы в утверждениях 1 и 2 можно считать ненулевыми, поскольку нулевой вектор не имеет определенного направления и может считаться параллельным любому другому вектору (в первом случае) или плоскости, определяемой двумя другими векторами (во втором), так что заключение теоремы выполняется. И, конечно, любая система векторов, содержащая нулевой вектор, линейно зависима.

1. Если $\overline{a}$ и $\overline{b}$ – ненулевые параллельные векторы, то $\overline{a} = \pm\frac{|\overline{a}|}{|\overline{b}|}\overline{b}$, где знак «+» или «–» выбирается в зависимости от того, направлены векторы $\overline{a}$ и $\overline{b}$ одинаково или противоположно. Тем самым равенство

$$1\overline{a} - \left(\pm\frac{|\overline{a}|}{|\overline{b}|}\right)\overline{b} = \overline{0}$$

доказывает линейную зависимость параллельных векторов $\overline{a}$ и $\overline{b}$.

Наоборот, если $\bar{a}$ и $\bar{b}$ линейно зависимы, то для некоторых чисел $\alpha, \beta \in \mathbf{R}$, из которых хотя бы одно отлично от нуля – пусть это будет α, имеет место равенство $\alpha\bar{a} + \beta\bar{b} = \bar{0}$ и тем самым равенства $\alpha\bar{a} = -\beta\bar{b}$ и $\bar{a} = \left(-\dfrac{\beta}{\alpha}\right)\bar{b}$, что в силу определения произведения вектора на число доказывает параллельность векторов $\bar{a}$ и $\bar{b}$.

2. Пусть векторы $\bar{a}$, $\bar{b}$ и $\bar{c}$ параллельны некоторой плоскости π. Если $\bar{a}$ и $\bar{b}$ параллельны, то в силу утверждения 1 найдутся числа α, β (хотя бы одно из которых отлично от нуля) такие, что $\alpha\bar{a} + \beta\bar{b} = \bar{0}$, но тогда равенство $\alpha\bar{a} + \beta\bar{b} + 0\bar{c} = \bar{0}$ доказывает линейную зависимость векторов $\bar{a}$, $\bar{b}$, $\bar{c}$. Если же среди векторов $\bar{a}$, $\bar{b}$, $\bar{c}$ нет параллельных, то пусть l_1, l_2, l_3 – любые прямые, проходящие в плоскости π параллельно векторам $\bar{a}$, $\bar{b}$, $\bar{c}$ соответственно и не пересекающиеся в одной точке. Пусть A – точка пересечения прямых l_1, l_3 (соответственно точка B – прямых l_1, l_2, а C – l_2, l_3, рис. 22).

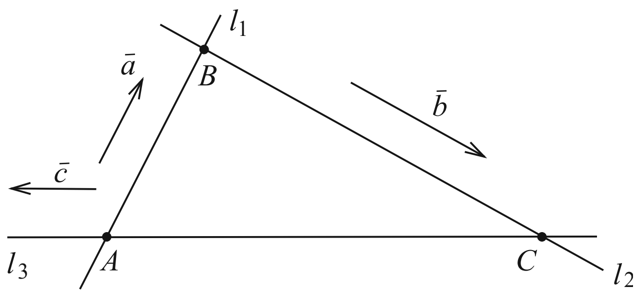


Рис. 22

Теперь в силу параллельности векторов $\bar{a}$ и $\overline{AB}$, $\bar{b}$ и $\overline{BC}$, $\bar{c}$ и $\overline{AC}$ найдутся отличные от нуля числа α, β, γ такие, что $\overline{AB} = \alpha\bar{a}$, $\overline{BC} = \beta\bar{b}$, $\overline{CA} = \gamma\bar{c}$. В силу же равенства $\overline{AB} + \overline{BC} + \overline{CA} = \bar{0}$ получаем искомую линейную зависимость

$$\alpha\bar{a} + \beta\bar{b} + \gamma\bar{c} = \bar{0}.$$

Покажем теперь обратное, пусть

$$\alpha\overline{a} + \beta\overline{b} + \gamma\overline{c} = \overline{0}$$

– некоторая линейная зависимость для системы $\overline{a}$, $\overline{b}$, $\overline{c}$ и, к примеру, $\alpha \neq 0$. Тогда

$$\overline{a} = \left(-\frac{\beta}{\alpha}\right)\overline{b} + \left(-\frac{\gamma}{\alpha}\right)\overline{c}.$$

Пусть π – некоторая плоскость, в которой лежит треугольник векторов $\overline{a}$, $\left(-\frac{\beta}{\alpha}\right)\overline{b}$, $\left(-\frac{\gamma}{\alpha}\right)\overline{c}$. Очевидно, что векторы $\overline{a}$, $\overline{b}$, $\overline{c}$ параллельны плоскости π, и утверждение 2 доказано.

3. Если векторы $\overline{a}$, $\overline{b}$, $\overline{c}$ параллельны какой-то одной плоскости, то в силу утверждения 2 они линейно зависимы и, так же как в начале доказательства утверждения 2, в этом случае получается линейная зависимость системы из четырех векторов $\overline{a}$, $\overline{b}$, $\overline{c}$, $\overline{d}$ для любого вектора $\overline{d}$.

В противном случае выберем три произвольные прямые l_1, l_2, l_3 (не лежащие в одной плоскости!), параллельные векторам $\overline{a}$, $\overline{b}$, $\overline{c}$ соответственно и проходящие через некоторую общую точку O (рис. 23).

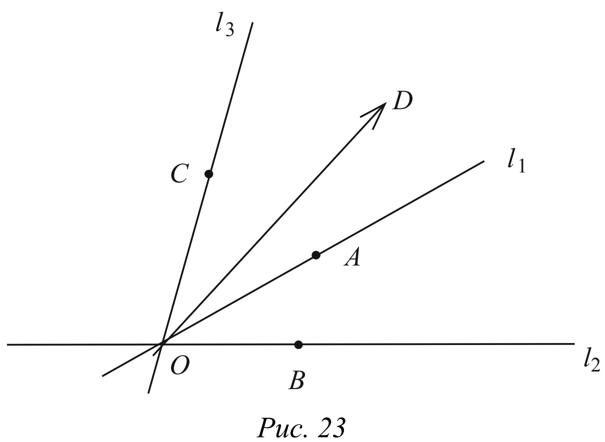


Рис. 23

Пусть D – такая точка, что $\overline{OD}=\overline{d}$. Пусть π_1 – плоскость, содержащая прямые l_2, l_3 ($\pi_2 - l_1$, l_3 и $\pi_3 - l_1$, l_2) соответственно.

Пусть также A – проекция точки D на прямую l_1 параллельно плоскости π_1, B – проекция точки D на прямую l_2 параллельно плоскости π_2 и C – проекция точки D на прямую l_3 параллельно плоскости π_3. Очевидно, что $\overline{d}=\overline{OA}+\overline{OB}+\overline{OC}$.

Из параллельности векторов $\overline{a}$ и $\overline{OA}$, $\overline{b}$ и $\overline{OB}$, $\overline{c}$ и $\overline{OC}$ вытекает, что найдутся числа α, β, γ такие, что $\overline{OA}=\alpha\overline{a}$, $\overline{OB}=\beta\overline{b}$, $\overline{OC}=\gamma\overline{c}$. Тем самым $\overline{d}=\alpha\overline{a}+\beta\overline{b}+\gamma\overline{c}$, т. е. $\alpha\overline{a}+\beta\overline{b}+\gamma\overline{c}+(-1)\overline{d}=\overline{0}$ и линейная зависимость векторов $\overline{a}$, $\overline{b}$, $\overline{c}$, $\overline{d}$ доказана и в этом случае.

3.4.2. В п. 3.3.18 было отмечено, что совокупности геометрических векторов, параллельных произвольной прямой l или плоскости π, образуют подпространства M_l и M_π линейного пространства L геометрических векторов.

Непосредственно из утверждений 1–3 теоремы п. 3.4.1 вытекают следующие утверждения.

3.4.3. Следствие. 1) ***Пространство*** L ***геометрических векторов трехмерно, а три вектора этого пространства образуют его базис тогда и только тогда, когда они все вместе не параллельны никакой плоскости*** (т. е. ***не компланарны***).

2) ***Для любой плоскости*** π ***пространство*** M_π ***двумерно и два вектора этого пространства образуют базис пространства тогда и только тогда, когда они не параллельны*** (т. е. ***не коллинеарны***).

3) ***Для любой прямой*** l ***пространство*** M_l ***одномерно и любой ненулевой вектор этого пространства образует его базис***.

4) ***Любое собственное*** (см. п. 3.3.18) ***подпространство*** M ***пространства*** L ***геометрических векторов имеет вид подпространства*** M_π ***либо*** M_l ***для некоторой плоскости*** π ***либо прямой*** l.

Утверждение 4 данного следствия вытекает из того, что согласно п. 3.3.26 любое собственное подпространство M пространства L геометрических векторов будет либо одномерно, либо двумерно.

Пусть в первом случае базис пространства M состоит из вектора $\overline{a}$. Очевидно, что если l – прямая, параллельная вектору $\overline{a}$, то в этом случае получим $M = M_l$. Если во втором случае $\{\overline{a}, \overline{b}\}$ – это базис пространства M и π – некоторая плоскость, параллельная векторам $\overline{a}$, $\overline{b}$, то пространство M совпадает с M_π.

Без труда замечается, что если пространство L геометрических векторов разложено в прямую сумму двух собственных подпространств, то одно из этих подпространств имеет вид M_π, другое – вид M_l для некоторых плоскости π и прямой l, не параллельной этой плоскости.

3.4.4. Пусть теперь $Oxyz$ – некоторая прямоугольная декартова система координат и $\overline{i}, \overline{j}, \overline{k}$ – векторы единичной длины, параллельные осям Ox, Oy, Oz соответственно (рис. 24).

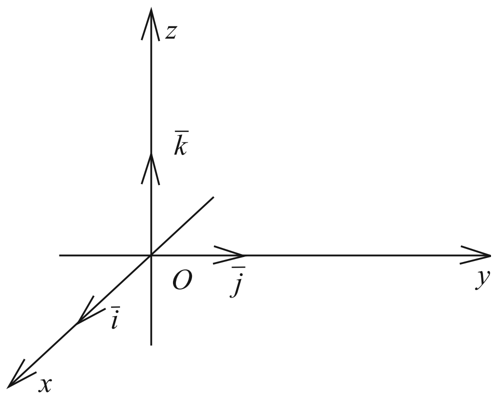


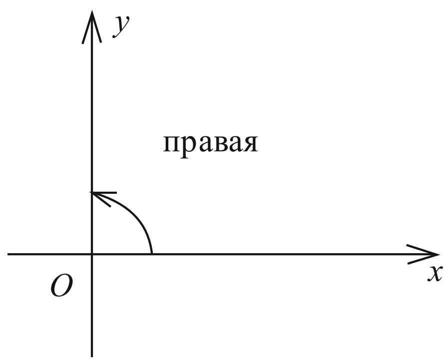
Рис. 24

Согласно утверждению 3 теоремы 3.4.1 векторы $\overline{i}$, $\overline{j}$, $\overline{k}$ образуют базис пространства геометрических векторов, называемый ***стандартным базисом***, связанным с декартовой прямоугольной системой координат $Oxyz$. Отметим также следующий факт, доказательство которого оставляется читателю: если (α, β, γ) – координаты вектора $\overline{a}$ в декартовой прямоугольной системе координат $Oxyz$, то эти же числа $(\alpha; \beta; \gamma)$ будут координатами вектора $\overline{a}$ в стандартном базисе $\overline{i}$, $\overline{j}$, $\overline{k}$, связанном с системой координат $Oxyz$.

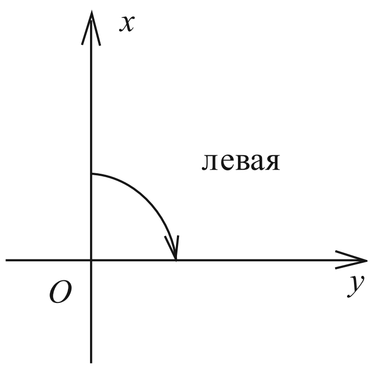
3.4.5. Рассмотрим также геометрический смысл ***матрицы перехода*** (см. п. 3.3.13) от стандартного базиса $\bar{i}$, $\bar{j}$, связанного с одной декартовой системой координат Oxy на плоскости, к стандартному базису $\bar{i}'$, $\bar{j}'$, связанному с другой декартовой системой координат $O'x'y'$ на той же плоскости.

Прежде всего заметим, что любая декартова прямоугольная (да и не только прямоугольная) система координат на плоскости относится к одному из двух классов: ***правая*** либо ***левая*** система координат.

Систему координат Oxy назовем ***правой***, если кратчайший поворот от оси Ox к оси Oy осуществляется против часовой стрелки, в противном случае назовем систему координат Oxy ***левой*** (рис. 25).



а



б

Рис. 25

Пусть теперь Oxy и $O'x'y'$ – две правые (либо левые) декартовы прямоугольные системы координат на плоскости, а $\overline{i}$, $\overline{j}$ и $\overline{i'}$, $\overline{j'}$ – стандартные базисы линейного пространства геометрических векторов, соответствующие системам координат Oxy и $O'x'y'$ (рис. 26).

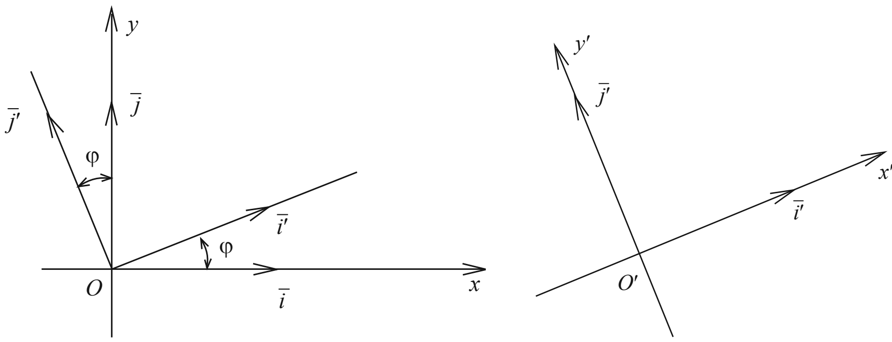


Рис. 26

Отложим векторы $\overline{i'}$ и $\overline{j'}$ из точки O; пусть φ – это угол поворота, которым совмещаются вектор $\overline{i}$ с вектором $\overline{i'}$ и вектор $\overline{j}$ с вектором $\overline{j'}$. Тогда координаты векторов старого базиса $\overline{i}$, $\overline{j}$, будучи проекциями отрезков единичной длины, совпадут с тригонометрическими функциями угла φ:

$$\overline{i} = \cos\varphi \cdot \overline{i'} + (-\sin\varphi)\overline{j'},$$

$$\overline{j} = \sin\varphi \cdot \overline{i'} + \cos\varphi \cdot \overline{j'}.$$

Тем самым матрица перехода от базиса $\overline{i}$, $\overline{j}$ к базису $\overline{i'}$, $\overline{j'}$ будет иметь вид

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix},$$

где φ – угол поворота от вектора $\overline{i}$ к вектору $\overline{i'}$.

В случае, когда системы координат Oxy и $O'x'y'$ будут разных классов, к примеру, Oxy – правая и $O'x'y'$ – левая системы координат, предоставляем читателю убедиться в том, что базис $\overline{i'}$, $\overline{j'}$

возникает из базиса $\bar{i}$, $\bar{j}$ поворотом на угол φ от вектора $\bar{i}$ к вектору $\bar{j}$ и зеркальным отражением относительно оси $O'x'$, а матрица перехода от базиса $\bar{i}$, $\bar{j}$ к базису $\bar{i}'$, $\bar{j}'$ примет вид

$$B = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$

3.4.6. В заключение выведем формулы перехода от координат точки в одной декартовой системе координат к координатам той же точки в другой декартовой системе координат.

Напомним, что координаты точки в фиксированной системе координат совпадают с координатами ее радиуса-вектора (п. 3.1.5).

Пусть точка O' имеет в системе Oxy координаты $O'(x_0, y_0)$.

Обозначим:

$$\bar{x} = (x;\ y),\ \bar{x}' = (x';\ y'),\ \bar{x}_0 = (x_0;\ y_0).$$

Тогда $\overline{OA} = (\bar{i}, \bar{j})\overset{|}{x}$, $\overline{O'A} = (\bar{i}', \bar{j}')\overset{|}{x'}$, $\overline{OO'} = (\bar{i}, \bar{j})\overset{|}{x_0}$ (рис. 27).

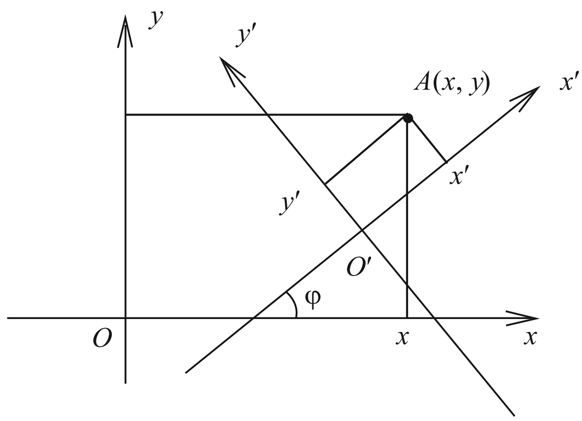


Рис. 27

Поскольку обе системы правые, матрицей перехода служит A, следовательно, $(\bar{i}, \bar{j}) = (\bar{i}', \bar{j}')\ A$ или $(\bar{i}', \bar{j}') = (\bar{i}, \bar{j})\ A^{-1}$. Теперь запишем очевидное векторное равенство $\overline{OA} = \overline{OO'} + \overline{O'A}$ в системе $(\bar{i}, \bar{j})$:

$$(\bar{i}, \bar{j})\overset{|}{x} = (\bar{i}, \bar{j})\overset{|}{x}_0 + (\bar{i}, \bar{j})A^{-1}\overset{|}{x}' = (\bar{i}, \bar{j})(\overset{|}{x}_0 + A^{-1}\overset{|}{x}') .$$

Отсюда по замечанию в конце п. 3.3.12 в силу линейной независимости векторов $\bar{i}$ и $\bar{j}$ следует матричное равенство $\overset{|}{x} = \overset{|}{x}_0 + A^{-1}\overset{|}{x}'$. В качестве легкого упражнения предоставляем читателю найти A^{-1} и выполнить действия в правой части этого равенства. Тогда, приравнивая соответствующие элементы, получаем:

$$x = x'\cos\varphi - y'\sin\varphi + x_0,$$
$$y = x'\sin\varphi + y'\cos\varphi + y_0.$$

Это и есть обещанные формулы.

ГЛАВА 4
ПРОИЗВЕДЕНИЯ ВЕКТОРОВ

4.1. СКАЛЯРНОЕ ПРОИЗВЕДЕНИЕ

Помимо операций сложения геометрических векторов и умножения их на числа, важную роль при решении различных геометрических и физических задач играют произведения векторов между собой. Мы рассмотрим три различные операции перемножения векторов: скалярное, векторное и смешанное произведения. Каждое из них может пригодиться при решении весьма разнообразных задач. И все же можно отметить некоторую «специализацию» этих операций: скалярное произведение играет наибольшую роль в задачах, связанных с нахождением углов, векторное произведение – в задачах о площадях фигур, смешанное – в задачах об объемах тел.

Везде ниже угол между векторами $\overline{a}$ и $\overline{b}$ обозначается $\angle\,\overline{a}\overline{b}$, так что если

$$\overline{a} = \overline{AB},\ \overline{b} = \overline{AC},\ \text{то}\ \angle\,\overline{a}\overline{b} = \angle\,BAC.$$

4.1.1. Напомним известное из школьного курса математики определение ***скалярного произведения*** векторов: любой паре векторов $\overline{a}$, $\overline{b}$ ставится в соответствие действительное число, обозначаемое далее как $(\overline{a},\ \overline{b})$ и именуемое скалярным произведением этих векторов по следующему правилу:

$$(\overline{a},\ \overline{b}) = |\overline{a}||\overline{b}|\cos\angle\overline{a}\,\overline{b}.$$

4.1.2. Отметим ряд простейших свойств операции скалярного произведения:

1) $(\overline{a},\ \overline{b}) = (\overline{b},\ \overline{a})$ (коммутативность скалярного произведения);

2) $(\overline{a}+\overline{b},\ \overline{c}) = (\overline{a},\ \overline{c}) + (\overline{b},\ \overline{c})$ (дистрибутивность скалярного произведения относительно сложения);

3) $(\alpha\overline{a},\ \overline{b}) = \alpha(\overline{a},\ \overline{b})$ для любого числа $\alpha \in \mathbf{R}$ (однородность относительно умножения на числа);

4) $(\overline{a},\ \overline{a}) > 0$, как только $\overline{a} \neq \overline{0}$ (положительная определенность) и $(\overline{a},\ \overline{a}) = 0 \Leftrightarrow \overline{a} = \overline{0}$.

Справедливость свойств 1 и 4 очевидна из определения скалярного произведения. Доказательство свойства 3 сводится к рассмотрению трех случаев:

если $\alpha > 0$, то $\angle\,\overline{a}\,\overline{b} = \angle\,(\alpha\overline{a})\,\overline{b}$ и $|\alpha\overline{a}| = \alpha\cdot|\overline{a}|$. Тем самым в этом случае

$$(\alpha\overline{a},\ \overline{b}) = |\alpha\overline{a}|\,|\overline{b}|\cos\angle(\alpha\overline{a})\overline{b} = \alpha|\overline{a}|\,|\overline{b}|\cos\angle\overline{a}\,\overline{b} = \alpha(\overline{a},\ \overline{b});$$

в случае, когда $\alpha < 0$, углы $\angle\,\overline{a}\,\overline{b}$ и $\angle\,(\alpha\overline{a})\,\overline{b}$ смежные; их косинусы различаются знаком и равны по модулю, и, конечно, $|\alpha\overline{a}| = -\alpha|\overline{a}|$.

Таким образом, в этом случае

$$(\alpha\overline{a},\ \overline{b}) = |\alpha\overline{a}||\overline{b}|\cos\angle(\alpha\,\overline{a})\,\overline{b} = -\alpha|\overline{a}||\overline{b}|\cos(\pi - \angle\overline{a}\,\overline{b}) =$$

$$= -\alpha|\overline{a}||\overline{b}|(-\cos\angle\overline{a}\,\overline{b}) = \alpha|\overline{a}||\overline{b}|\cos\angle\overline{a}\,\overline{b} = \alpha(\overline{a},\ \overline{b});$$

если же $\alpha = 0$, свойство 3 очевидно; тем самым оно доказано нами полностью.

Докажем свойство 2. Пусть $l_{\overline{b}}$ – произвольная прямая, параллельная вектору $\overline{b}$, тогда проекция вектора $\overline{a}$ на прямую $l_{\overline{b}}$ выражается так: $\text{пр}_{l_{\overline{b}}}\overline{a} = |\overline{a}|\cos\angle\overline{a}\,\overline{b}$ и имеет место следующая формула:

$$(\overline{a},\ \overline{b}) = |\overline{b}|\,\text{пр}_{l_{\overline{b}}}\overline{a}\,.$$

В силу указанной в разделе 3.1.4 формулы

$$\text{пр}_{l_{\overline{c}}}(\overline{a}+\overline{b}) = \text{пр}_{l_{\overline{c}}}\overline{a} + \text{пр}_{l_{\overline{c}}}\overline{b}$$

имеем следующую цепочку равенств:

$$(\overline{a},\ \overline{c})+(\overline{b},\ \overline{c})=|\overline{a}|\ |\overline{c}|\cos\angle\overline{a}\ \overline{c}+|\overline{b}|\ |\overline{c}|\cos\angle\overline{b}\ \overline{c}=$$

$$=|\overline{c}|\left(|\overline{a}|\cos\angle\overline{a}\ \overline{c}+|\overline{b}|\cos\angle\overline{b}\ \overline{c}\right)=$$

$$=|\overline{c}|\left(\text{пр}_{l_{\overline{c}}}\overline{a}+\text{пр}_{l_{\overline{c}}}\overline{b}\right)=|\overline{c}|\ \ \text{пр}_{l_{\overline{c}}}(\overline{a}+\overline{b})=(\overline{a}+\overline{b},\ \overline{c}).$$

Тем самым доказано и свойство 3.

4.1.3. Из свойств 1–3 вытекают также следующие равенства:

$$(\overline{a},\ \overline{b}+\overline{c})=(\overline{a},\ \overline{b})+(\overline{a},\ \overline{c}),$$

$$(\overline{a},\ \alpha\overline{b})=\alpha(\overline{a},\ \overline{b}).$$

Действительно,

$$(\overline{a},\ \overline{b}+\overline{c})=(\overline{b}+\overline{c},\ \overline{a})=(\overline{b},\ \overline{a})+(\overline{c},\ \overline{a})=(\overline{a},\ \overline{b})+(\overline{a},\ \overline{c}),$$

$$(\overline{a},\ \alpha\overline{b})=(\alpha\overline{b},\ \overline{a})=\alpha(\overline{b},\ \overline{a})=\alpha(\overline{a},\ \overline{b}).$$

4.1.4. Из доказанных свойств вытекает формула, позволяющая вычислить скалярное произведение векторов, заданных своими координатами в произвольной декартовой прямоугольной системе координат. Заметим, что для стандартного базиса $\overline{i}$, $\overline{j}$, $\overline{k}$ (п. 3.4.4), соответствующего декартовой прямоугольной системе координат *Oxyz*, непосредственно из определения скалярного произведения вытекает такая «таблица умножения»:

$$(\overline{i},\ \overline{i})=(\overline{j},\overline{j})=(\overline{k},\ \overline{k})=1,$$

$$(\overline{i},\overline{j})=(\overline{i},\ \overline{k})=(\overline{j},\ \overline{i})=(\overline{j},\ \overline{k})=(\overline{k},\ \overline{i})=(\overline{k},\overline{j})=0.$$

Пусть $(x_1;\ y_1;\ z_1)$ и $(x_2;\ y_2;\ z_2)$ – координаты векторов $\overline{a}$ и $\overline{b}$ в системе координат *Oxyz*, т. е.

$$\overline{a}=x_1\overline{i}+y_1\overline{j}+z_1\overline{k}\quad \text{и}\quad \overline{b}=x_2\overline{i}+y_2\overline{j}+z_2\overline{k}.$$

Тогда с помощью свойств 2, 3 скалярного произведения и «таблицы умножения» получаем:

$$(\overline{a},\ \overline{b})=\left(x_1\overline{i}+y_1\overline{j}+z_1\overline{k},\ x_2\overline{i}+y_2\overline{j}+z_2\overline{k}\right)=\left(x_1\overline{i},\ x_2\overline{i}+y_2\overline{j}+z_2\overline{k}\right)+$$

$$+\left(y_1\overline{j},\ x_2\overline{i}+y_2\overline{j}+z_2\overline{k}\right)+\left(z_1\overline{k},\ x_2\overline{i}+y_2\overline{j}+z_2\overline{k}\right)=$$

$$=\left(x_1\overline{i},\ x_2\overline{i}\right)+\left(x_1\overline{i},\ y_2\overline{j}\right)+\left(x_1\overline{i},\ z_2\overline{k}\right)+\left(y_1\overline{j},\ x_2\overline{i}\right)+$$

$$+\left(y_1\overline{j},\ y_2\overline{j}\right)+\left(y_1\overline{j},\ z_2\overline{k}\right)+\left(z_1\overline{k},\ x_2\overline{i}\right)+\left(z_1\overline{k},\ y_2\overline{j}\right)+\left(z_1\overline{k},\ z_2\overline{k}\right)=$$

$$=x_1x_2(\overline{i},\overline{i})+x_1y_2(\overline{i},\overline{j})+x_1z_2(\overline{i},\ \overline{k})+y_1x_2(\overline{j},\ \overline{i})+y_1y_2(\overline{j},\overline{j})+$$

$$+\ y_1z_2(\overline{j},\ \overline{k})+z_1x_2(\overline{k},\ \overline{i})+z_1y_2(\overline{k},\overline{j})+z_1z_2(\overline{k},\ \overline{k})=x_1x_2+y_1y_2+z_1z_2.$$

4.1.5. Итак, если $(x_1;\ y_1;\ z_1)$, $(x_2;\ y_2;\ z_2)$ – координаты векторов $\overline{a}$ и $\overline{b}$ в декартовой прямоугольной системе координат, то

$$(\overline{a},\ \overline{b})=x_1x_2+y_1y_2+z_1z_2.$$

4.1.6. Из определения скалярного произведения вытекают равенства:

$$|\overline{a}|=\sqrt{(\overline{a},\ \overline{a})},$$

$$\cos\angle\overline{a}\,\overline{b}=\frac{(\overline{a},\ \overline{b})}{|\overline{a}||\overline{b}|}.$$

Если $(x_1;\ y_1;\ z_1)$ и $(x_2;\ y_2;\ z_2)$ – координаты векторов $\overline{a}$ и $\overline{b}$ в декартовой системе координат, то эти формулы приводят к следующим равенствам:

$$|\overline{a}|=\sqrt{x_1^2+y_1^2+z_1^2},$$

$$\cos\angle\overline{a}\,\overline{b}=\frac{x_1x_2+y_1y_2+z_1z_2}{\sqrt{x_1^2+y_1^2+z_1^2}\sqrt{x_2^2+y_2^2+z_2^2}}.$$

4.1.7. В заключение этого раздела отметим также следующее свойство, непосредственно вытекающее из определения скалярного произведения:

$(\overline{a},\ \overline{b})=0$ тогда и только тогда, когда $\overline{a}$ и $\overline{b}$ перпендикулярны (включая тот случай, когда один из векторов равен нулю);

$(\overline{a},\ \overline{b}) > 0$ тогда и только тогда, когда угол $\angle\,\overline{a}\,\overline{b}$ острый;

$(\overline{a},\ \overline{b}) < 0$ тогда и только тогда, когда угол $\angle\,\overline{a}\,\overline{b}$ тупой.

4.1.8. Одно из физических приложений скалярного произведения известно из школьного курса физики. А именно, работа силы $\overline{F}$, действующей на материальную точку при перемещении ее вдоль вектора $\overline{s}$, равна скалярному произведению этих векторов:

$$A = (\overline{F},\ \overline{s})\,.$$

4.2. ВЕКТОРНОЕ ПРОИЗВЕДЕНИЕ И ЕГО ПРИЛОЖЕНИЯ

4.2.1. Прежде всего нам понадобятся понятия ***правых*** и ***левых*** упорядоченных троек векторов. Существуют различные определения этих понятий. Мы приведем здесь в первую очередь наиболее простое и очевидное, «анатомическое». Упорядоченную тройку некомпланарных векторов $\overline{a},\ \overline{b},\ \overline{c}$ будем называть ***правой*** (***левой***), если векторы $\overline{a},\ \overline{b},\ \overline{c}$ расположены так же, как расставленные большой, указательный и средний пальцы правой (левой) руки; например, если кисть расположена ладонью вверх (рис. 28).

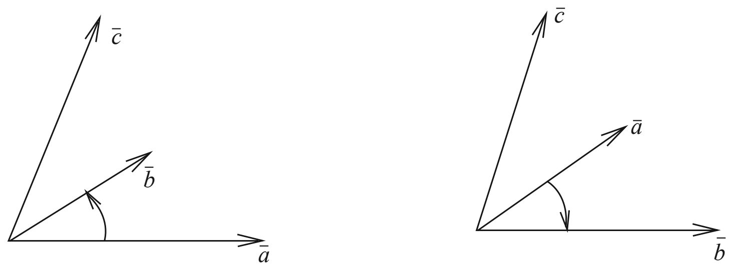


Рис. 28

Правую и левую тройки векторов можно описать иначе: ***упорядоченная тройка*** $\overline{a},\ \overline{b},\ \overline{c}$ ***является правой*** (***левой***), **если *кратчайший поворот от вектора*** $\overline{a}$ ***к вектору*** $\overline{b}$ ***осуществляется против часовой стрелки*** (***по часовой стрелке***), **если наблюдать за этим поворотом из конца вектора** $\overline{c}$ **.**

4.2.2. Определение. Любой паре векторов $\overline{a}$, $\overline{b}$ поставим в соответствие третий вектор $\overline{c}$, называемый их ***векторным произведением*** и обозначаемый как $\overline{a}\times\overline{b}$ по следующим правилам:

1) $\overline{c}$ перпендикулярен векторам $\overline{a}$ и $\overline{b}$;

2) $|\overline{c}| = |\overline{a}|\,|\overline{b}|\sin\angle\overline{a}\,\overline{b}$;

3) упорядоченная тройка векторов $\overline{a}$, $\overline{b}$, $\overline{c}$ является правой.

4.2.3. Из свойства 2 этого определения непосредственно вытекает, что длина векторного произведения $\overline{c}=\overline{a}\times\overline{b}$ равна площади параллелограмма, построенного на векторах $\overline{a}$ и $\overline{b}$ (рис. 29).

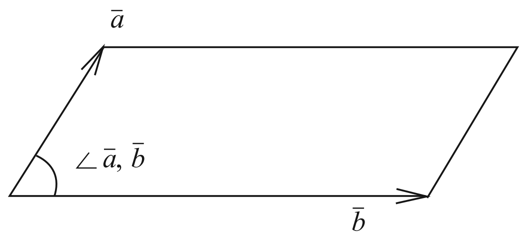


Рис. 29

4.2.4. Заметим также, что из определения векторного произведения (точнее, из определения его длины) прямо следует, что для любых векторов $\overline{a}$ и $\overline{b}$ равенство $\overline{a}\times\overline{b}=\overline{0}$ верно тогда и только тогда, когда $\overline{a}$ и $\overline{b}$ параллельны (включая случай, когда один из векторов равен нулю).

4.2.5. Докажем теперь основные алгебраические свойства векторного произведения:

1) $\overline{b}\times\overline{a}=-\overline{a}\times\overline{b}$;

2) $(\lambda\overline{a})\times\overline{b}=\lambda(\overline{a}\times\overline{b})$;

3) $(\overline{a}+\overline{b})\times\overline{c}=(\overline{a}\times\overline{c})+(\overline{b}\times\overline{c})$.

В силу того, что векторное произведение параллельных векторов равно $\overline{0}$ (п. 4.2.4), при доказательстве свойств 1 и 2 достаточно рассматривать случай, когда $\overline{a}$ и $\overline{b}$ не параллельны.

Итак, пусть $\overline{a}$ и $\overline{b}$ не параллельны. Тогда параллельными будут векторы $\overline{a}\times\overline{b}$ и $\overline{b}\times\overline{a}$ как перпендикуляры к паре непараллельных векторов $\overline{a}$ и $\overline{b}$. Длины векторов $\overline{a}\times\overline{b}$ и $\overline{b}\times\overline{a}$ совпадают, так как

$$\left|\overline{a}\times\overline{b}\right| = \left|\overline{a}\right|\left|\overline{b}\right|\sin\angle\overline{a}\,\overline{b} = \left|\overline{b}\right|\left|\overline{a}\right|\sin\angle\overline{b}\,\overline{a} = \left|\overline{b}\times\overline{a}\right|.$$

Наконец, тройка $\overline{b}$, $\overline{a}$, $\overline{b}\times\overline{a}$ правая и, значит, тройка $\overline{a}$, $\overline{b}$, $\overline{b}\times\overline{a}$ левая. Следовательно, векторы $\overline{a}\times\overline{b}$ и $\overline{b}\times\overline{a}$ параллельны, имеют равную длину и противоположно направлены, т. е. действительно

$$\overline{b}\times\overline{a} = -\overline{a}\times\overline{b}.$$

При доказательстве свойства 2 по-прежнему считаем, что $\overline{a}$ и $\overline{b}$ не параллельны, а также, что $\lambda \neq 0$. Рассмотрим два случая:

а) $\lambda > 0$. Векторы $\overline{a}$ и $\lambda\overline{a}$ параллельны и одинаково направлены. Поэтому упорядоченные тройки векторов: $\overline{a}$, $\overline{b}$, $\overline{a}\times\overline{b}$, затем $\overline{a}$, $\overline{b}$, $\lambda(\overline{a}\times\overline{b})$ и, наконец, $\lambda\overline{a}$, $\overline{b}$, $(\lambda\overline{a})\times\overline{b}$ являются правыми. Отсюда и из определения векторного произведения следует, что векторы $\overline{a}\times\overline{b}$, $\lambda(\overline{a}\times\overline{b})$, $(\lambda\overline{a})\times\overline{b}$ перпендикулярны векторам $\overline{a}$, $\overline{b}$, а значит, параллельны между собой и одинаково направлены. Кроме того,

$$\left|\lambda(\overline{a}\times\overline{b})\right| = \lambda\left|\overline{a}\times\overline{b}\right| = \lambda\left|\overline{a}\right|\left|\overline{b}\right|\sin\angle\overline{a}\,\overline{b} =$$

$$= \left|\lambda\overline{a}\right|\left|\overline{b}\right|\sin\angle(\lambda\overline{a})\,\overline{b} = \left|(\lambda\overline{a})\times\overline{b}\right|.$$

Из доказанного и следует равенство 2 в случае, когда $\lambda > 0$.

б) $\lambda < 0$. Рассматривается аналогично, и здесь доказательство свойства 2 предоставляется читателю.

4.2.6. Для доказательства свойства 3 нам предварительно придется ввести понятие проекции вектора $\overline{a}$ на плоскость π. Пусть A и B – некоторые точки пространства, такие, что $\overline{AB} = \overline{a}$, и пусть A_0, B_0 – проекции точек A и B на плоскость π. Вектор $\overline{A_0B_0}$ назовем проекцией вектора $\overline{a}$ на плоскость π и будем впредь обозначать как $\text{пр}_\pi\overline{a}$ (рис. 30).

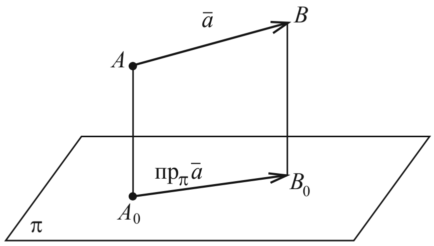


Рис. 30

Поскольку равенство $\overline{AB}+\overline{BC}=\overline{AC}$ очевидно переходит в равенство $\overline{A_0B_0}+\overline{B_0C_0}=\overline{A_0C_0}$, для любых векторов $\overline{a}$, $\overline{b}$ и любой плоскости π имеет место равенство

$$\text{пр}_\pi(\overline{a}+\overline{b})=\text{пр}_\pi\overline{a}+\text{пр}_\pi\overline{b},$$

аналогичное равенству п. 3.1.4.

Пусть теперь даны векторы $\overline{a}$ и $\overline{b}$. Обозначим через $\pi_{\overline{b}}$ плоскость, перпендикулярную вектору $\overline{b}$. Тогда

$$\overline{a}\times\overline{b}=\text{пр}_{\pi_{\overline{b}}}\overline{a}\times\overline{b}.$$

Действительно, как непосредственно видно, векторы $\overline{a}$, $\overline{b}$ и $\text{пр}_{\pi_{\overline{b}}}\overline{a}$ компланарны, так что общий перпендикуляр к паре $\overline{a}$, $\overline{b}$ тот же, что и к паре $\overline{b}$, $\text{пр}_{\pi_{\overline{b}}}\overline{a}$; более того, векторы $\overline{a}\times\overline{b}$ и $\text{пр}_{\pi_{\overline{b}}}\overline{a}\times\overline{b}$ имеют одинаковое направление. А длины этих векторов равны, так как параллелограммы, построенные на векторах $\overline{a}$, $\overline{b}$ и на векторах $\text{пр}_{\pi_{\overline{b}}}\overline{a}$, $\overline{b}$, равновелики (основание $\overline{b}$ обоих параллелограммов одно и то же, а сторона второго $\text{пр}_{\pi_{\overline{b}}}\overline{a}$, перпендикулярная стороне $\overline{b}$, является высотой первого, рис. 31).

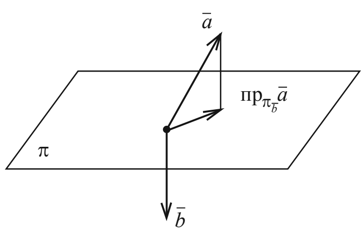


Рис. 31

4.2.7. Вернемся теперь к доказательству свойства 3. Если $\overline{c} = \overline{0}$, то равенство 3 очевидно. Заметим также, что равенство 3 достаточно доказать для случая $|\overline{c}| = 1$, так как для вектора $\overline{c}$ произвольной длины оно будет следовать из этого случая и свойства 2.

Итак, пусть $|\overline{c}| = 1$ и $\overline{a} \perp \overline{c}$; $\pi_{\overline{c}}$ – плоскость, перпендикулярная вектору $\overline{c}$. Выясним, как вектор $\overline{a} \times \overline{c}$ получается из вектора $\overline{a}$. По определению, три вектора $\overline{a}$, $\overline{c}$ и $\overline{a} \times \overline{c}$ попарно перпендикулярны и образуют правую тройку, при этом $|\overline{a} \times \overline{c}| = |\overline{a}| \cdot |\overline{c}| \cdot \sin \angle \overline{ac} = |\overline{a}|$. Таким образом, вектор $\overline{a} \times \overline{c}$ получается из вектора $\overline{a}$ поворотом на прямой угол вокруг вектора $\overline{c}$, причем по часовой стрелке (если смотреть с конца вектора $\overline{c}$ – это легко понять, представив векторы $\overline{a}$, $\overline{c}$ и $\overline{a} \times \overline{c}$ лежащими на декартовых осях *Ox*, *Oy* и *Oz* соответственно). Следовательно, векторы $\text{пр}_{\pi_{\overline{c}}}\overline{a} \times \overline{c}$, $\text{пр}_{\pi_{\overline{c}}}\overline{b} \times \overline{c}$ и $\left(\text{пр}_{\pi_{\overline{c}}}\overline{a} + \text{пр}_{\pi_{\overline{c}}}\overline{b}\right) \times \overline{c}$ получаются из векторов $\text{пр}_{\pi_{\overline{c}}}\overline{a}$, $\text{пр}_{\pi_{\overline{c}}}\overline{b}$ и $\text{пр}_{\pi_{\overline{c}}}\overline{a}$ + $\text{пр}_{\pi_{\overline{c}}}\overline{b}$ поворотом в плоскости $\pi_{\overline{c}}$ вокруг вектора $\overline{c}$ на угол $\dfrac{\pi}{2}$ по часовой стрелке, если смотреть с конца вектора $\overline{c}$ (тройка взаимно перпендикулярных векторов $\text{пр}_{\pi_{\overline{c}}}\overline{a}$, $\overline{c}$, $\left(\text{пр}_{\pi_{\overline{c}}}\overline{a}\right) \times \overline{c}$ – правая по определению п. 4.2.1, значит, тройка $\text{пр}_{\pi_{\overline{c}}}\overline{a}$, $\left(\text{пр}_{\pi_{\overline{c}}}\overline{a}\right) \times \overline{c}$, $\overline{c}$ – левая, рис. 32).

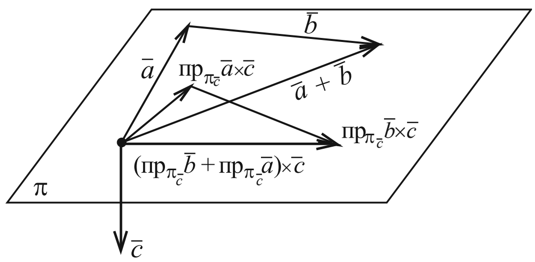


Рис. 32

Поэтому

$$\left(\text{пр}_{\pi_{\overline{c}}}\overline{a} + \text{пр}_{\pi_{\overline{c}}}\overline{b}\right) \times \overline{c} = \left(\text{пр}_{\pi_{\overline{c}}}\overline{a} \times \overline{c}\right) + \left(\text{пр}_{\pi_{\overline{c}}}\overline{b} \times \overline{c}\right).$$

Так как

$$\text{пр}_{\pi_c}\overline{a}\times\overline{c} = \overline{a}\times\overline{c},\ \text{пр}_{\pi_{\overline{b}}}\overline{b}\times\overline{c} = \overline{b}\times\overline{c},\ \left(\text{пр}_{\pi_{\overline{c}}}\overline{a}+\text{пр}_{\pi_{\overline{c}}}\overline{b}\right)\times\overline{c} =$$

$$= \left(\text{пр}_{\pi_{\overline{c}}}(\overline{a}+\overline{b})\right)\times\overline{c} = (\overline{a}+\overline{b})\times\overline{c},$$

то $(\overline{a}+\overline{b})\times\overline{c} = (\overline{a}\times\overline{c})+(\overline{b}\times\overline{c})$, что и завершает доказательство равенства 3.

4.2.8. Заметим, что из свойств 1–3 непосредственно вытекают равенства

4) $\overline{a}\times(\lambda\overline{b}) = \lambda(\overline{a}\times\overline{b})$,

5) $\overline{a}\times(\overline{b}+\overline{c}) = (\overline{a}\times\overline{b})+(\overline{a}\times\overline{c})$.

Действительно,

$$\overline{a}\times(\lambda\overline{b}) = -(\lambda\overline{b})\times\overline{a} = -\lambda(\overline{b}\times\overline{a}) = \lambda(-\overline{b}\times\overline{a}) = \lambda(\overline{a}\times\overline{b}),$$

$$\overline{a}\times(\overline{b}+\overline{c}) = -(\overline{b}+\overline{c})\times\overline{a} = -((\overline{b}\times\overline{a})+(\overline{c}\times\overline{a})) =$$

$$= -(-(\overline{a}\times\overline{b})-(\overline{a}\times\overline{c})) = (\overline{a}\times\overline{b})+(\overline{a}\times\overline{c}).$$

4.2.9. Рассмотрим теперь вопрос о выражении координат векторного произведения через известные координаты сомножителей в правой декартовой системе координат.

Прежде всего отметим равенства, непосредственно следующие из определения векторного произведения (здесь $\overline{i}$, $\overline{j}$, $\overline{k}$, как всегда, стандартный базис, соответствующий некоторой правой декартовой прямоугольной системе координат $Oxyz$):

$$\overline{i}\times\overline{i} = \overline{0},\quad \overline{j}\times\overline{i} = -\overline{k},\quad \overline{k}\times\overline{i} = \overline{j},$$

$$\overline{i}\times\overline{j} = \overline{k},\quad \overline{j}\times\overline{j} = \overline{0},\quad \overline{k}\times\overline{j} = -\overline{i},$$

$$\overline{i}\times\overline{k} = -\overline{j},\quad \overline{j}\times\overline{k} = \overline{i},\quad \overline{k}\times\overline{k} = \overline{0}.$$

Пусть теперь $(x_1;\ y_1;\ z_1)$ и $(x_2;\ y_2;\ z_2)$ – координаты векторов $\overline{a}$ и $\overline{b}$ в системе координат $Oxyz$, т. е. $\overline{a} = x_1\overline{i}+y_1\overline{j}+z_1\overline{k}$,

$\overline{b} = x_2\overline{i} + y_2\overline{j} + z_2\overline{k}$. Тогда на основе свойств 1–5 векторного произведения получаем цепочку равенств

$$\overline{a} \times \overline{b} = \left(x_1\overline{i} + y_1\overline{j} + z_1\overline{k}\right) \times \left(x_2\overline{i} + y_2\overline{j} + z_2\overline{k}\right) =$$

$$= x_1\overline{i} \times \left(x_2\overline{i} + y_2\overline{j} + z_2\overline{k}\right) + y_1\overline{j} \times \left(x_2\overline{i} + y_2\overline{j} + z_2\overline{k}\right) +$$

$$+ z_1\overline{k} \times \left(x_2\overline{i} + y_2\overline{j} + z_2\overline{k}\right) = x_1x_2(\overline{i} \times \overline{i}) +$$

$$+ x_1y_2(\overline{i} \times \overline{j}) + x_1z_2(\overline{i} \times \overline{k}) + y_1x_2(\overline{j} \times \overline{i}) + y_1y_2(\overline{j} \times \overline{j}) +$$

$$+ y_1z_2(\overline{j} \times \overline{k}) + z_1x_2(\overline{k} \times \overline{i}) + z_1y_2(\overline{k} \times \overline{j}) + z_1z_2(\overline{k} \times \overline{k}) =$$

$$= (x_1y_2)\overline{k} - (x_1z_2)\overline{j} - (y_1x_2)\overline{k} + (y_1z_2)\overline{i} + (z_1x_2)\overline{j} - (z_1y_2)\overline{i} =$$

$$= (y_1z_2 - z_1y_2)\overline{i} + (z_1x_2 - x_1z_2)\overline{j} + (x_1y_2 - y_1x_2)\overline{k} =$$

$$= \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} \overline{i} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix} \overline{j} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \overline{k}.$$

Или, по аналогии с формулой разложения определителя по первой строке,

$$\overline{a} \times \overline{b} = \begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}.$$

4.2.10. В качестве одного из физических приложений понятия векторного произведения отметим связь момента $\overline{M}$ силы $\overline{F}$ и плеча $\overline{r}$ ее приложения относительно точки (или оси), вокруг которой происходит вращение тела:

$$\overline{M} = \overline{F} \times \overline{r}.$$

Так же связаны линейная $\overline{v}$ и угловая $\overline{w}$ скорости:

$$\overline{v} = \overline{w} \times \overline{r},$$

где $\overline{r}$ – радиус-вектор, направленный с оси вращения в точку, где вычисляется скорость. Схожим образом магнитная индукция $\overline{B}$ выражается через элемент тока $\overline{I}$, протекающего в прямолинейном проводнике небольшой длины l, и вектор $\overline{r}$ от его центра до точки, где вычисляется индукция:

$$\overline{B} = \frac{\mu l}{r^3} \overline{I} \times \overline{r},$$

здесь r – длина вектора $\overline{r}$ (она гораздо больше, чем l), а μ – магнитная постоянная, зависящая от свойств среды.

4.3. СМЕШАННОЕ ПРОИЗВЕДЕНИЕ И ЕГО СВОЙСТВА

4.3.1. Определение. ***Смешанным произведением*** упорядоченной тройки векторов $\overline{a}$, $\overline{b}$, $\overline{c}$ называется число, равное скалярному произведению вектора $\overline{a} \times \overline{b}$ на вектор $\overline{c}$.

Обозначим смешанное произведение векторов $\overline{a}, \overline{b}, \overline{c}$ как $(\overline{a}, \overline{b}, \overline{c})$, и тогда по определению

$$(\overline{a}, \overline{b}, \overline{c}) = (\overline{a} \times \overline{b}, \ \overline{c}).$$

4.3.2. Прежде всего остановимся на геометрической интерпретации смешанного произведения. Пусть тройка векторов $\overline{a}, \overline{b}, \overline{c}$ является правой. Через V обозначим объем параллелепипеда, построенного на векторах $\overline{a}, \overline{b}, \overline{c}$ (при совмещении начал всех векторов $\overline{a}, \overline{b}, \overline{c}$). Тогда $V = Sh$, где площадь основания S – это площадь параллелограмма, построенного на векторах $\overline{a}$ и $\overline{b}$, а высота параллелепипеда h – это длина отрезка CD, перпендикулярного плоскости π основания этого параллелограмма (рис. 33).

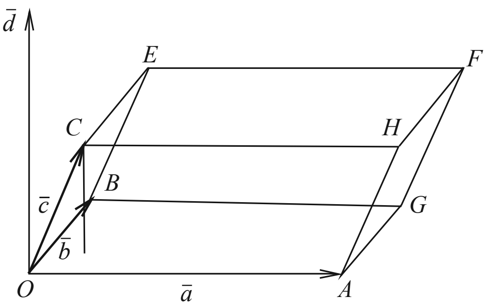


Рис. 33

Пусть $\overline{d} = \overline{a} \times \overline{b}$, тогда по определению п. 4.2.1 $|\overline{d}| = S$, а вектор $\overline{d}$ перпендикулярен плоскости π, проходящей через векторы $\overline{a}$, $\overline{b}$, и, наконец, тройка векторов $\overline{a}$, $\overline{b}$, $\overline{d}$ правая. Отсюда следует, что векторы $\overline{c}$ и $\overline{d}$ будут расположены по одну сторону от плоскости π. Так как при этом вектор $\overline{d}$ перпендикулярен плоскости π, то угол $\angle\overline{c}\overline{d}$ острый, т. е. $\cos\angle\overline{c}\overline{d} \geq 0$. Тем самым $h = |\overline{c}|\cos\angle\overline{c}\overline{d}$ и

$$(\overline{a} \times \overline{b},\ \overline{c}) = (\overline{d},\ \overline{c}) = |\,\overline{d}\,||\,\overline{c}\,|\cos\angle\overline{d}\overline{c} = Sh = V.$$

Итак, в случае, когда $\overline{a}$, $\overline{b}$, $\overline{c}$ – правая тройка векторов, имеет место равенство $V = (\overline{a},\ \overline{b},\ \overline{c})$.

В случае, когда $\overline{a}$, $\overline{b}$, $\overline{c}$ – левая тройка, аналогичные рассуждения приводят к равенству $V = -(\overline{a},\ \overline{b},\ \overline{c})$.

Обобщая полученные формулы на оба случая, получаем формулу

$$V = |(\overline{a},\ \overline{b},\ \overline{c})|,$$

т. е. модуль смешанного произведения векторов $\overline{a}, \overline{b}, \overline{c}$ ***равен объему параллелепипеда, построенного на векторах*** $\overline{a}$, $\overline{b}$, $\overline{c}$.

4.3.3. Из утверждения и доказательства 4.3.2 вытекает ряд следствий:

1) ***векторы*** $\overline{a}, \overline{b}, \overline{c}$ ***компланарны*** (т. е. параллельны некоторой одной плоскости) **тогда и только тогда, когда** $(\overline{a},\ \overline{b},\ \overline{c}) = 0$;

2) ***тройка векторов*** $\overline{a}, \overline{b}, \overline{c}$ ***правая*** (***левая***) ***тогда и только тогда, когда*** $(\overline{a},\ \overline{b},\ \overline{c}) > 0$ $((\overline{a},\ \overline{b},\ \overline{c}) < 0)$;

3) ***для любых векторов*** $\overline{a}, \overline{b}, \overline{c}$ ***имеет место равенство*** $(\overline{a} \times \overline{b},\ \overline{c}) = (\overline{a},\ \overline{b} \times \overline{c})$.

Для доказательства последнего равенства достаточно заметить, что тройки векторов $\overline{a}$, $\overline{b}$, $\overline{c}$ и $\overline{b}$, $\overline{c}$, $\overline{a}$ одновременно правые либо левые, что влечет совпадение знаков чисел $(\overline{a} \times \overline{b},\ \overline{c})$ и $(\overline{b} \times \overline{c},\ \overline{a})$ = $= (\overline{a},\ \overline{b} \times \overline{c})$. Кроме того, оба числа $|(\overline{a} \times \overline{b},\ \overline{c})|$ и $|(\overline{b} \times \overline{c},\ \overline{a})|$ равны объему параллелепипеда, построенного на векторах $\overline{a}$, $\overline{b}$, $\overline{c}$.

Все это и влечет доказываемое равенство.

4.3.4. В заключение получим формулу для вычисления смешанного произведения векторов $\overline{a}$, $\overline{b}$, $\overline{c}$ в декартовых координатах.

Итак, пусть $\overline{a} = (x_1;\ y_1;\ z_1)$, $\overline{b} = (x_2;\ y_2;\ z_2)$, $\overline{c} = (x_3;\ y_3;\ z_3)$ в правой декартовой системе координат. Тогда, как доказано в разделе 4.2.9,

$$\overline{a} \times \overline{b} = \left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix};\ -\begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix};\ \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right).$$

Тем самым

$$(\overline{a},\ \overline{b},\ \overline{c}) = (\overline{a} \times \overline{b},\ \overline{c}) =$$

$$= x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} - y_3 \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix} + z_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

(если разложить последний определитель по третьей строке, см. п. 2.4.3).

4.3.5. В качестве одного из следствий (вытекающих из доказанной формулы и свойств определителей) отметим, что ***при перестановке любых двух сомножителей смешанное произведение меняет знак***. Например,

$$(\overline{a},\ \overline{b},\ \overline{c}) = -(\overline{b},\ \overline{a},\ \overline{c}).$$

ГЛАВА 5
ЛИНЕЙНЫЕ ГЕОМЕТРИЧЕСКИЕ ОБЪЕКТЫ

5.1. МНОЖЕСТВА ТОЧЕК И ИХ УРАВНЕНИЯ

Эта глава посвящена аналитической геометрии – дисциплине, изучающей геометрические объекты аналитическими числовыми методами, по преимуществу методами алгебры. В этой связи мы начнем с напоминания некоторых понятий и ситуаций, знакомых нам еще по школьному курсу математики.

5.1.1. В главе 3 для аналитического описания точек плоскости (пространства) нами было введено понятие декартовой прямоугольной системы координат. При наличии фиксированной декартовой прямоугольной системы координат на плоскости (в пространстве) мы установили взаимно однозначное соответствие между точками плоскости (пространства) и упорядоченными парами (тройками) действительных чисел – координатами точек в данной системе координат.

Мы будем далее говорить, что уравнение $\varphi(x, y) = 0$ $(\varphi(x, y, z) = 0)$ от двух переменных (трех переменных) является ***уравнением множества*** M точек плоскости (пространства) в системе координат Oxy $(Oxyz)$, если:

1) для любой точки $A(\alpha;\ \beta)$ плоскости (или соответственно точки $A(\alpha;\ \beta;\ \gamma)$ пространства) из множества M имеет место равенство $\varphi(\alpha,\ \beta) = 0$ (или $\varphi(\alpha,\ \beta,\ \gamma) = 0$);

2) для любой пары (тройки) чисел $(\alpha,\ \beta)$ (или $(\alpha,\ \beta,\ \gamma)$) такой, что $\varphi(\alpha,\ \beta) = 0$ $(\varphi(\alpha,\ \beta,\ \gamma) = 0)$, точка $A(\alpha;\ \beta)$ $(A(\alpha;\ \beta;\ \gamma))$ принадлежит множеству M; иначе говоря, множество M есть геометри-

ческое место точек, координаты которых являются решением данного уравнения.

5.1.2. Если левая часть уравнения $\varphi(x, y) = 0$ (или $\varphi(x, y, z) = 0$) является многочленом, то M называется ***алгебраическим множеством***. Если число n – это наименьшая из степеней многочленов, стоящих в левых частях уравнений, описывающих множество M, то число n называется ***порядком алгебраического множества*** M.

5.1.3. Выше, в главе 3, были получены формулы взаимосвязи координат точки в разных декартовых прямоугольных системах координат.

Если $\varphi(x, y) = 0$ является уравнением некоторого множества точек плоскости в системе координат Oxy, то для получения уравнения этого множества в системе координат $O'x'y'$ достаточно подставить в уравнение $\varphi(x, y) = 0$ вместо переменных x и y их выражения через переменные x', y' согласно формулам п. 3.4.6. Обозначим полученное уравнение так: $\varphi'(x', y') = 0$.

Можно показать, что если левая часть $\varphi(x, y)$ была многочленом степени n, то многочленом той же степени останется и выражение $\varphi'(x', y')$. Таким образом, порядок алгебраического множества точек плоскости является ***инвариантом***, т. е. не зависит от выбора конкретной декартовой прямоугольной системы координат. Это же верно и для пространственных алгебраических множеств.

Эта глава будет посвящена алгебраическим множествам точек плоскости (пространства), имеющим наименьший порядок – равный единице, т.е. множествам точек плоскости (пространства), описываемых в фиксированной декартовой прямоугольной системе координат уравнениями вида $ax + by + c = 0$ ($ax + by + cz + d = 0$) соответственно.

5.2. ПРЯМАЯ НА ПЛОСКОСТИ

5.2.1. Основным содержанием этого раздела является следующее утверждение.

Теорема. ***Множества первого порядка на плоскости суть прямые, и только они.***

Доказательство. Пусть фиксирована некоторая декартова прямоугольная система координат Oxy на плоскости. Мы должны показать, что любая прямая на плоскости определяется уравнением первой степени и, наоборот, любое уравнение первой степени задает в системе координат Oxy некоторую прямую.

Пусть l – прямая на плоскости Oxy и $A(x_0;\ y_0)$ – некоторая точка на прямой l (рис. 34).

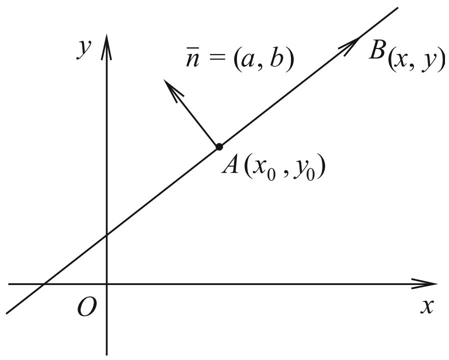


Рис. 34

Выберем и зафиксируем произвольный вектор $\bar{n} = (a;\ b)$, перпендикулярный прямой l. При этом любая точка $B(x; y)$ плоскости лежит на прямой l тогда и только тогда, когда векторы $\bar{n}$ и $\overline{AB}$ перпендикулярны. Последнее же равносильно равенству нулю скалярного произведения векторов $\bar{n}$ и $\overline{AB}$ (см. 4.1.7). Итак,

$$B(x;\ y) \in l \Leftrightarrow (\bar{n},\ \overline{AB}) = 0 \Leftrightarrow$$

$$\Leftrightarrow a(x - x_0) + b(y - y_0) = 0 \Leftrightarrow ax + by + c = 0,$$

где $c = -ax_0 - by_0$. Тем самым прямая l действительно определяется уравнением первой степени в системе координат Oxy, т. е. l есть множество первого порядка.

5.2.2. Покажем теперь обратное. Пусть M – некоторое множество точек плоскости, определяемое в декартовой прямоугольной системе координат Oxy уравнением $ax + by + c = 0$, где хотя бы один из коэффициентов a или b отличен от нуля. Выберем некоторое решение $(x_0,\ y_0)$ уравнения $ax + by + c = 0$ [такое найдется:

например, если $b \neq 0$, то в качестве решения этого уравнения можно взять пару $(0, -c/b)$]. Таким образом, точка $A(x_0; y_0)$ является точкой множества M и $ax_0 + by_0 + c = 0$. Вычитая из уравнения $ax + by + c = 0$ равенство $ax_0 + by_0 + c = 0$, получаем эквивалентное уравнение $a(x - x_0) + b(y - y_0) = 0$, определяющее на плоскости то же самое множество M. Если $B(x; y)$ – произвольная точка плоскости и $\overline{n} = (a; b)$, то уравнение $a(x - x_0) + b(y - y_0) = 0$ соответствует равенству нулю скалярного произведения $(\overline{n}, \overline{AB})$. Очевидно, что совокупность точек B плоскости таких, что вектор $\overline{AB}$ (где A – фиксированная нами точка) перпендикулярен некоторому фиксированному вектору $\overline{n}$, образует прямую.

Итак, произвольное уравнение первой степени $ax + by + c = 0$ задает в любой фиксированной декартовой прямоугольной системе координат на плоскости некоторую прямую. Теорема доказана.

5.2.3. Вектор, перпендикулярный к некоторой прямой l, называется ***нормальным*** к прямой l. Из доказательства теоремы 5.2.1 вытекает, что если $ax + by + c = 0$ – уравнение прямой l, то вектор $\overline{n} = (a;\ b)$ является нормальным к этой прямой.

5.2.4. Заметим, что соответствие между прямой и ее уравнением в данной фиксированной декартовой прямоугольной системе координат не является взаимно однозначным: различные уравнения могут определять одну и ту же прямую. Действительно, если $ax + by + c = 0$ – уравнение прямой l, то для любого действительного числа $d \neq 0$ уравнение $dax + dby + dc = 0$ имеет те же решения, что и уравнение $ax + by + c = 0$ и, значит, определяет ту же прямую l.

5.2.5. Пусть прямые l_1 и l_2 заданы соответственно уравнениями

$$a_1x + b_1y + c_1 = 0$$

и

$$a_2x + b_2y + c_2 = 0,$$

тогда векторы $\overline{n}_1 = (a_1;\ b_1)$, $\overline{n}_2 = (a_2;\ b_2)$ являются нормальными к l_1 и l_2.

Тем самым угол φ между прямыми l_1 и l_2 и угол между векторами $\overline{n}_1$ и $\overline{n}_2$ равны как углы с соответственно перпендикулярными сторонами и, значит,

$$\cos\varphi = \frac{a_1b_1 + a_2b_2}{\sqrt{a_1^2 + b_1^2}\sqrt{a_2^2 + b_2^2}}.$$

В частности, ***прямые*** l_1 и l_2 ***перпендикулярны*** тогда и только тогда, когда $\cos\varphi = 0$, т. е. тогда и только тогда, когда

$$a_1b_1 + a_2b_2 = 0.$$

Прямые l_1 и l_2 ***параллельны*** тогда и только тогда, когда параллельны векторы $\overline{n}_1$ и $\overline{n}_2$, т. е. тогда и только тогда, когда

$$\frac{a_1}{a_2} = \frac{b_1}{b_2}.$$

Любопытно, что если один из знаменателей здесь оказывается равен нулю, то и соответствующий числитель тоже равен нулю, это имеет место при параллельности обеих прямых одной из координатных осей.

5.2.6. Совпадение же прямых l_1 и l_2, заданных указанными выше уравнениями, равносильно параллельности векторов $\overline{n}_1$ и $\overline{n}_2$ и наличию общей точки этих прямых, т. е. общего решения (x_0, y_0) для уравнений

$$a_1x + b_1y + c_1 = 0,$$
$$a_2x + b_2y + c_2 = 0.$$

Параллельность векторов $\overline{n}_1$ и $\overline{n}_2$ выражается в пропорциональности координат: $\dfrac{a_1}{a_2} = \dfrac{b_1}{b_2}$. Обозначим значение этих дробей через k, тогда $a_1 = a_2k,\ b_1 = b_2k$. В то же время справедливы равенства $a_1x_0 + b_1y_0 + c_1 = 0,$ а также $a_2x_0 + b_2y_0 + c_2 = 0.$ Умножая второе равенство на число k и вычитая результат из первого равенства, с учетом того, что $a_1 = a_2k,\ b_1 = b_2k,$ получаем равенство $c_1 = c_2k.$

Таким образом, если уравнения $a_1x + b_1y + c_1 = 0$ и $a_2x + b_2y + c_2 = 0$ определяют одну и ту же прямую, то имеет место пропорциональность

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}.$$

Мы уже видели в п. 5.2.4 справедливость обратного утверждения, и тем самым ***уравнения*** $a_1x + b_1y + c_1 = 0$ и $a_2x + b_2y + c_2 = 0$ ***определяют одну и ту же прямую тогда и только тогда, когда выполняются равенства***

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}.$$

5.2.7. Мы хорошо помним из школьного курса математики, что через любые две различные точки плоскости $A(x_1;\, y_1)$ и $B(x_2;\, y_2)$ проходит единственная прямая l_{AB}. Возникает вопрос об уравнении этой прямой. Очевидно, что произвольная точка $C(x;\, y)$ лежит на прямой l_{AB} тогда и только тогда, когда векторы $\overline{AB}$ и $\overline{AC}$ параллельны, т. е. тогда и только тогда, когда пропорциональны координаты векторов $\overline{AB} = (x_2 - x_1;\, y_2 - y_1)$ и $\overline{AC} = (x - x_1;\, y - y_1)$. Таким образом, уравнение

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}$$

является уравнением прямой, проходящей через точки $A(x_1;\, y_1)$ и $B(x_2;\, y_2)$. Вектор $\overline{AB}$, параллельный рассматриваемой прямой, называется ***направляющим вектором*** прямой.

5.2.8. Рассмотрим, наконец, еще одну форму уравнения прямой l на плоскости – так называемое ***нормальное уравнение прямой*** l. Нормирующим множителем μ уравнения $ax + by + c = 0$ назовем число $\pm\dfrac{1}{\sqrt{a^2 + b^2}}$, где знак нормирующего множителя μ выбирается так, чтобы число μc было отрицательным (если же $c = 0$, знак μ произволен). Как уже отмечалось в 5.2.6, уравнение $\mu ax+\mu by+\mu c=0$, получаемое из первоначального уравнения $ax+by+c=0$ умножением обеих частей на число μ, определяет ту же прямую l. Назовем полученное уравнение $\mu ax+\mu by+\mu c=0$ ***нормальным уравнением прямой*** l. Обозначим через α (или β) угол, образуемый вектором нормали к прямой l, отложенным из начала координат в сторону прямой и осью Ox (или Oy).

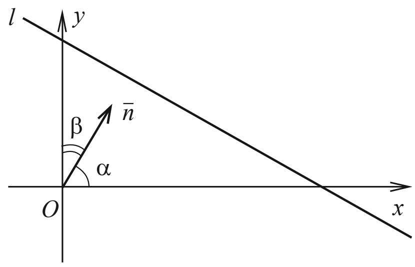


Рис. 35

Теорема. ***Пусть*** $ax+by+c=0$ ***– нормальное уравнение прямой*** l***, тогда*** $a=\cos\alpha$; $b=\cos\beta$, ***число*** $(-c)$ ***равно расстоянию от начала координат до прямой*** l ***и если*** $A(x_0;\, y_0)$ ***– произвольная точка плоскости, то число*** $|ax_0+by_0+c|$ ***(равное модулю результата, полученного после подстановки координат точки*** A ***в левую часть уравнения прямой) совпадает с расстоянием от точки*** A ***до прямой*** l***.***

Доказательство. Уравнение $ax+by+c=0$ получено из произвольного уравнения $a_1x+b_1y+c_1=0$ прямой l умножением на нормирующий множитель

$$\mu=\pm\frac{1}{\sqrt{a_1^2+b_1^2}}.$$

Таким образом, $a=\mu a_1$, $b=\mu b_1$. Вектор $\bar{n}=(a;\, b)$ нормали к прямой l перпендикулярен самой прямой l и имеет длину, равную единице.

Пусть точка $K(x_1;\, y_1)$ – это основание перпендикуляра, опущенного из начала координат O на прямую l (в случае, когда прямая l проходит через начало координат, K совпадает с O, рис. 36).

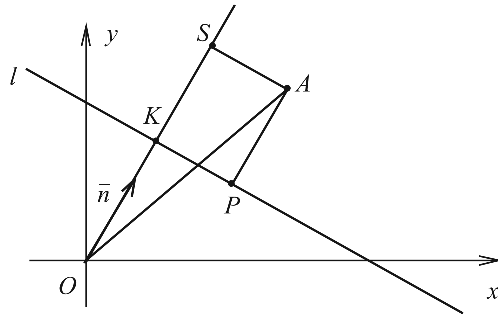


Рис. 36

Тогда $ax_1 + by_1 + c = 0$, т. е.

$$-c = ax_1 + by_1 = ((a;\ b),\ (x_1;\ y_1)) =$$

$$= (\overline{n},\ \overline{OK}) = |\overline{n}| \cdot |\overline{OK}| \cos\angle\overline{n}\,\overline{OK} = |\overline{OK}|,$$

так как по выбору множителя μ число $(-c)$ неотрицательно; следовательно, угол между векторами $\overline{n}$ и $\overline{OK}$ равен нулю: во-первых, они параллельны как два перпендикуляра к прямой l; во-вторых, $\cos\angle\overline{n}\,\overline{OK} > 0$. Итак, векторы $\overline{n}$ и $\overline{OK}$ сонаправлены. А число $-c$ равно расстоянию от начала координат до прямой l и

$$\cos\alpha = \frac{a}{|\overline{n}|} = a, \quad \cos\beta = \frac{b}{|\overline{n}|} = b.$$

Первое утверждение теоремы доказано.

Пусть $A(x_0;\ y_0)$ – произвольная точка плоскости, а P – основание перпендикуляра, опущенного из точки A на прямую l; S – основание перпендикуляра, опущенного из точки A на прямую, перпендикулярную прямой l и проходящую через начало координат. Имеем цепочку равенств

$$ax_0 + by_0 = (\overline{n},\ \overline{OA}) = |\overline{n}||\overline{OA}| \cos\angle\,\overline{n}\,\overline{OA} = |\overline{OA}| \cos\angle\,\overline{n}\,\overline{OA} = |\overline{OS}|.$$

Расстояние от точки A до прямой l равно

$$\rho(A,\ l) = \Big|\, |\overline{OS}| - |\overline{OK}|\, \Big| = |\, ax_0 + by_0 + c\,|,$$

что и требовалось доказать. В общем случае с нормалью $\overline{n} = (a;\ b)$ произвольной длины $\rho(A(x_0;\ y_0),\ l) = \dfrac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$.

5.2.9. Прямая l делит плоскость на две полуплоскости, и естественным является вопрос: лежат ли две данные точки $A(x_1;\ y_1)$, $B(x_2;\ y_2)$ по одну сторону от прямой l или по разные? Пусть $ax + by + c = 0$ – произвольное уравнение прямой l и $K(x_0;\ y_0)$ – основание перпендикуляра, опущенного из начала координат на прямую l. Тогда $ax_0 + by_0 + c = 0$ и, значит, $-c = ax_0 + by_0 = (\overline{n}, \overline{OK})$, где $\overline{n}$ – это вектор нормали $(a;\ b)$ к прямой l. Подставив координаты точки $A(x_1,\ y_1)$ в левую часть уравнения прямой l, получаем

$$ax_1 + by_1 + c = ax_1 + by_1 - (-c) = (\overline{n},\ \overline{OA}) - (\overline{n},\ \overline{OK}) =$$

$$= (\overline{n},\ \overline{OA} - \overline{OK}) = (\overline{n},\ \overline{KA}) = |\overline{n}|\left|\overline{KA}\right| \cos\angle\overline{n}\ \overline{KA}.$$

Аналогично, подставив координаты точки $B(x_2;\ y_2)$ в левую часть уравнения прямой l, получаем

$$ax_2 + by_2 + c = |\overline{n}|\left|\overline{KB}\right| \cos\angle\overline{n}\ \ \overline{KB}.$$

Тем самым числа $(ax_1 + by_1 + c)$ и $(ax_2 + by_2 + c)$ имеют одинаковые знаки тогда и только тогда, когда совпадают знаки $\cos\angle\overline{n}\,\overline{KA}$ и $\cos\angle\overline{n}\,\overline{KB}$, т. е. тогда и только тогда, когда углы $\angle\overline{n}\,\overline{KA}$ и $\angle\overline{n}\,\overline{KB}$ либо одновременно острые, либо одновременно тупые. Таким образом, числа, получаемые в результате подстановки координат пары точек в левую часть уравнения прямой, имеют одинаковые знаки тогда и только тогда, когда эти точки лежат по одну сторону от прямой.

5.3. ПЛОСКОСТЬ В ПРОСТРАНСТВЕ

Ситуация, связанная с геометрией плоскости в пространстве и уравнениями плоскости, полностью аналогична ситуации, рассмотренной в предыдущем пункте относительно прямой на плоскости. В связи с этим мы лишь сформулируем соответствующие результаты, оставляя читателю их доказательства, практически дословно совпадающие с доказательствами аналогичных результатов из раздела 5.2.

5.3.1. Теорема. ***Множества первого порядка в пространстве – это плоскости и только они.***

Таким образом, уравнения плоскостей в пространстве имеют вид

$$ax + by + cz + d = 0,$$

где хотя бы одно из чисел a, b, c отлично от нуля.

Вектор, перпендикулярный плоскости π, называется ***нормальным*** к плоскости π вектором.

Если плоскость π задана уравнением $ax + by + cz + d = 0$, то вектор $\overline{n} = (a;\ b;\ c)$ является нормальным к этой плоскости.

Пусть плоскости π_1 и π_2 заданы уравнениями

$$a_1x + b_1y + c_1z + d_1 = 0 \quad \text{и} \quad a_2x + b_2y + c_2z + d_2 = 0$$

соответственно. Тогда, если φ – угол между плоскостями π_1 и π_2, то

$$\cos\varphi = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}\sqrt{a_2^2 + b_2^2 + c_2^2}}.$$

Плоскости π_1 и π_2 перпендикулярны тогда и только тогда, когда

$$a_1a_2 + b_1b_2 + c_1c_2 = 0.$$

Плоскости π_1 и π_2 параллельны тогда и только тогда, когда

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$$

(если здесь один или два из знаменателей равны нулю, то и в соответствующих числителях тоже стоит нуль).

Плоскости π_1 и π_2 совпадают тогда и только тогда, когда

$$\frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{c_1}{c_2}=\frac{d_1}{d_2}.$$

5.3.2. ***Нормирующим множителем*** μ уравнения

$$ax+by+cz+d=0$$

назовем число

$$\mu=\pm\frac{1}{\sqrt{a^2+b^2+c^2}},$$

где знак нормирующего множителя выбирается таким образом, чтобы число μd было отрицательным (если $d=0$, то знак μ произволен). Уравнение $\mu ax+\mu by+\mu cz+\mu d=0$, получаемое умножением уравнения $ax+by+cz+d=0$ плоскости π на нормирующий множитель, называется ***нормальным уравнением плоскости*** π. Через α (β, γ) обозначим угол, образуемый вектором нормали к плоскости π, отложенным из начала координат в сторону плоскости π, с осью Ox (Oy, Oz соответственно).

Теорема. ***Пусть*** $ax+by+cz+d=0$ ***– нормальное уравнение плоскости*** π, ***тогда*** $a=\cos\alpha$, $b=\cos\beta$, $c=\cos\gamma$, ***а число*** $-d$ ***равно расстоянию от начала координат до плоскости*** π; ***и если*** $A(x_0;\, y_0;\, z_0)$ ***– произвольная точка пространства, то число*** $\rho(A,\ \pi)=|ax_0+by_0+cz_0+d|$ **(*равное модулю результата, получающегося от подстановки координат точки*** A ***в левую часть нормального уравнения*** π**) *совпадает с расстоянием от точки*** A ***до плоскости*** π. ***Если*** $ax+by+cz+d=0$ ***– произвольное уравнение плоскости*** π, ***то*** $\rho(A,\ \pi)=\dfrac{|ax_0+by_0+cz_0+d|}{\sqrt{a^2+b^2+c^2}}$.

5.3.3. Плоскость π с уравнением $ax+by+cz+d=0$ делит пространство на два полупространства. Точки $A(x_0;\, y_0;\, z_0)$ и $B(x_1;\, y_1;\, z_1)$ лежат по одну сторону от плоскости π тогда и только тогда, когда числа $ax_0 + by_0 + cz_0 + d$ и $ax_1 + by_1 + cz_1 + d$ имеют одинаковые знаки.

5.3.4. В заключение этого раздела найдем уравнение плоскости, проходящей через три точки пространства $A(x_1;\ y_1;\ z_1)$, $B(x_2;\ y_2;\ z_2)$, $C(x_3;\ y_3;\ z_3)$, не лежащие на одной прямой. Обозначим плоскость, проходящую через точки A, B, C, как π_{ABC}. Заметим, что условие того, что точки A, B, C не лежат на одной прямой, равносильно условию о том, что векторы $\overline{AB}$ и $\overline{AC}$ не параллельны, т. е. равенства

$$\frac{x_2 - x_1}{x_3 - x_1} = \frac{y_2 - y_1}{y_3 - y_1} = \frac{z_2 - z_1}{z_3 - z_1}$$

нарушаются хотя бы в одном месте. Пусть $D(x;\ y;\ z)$ – произвольная точка пространства. Тогда $D \in \pi_{ABC}$ тогда и только тогда, когда векторы $\overline{AB}$, $\overline{AC}$, $\overline{AD}$ компланарны. Последнее же имеет место тогда и только тогда, когда смешанное произведение векторов

$$\overline{AB} = (x_2 - x_1;\ y_2 - y_1;\ z_2 - z_1),$$

$$\overline{AC} = (x_3 - x_1;\ y_3 - y_1;\ z_3 - z_1),$$

$$\overline{AD} = (x - x_1;\ y - y_1;\ z - z_1)$$

равно нулю (см. следствие 1 п. 4.3.3(1)). Тем самым уравнение плоскости π_{ABC} имеет вид

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

Из этого уравнения легко получается так называемое ***уравнение плоскости в отрезках на осях:***

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

Здесь a, b, c – ненулевые отрезки, отсекаемые плоскостью на осях координат (если плоскость не пересекает какую-то из осей, соответствующее слагаемое в левой части отсутствует).

Для решения многих задач полезно ***уравнение плоскости, проходящей через данную точку*** $M_0(x_0; y_0; z_0)$ ***перпендикулярно данному вектору*** $\overline{n} = (a; b; c)$:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

5.4. ПРЯМАЯ В ПРОСТРАНСТВЕ

5.4.1. Рассмотрим в пространстве пару плоскостей π_1 и π_2, имеющих в некоторой декартовой прямоугольной системе координат уравнения

$$a_1x + b_1y + c_1z + d_1 = 0$$

и (1)

$$a_2x + b_2y + c_2z + d_2 = 0$$

соответственно. Условием того, что плоскости π_1 и π_2 не параллельны, является то, что в последовательности равенств

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$$

хотя бы одно нарушается (см. 5.3.1) Запишем этот факт так:

$$\cancel{\left(\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}\right)}. \tag{2}$$

Пересечением непараллельных плоскостей π_1 и π_2 является некоторая прямая l. Таким образом, система уравнений первой степени

$$\begin{cases} a_1x + b_1y + c_1z + d_1 = 0, \\ a_2x + b_2y + c_2z + d_2 = 0 \end{cases}$$

при условии

$$\cancel{\left(\frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{c_1}{c_2}\right)}$$

определяет прямую в пространстве. Очевидно верно и обратное: любую прямую l в пространстве можно представить как пересечение двух непараллельных плоскостей π_1 и π_2 и, значит, можно определить некоторой парой уравнений вида (1) при условии (2).

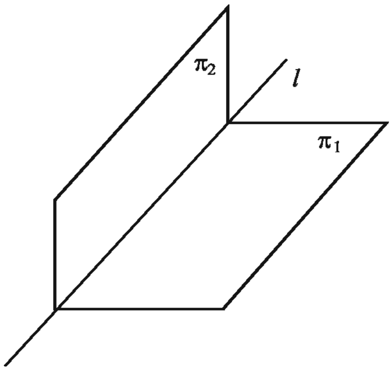


Рис. 37

Поскольку выбор плоскостей π_1 и π_2, по существу, случаен (в чем мы убеждаемся, «вращая» эти плоскости вокруг оси l), система уравнений (1) несет в себе не так много непосредственной информации о «геометрии» прямой l, как хотелось бы (рис. 37).

5.4.2. Более удобными представляются уравнения прямой в пространстве, задаваемой как ***прямая, проходящая через две различные точки***. Пусть этими фиксированными точками будут $A(x_1;\, y_1;\, z_1)$ и $B(x_2;\, y_2;\, z_2)$. Тогда произвольная точка $C(x;\, y;\, z)$ принадлежит прямой l_{AB}, проходящей через точки A и B, если и только если векторы $\overline{AB}$ и $\overline{AC}$ параллельны. Таким образом, в соответствии с п. 3.1.5 уравнения (т.е. *система уравнений*) прямой l_{AB}, проходящей через точки A и B, имеют вид

$$\frac{x-x_1}{x_2-x_1}=\frac{y-y_1}{y_2-y_1}=\frac{z-z_1}{z_2-z_1}. \tag{3}$$

Аналогичным образом выглядят уравнения прямой l, определенной такими условиями: прямая l содержит некоторую точку $A(x_1;\, y_1;\, z_1)$ и параллельна некоторому вектору $\overline{a}=(m;\, n;\, p)$:

$$\frac{x-x_1}{m}=\frac{y-y_1}{n}=\frac{z-z_1}{p}. \tag{4}$$

Это так называемые ***канонические уравнения прямой в пространстве.*** Как всегда, если один или два знаменателя обращаются в нуль, то соответствующие числители равны нулю.

Из этих уравнений легко выводятся ***параметрические уравнения прямой в пространстве.*** Для этого достаточно все отношения приравнять к *параметру* t и выразить из полученных равенств x, y, z:

$$\begin{aligned} x &= x_0 + mt, \\ y &= y_0 + nt, \\ z &= z_0 + pt, \end{aligned} \tag{5}$$

причем здесь параметр t пробегает всевозможные действительные значения.

5.4.3. Естественным представляется вопрос о переходе от уравнений вида (1), (2) к уравнениям вида (4).

Так как вектор $\bar{n}_1 = (a_1;\ b_1;\ c_1)$ перпендикулярен плоскости π_1, то он перпендикулярен и прямой l. По той же причине перпендикулярен прямой и вектор $\bar{n}_2 = (a_2;\ b_2;\ c_2)$. Но тогда векторное произведение этих векторов, а именно вектор

$$\bar{n}_1 \times \bar{n}_2 = \left(\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix},\ -\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix},\ \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \right)$$

будет параллелен прямой l и окажется ее направляющим вектором (рис. 38).

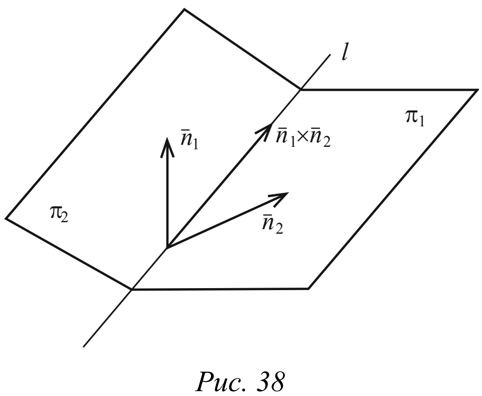


Рис. 38

Тем самым, если $(x_0; y_0; z_0)$ – некоторое решение системы (1), то прямая l, заданная уравнениями (1), может быть задана и уравнениями

$$\frac{x-x_0}{\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}} = \frac{y-y_0}{-\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}} = \frac{z-z_0}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}.$$

5.4.4. Пусть теперь пара прямых l_1 и l_2 задана уравнениями

$$\frac{x-x_1}{m_1} = \frac{y-y_1}{n_1} = \frac{z-z_1}{p_1}$$

и

$$\frac{x-x_2}{m_2} = \frac{y-y_2}{n_2} = \frac{z-z_2}{p_2}$$

соответственно.

Угол φ между прямыми l_1 и l_2 совпадает с углом между их направляющими векторами $\overline{a}_1 = (m_1; n_1; p_1)$ и $\overline{a}_2 = (m_2; n_2; p_2)$ и тем самым

$$\cos\varphi = \frac{m_1 m_2 + n_1 n_2 + p_1 p_2}{\sqrt{m_1^2 + n_1^2 + p_1^2}\sqrt{m_2^2 + n_2^2 + p_2^2}}.$$

Заметим, что при взятии числителя под модуль угол φ окажется острым или прямым, даже если угол $\angle \overline{a}_1 \overline{a}_2$ тупой.

Прямые l_1 и l_2 ***перпендикулярны*** (или, что то же самое, перпендикулярны векторы $\overline{a}_1$ и $\overline{a}_2$) тогда и только тогда, когда

$$m_1 m_2 + n_1 n_2 + p_1 p_2 = 0.$$

Прямые l_1 и l_2 ***параллельны*** (или, что то же самое, параллельны векторы $\overline{a}_1$ и $\overline{a}_2$) тогда и только тогда, когда

$$\frac{m_1}{m_2} = \frac{n_1}{n_2} = \frac{p_1}{p_2}.$$

Так как точки $A_1(x_1;\, y_1;\, z_1)$ и $A_2(x_2;\, y_2;\, z_2)$ лежат на прямых l_1 и l_2 соответственно, то прямые l_1 и l_2 лежат в одной плоскости тогда и только тогда, когда компланарны векторы $\overline{A_1A_2}$, $\bar{a}_1$, $\bar{a}_2$, т. е. когда смешанное произведение этих векторов равно нулю:

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ m_1 & n_1 & p_1 \\ m_2 & n_2 & p_2 \end{vmatrix} = 0.$$

Таким образом, прямые l_1 и l_2 ***пересекаются*** (лежат в одной плоскости и не параллельны) тогда и только тогда, когда

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ m_1 & n_1 & p_1 \\ m_2 & n_2 & p_2 \end{vmatrix} = 0 \quad \text{и} \quad \cancel{\frac{m_1}{m_2} = \frac{n_1}{n_2} = \frac{p_1}{p_2}}\,.$$

Если же прямые не лежат в одной плоскости, или, иными словами, скрещиваются, выполняется необходимое и достаточное условие

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ m_1 & n_1 & p_1 \\ m_2 & n_2 & p_2 \end{vmatrix} \neq 0.$$

5.4.5. Рассмотрим теперь прямую l с уравнениями

$$\frac{x - x_1}{m} = \frac{y - y_1}{n} = \frac{z - z_1}{p}$$

и плоскость π с уравнением

$$ax + by + cz + d = 0.$$

Предоставляем читателю самому убедиться, что:

1) если φ – это угол между прямой l и плоскостью π, то

$$\sin\varphi = \frac{|\, am + bn + cp \,|}{\sqrt{a^2 + b^2 + c^2}\sqrt{m^2 + n^2 + p^2}}\,;$$

2) прямая l ***параллельна плоскости*** π (включая случай, когда l лежит в π) тогда и только тогда, когда

$$am + bn + cp = 0;$$

3) прямая l ***перпендикулярна плоскости*** π тогда и только тогда, когда

$$\frac{a}{m} = \frac{b}{n} = \frac{c}{p};$$

4) прямая l ***лежит в плоскости*** π тогда и только тогда, когда

$$am + bn + cp = 0 \quad \text{и} \quad ax_1 + by_1 + cz_1 + d = 0.$$

ГЛАВА 6
ОБЩАЯ ТЕОРИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

6.1. ГЕОМЕТРИЧЕСКОЕ ВВЕДЕНИЕ

При решении разнообразных задач аналитической геометрии нередко возникают две проблемы, после некоторой формализации принимающие совершенно одинаковый вид.

6.1.1. Установить, возможно ли представить вектор $\overline{d} = (d_1;\ d_2;\ d_3)$ в виде линейной комбинации векторов $\overline{a} = (a_1;\ a_2;\ a_3)$, $\overline{b} = (b_1;\ b_2;\ b_3)$ и $\overline{c} = (c_1;\ c_2;\ c_3)$, и найти коэффициенты такой линейной комбинации, т. е. мы должны узнать, разрешимо ли векторное уравнение

$$x\overline{a} + y\overline{b} + z\overline{c} = \overline{d}$$

и при каких численных значениях букв x, y и z оно удовлетворяется.

Поскольку векторы заданы в виде координатных троек, одно векторное равенство можно переписать в виде трех числовых:

$$\begin{cases} xa_1 + yb_1 + zc_1 = d_1, \\ xa_2 + yb_2 + zc_2 = d_2, \\ xa_3 + yb_3 + zc_3 = d_3. \end{cases}$$

Задача заключается в том, чтобы найти такие числа x, y и z, чтобы все три равенства удовлетворялись одновременно, или установить, что их не существует.

6.1.2. Найти множество общих точек, т. е. пересечение нескольких плоскостей, задаваемых общими уравнениями:

$$a_i x + b_i y + c_i z = d_i \quad i = 1, \ldots, k.$$

Для случая $k = 2$ это задача перехода от общих уравнений прямой (см. п. 5.4.1):

$$\begin{cases} a_1 x + b_1 y + c_1 z = d_1, \\ a_2 x + b_2 y + c_2 z = d_2 \end{cases}$$

к каноническим или параметрическим.

Для случая $k = 3$ получим полное соответствие с задачей 6.1.1:

$$\begin{cases} a_1 x + b_1 y + c_1 z = d_1, \\ a_2 x + b_2 y + c_2 z = d_2, \\ a_3 x + b_3 y + c_3 z = d_3. \end{cases}$$

В случае двух плоскостей понятно, что их пересечение даст прямую, если плоскости действительно пересекаются в обычном геометрическом смысле, а для этого нормали их должны быть неколлинеарны (см. п. 5.4.1):

$$\overline{n}_1 = (a_1;\ b_1;\ c_1) \not\parallel \overline{n}_2 = (a_2;\ b_2;\ c_2)$$

или, что то же самое, координаты нормалей непропорциональны, т. е. все три равенства

$$a_1 = sa_2,\ b_1 = sb_2,\ c_1 = sc_2$$

не выполняются одновременно ни для какого числа s, тем самым выполняется условие (2) из п. 5.4.1.

Если же нормали коллинеарны (что можно распознать по равенству $\overline{n}_1 \times \overline{n}_2 = \overline{0}$, см. 4.2.4), то существуют две возможности, указанные в п. 5.3.1:

• либо обе плоскости параллельны – это наблюдается в том случае, когда пропорциональность координат нормалей (и, следовательно, левых частей уравнений плоскостей) не продолжается пропорциональностью правых частей

$$a_1 x + b_1 y + c_1 z = sa_2 x + sb_2 y + sc_2 z = s(a_2 x + b_2 y + c_2 z) = sd_2 \neq d_1$$

и, значит, оба они одновременно не удовлетворяются никаким набором значений $(x; y; z)$;

• либо $d_1 = sd_2$, и оба уравнения пропорциональны с коэффициентом пропорциональности s, а значит, удовлетворяются только одновременно и на самом деле описывают одну и ту же плоскость.

Итак, в случае $k = 2$ мы имеем три возможности для множества общих точек плоскостей:

• прямая, если коэффициенты уравнений не пропорциональны;

• пустое множество, если пропорциональность коэффициентов уравнений нарушается их правыми частями;

• плоскость, если и коэффициенты уравнений, и правые части пропорциональны.

В случае $k = 3$, когда ищется пересечение трех плоскостей, к этим возможностям добавляется четвертая: единственная точка (точные формулировки, соответствующие этим возможностям, оставляются читателю).

Однако вопрос о пересечении более чем трех плоскостей с чисто геометрической точки зрения может привести в смущение. Хотя ясно, что возможен только один из четырех указанных выше ответов и в каждом конкретном случае его отыскание не представляет трудностей, сформулировать геометрические условия каждой из этих возможностей достаточно сложно.

По этой и многим другим причинам изучаются системы линейных алгебраических уравнений с произвольным числом уравнений и неизвестных, к чему мы и переходим.

6.2. ПОНЯТИЕ И ФОРМЫ ЗАПИСИ СИСТЕМЫ. ТЕОРЕМА КРОНЕКЕРА–КАПЕЛЛИ

6.2.1. ***Системой линейных алгебраических уравнений*** называется совокупность одновременно рассматриваемых равенств

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

Левые части этих равенств представляют собой линейные комбинации букв $x_1,\ x_2,\ldots,\ x_n$, называемых ***неизвестными***, с коэффициентами a_{ij} из числового поля (заметим, что в паре номеров ij, стоящей возле каждого из таких коэффициентов, левый номер i соответствует номеру уравнения, правый номер j – номеру неизвестного). ***Правые части*** $b_1,\ldots,\ b_n$ берутся из того же числового поля и называются ***свободными членами***.

Например, в пунктах 6.1.1 и 6.1.2 возникали системы из трех уравнений с тремя неизвестными $x_1 = x,\ x_2 = y,\ x_3 = z$; коэффициентами системы были координаты векторов – базисных или нормальных; в п. 6.1.2 рассматривалась также система из двух уравнений с тремя неизвестными.

Набор чисел $x_1^0,\ldots,\ x_n^0$ называется ***частным решением*** системы уравнений, если при подстановке этих чисел вместо букв с теми же номерами (т. е. если положить $x_j = x_j^0,\ j = 1,\ldots,n$) все уравнения системы обратятся в верные равенства.

Общим решением системы уравнений называется множество всех ее частных решений.

Скажем, для двух плоскостей

$$\pi_1 : 2x - y + z = 1,$$

$$\pi_2 : x + y - 3z = 0$$

принадлежащая им обеим точка с координатами (1; 2; 1) будет частным решением системы этих двух уравнений; а общее решение системы представляет собой прямую – пересечение этих плоскостей:

$$\begin{cases} x = 1 + 2t, \\ y = 2 + 7t, \\ z = 1 + 3t \end{cases} \quad \text{или} \quad \frac{x-1}{2} = \frac{y-2}{7} = \frac{z-1}{3}.$$

6.2.2. При работе с системами уравнений иногда бывают удобны другие формы записи системы, кроме той ***полной,*** или ***развернутой***, которая приведена выше в п. 6.2.1.

Прежде всего напрашивается использование значка суммы в левых частях уравнений, а вместо того, чтобы выписывать их все, достаточно использовать переменный индекс $i = 1,\ldots, m$:

$$\sum_{j=1}^{n} a_{ij} x_j = b_i.$$

Такую форму записи будем называть ***сокращенной***.

Если вспомнить, как возникла система уравнений в п. 6.1.1, то можно представить набор коэффициентов при каждом неизвестном x_j как m-координатный вектор

$$\vec{a}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Точно так же из правых частей образуем вектор $\vec{b}$ и m числовых равенств превратим в одно векторное:

$$\vec{a}_1 x_1 + \cdots + \vec{a}_n x_n = \vec{b}.$$

Такую форму записи системы назовем ***векторно-столбцовой***.

Наконец, можно привлечь матричные операции. Поскольку в левых частях уравнений системы стоят матричные произведения (см. п. 2.1.5): числовых строк $(a_{i1} \cdots a_{in})$ на буквенный столбец $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, всю систему можно представить одним матричным уравнением

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Числовая матрица $A = (a_{ij})_{m\times n}$ называется ***матрицей системы***, одностолбцовая буквенная матрица $X = (x_j)_{n\times 1}$, как и выше, называется ***столбцом неизвестных***, а одностолбцовая же матрица $B=(b_i)_{m\times 1}$ – ***столбцом правых частей*** или ***столбцом свободных членов***.

Такая форма записи будет называться ***матричной***: $AX=B$ (мы уже пользовались ею в п. 2.5.5).

6.2.3. С геометрической точки зрения параллельные плоскости ничем не хуже пересекающихся или совпадающих плоскостей. С точки зрения других, чрезвычайно разнообразных приложений систем линейных уравнений две эти возможности отличаются принципиально: уравнения параллельных плоскостей образуют систему, не имеющую решений, или ***несовместную***, уравнения пересекающихся или совпадающих плоскостей – систему, имеющую решение (и даже множество решений), или ***совместную***. Первый вопрос, который возникает в этой связи, таков: как распознавать совместные и несовместные системы? Оказывается, это можно сделать, используя понятие ***ранг***.

Напомним, что (конечное) множество векторов $a_1,\ldots,a_n$ векторного пространства $L_{\mathbf{F}}$ над полем **F** называется линейно ***независимым***, если из равенства

$$\alpha_1 a_1+\ldots+\alpha_n a_n=\underline{0}$$

для некоторых чисел $\alpha_1,\ldots,\alpha_n\in\mathbf{F}$ следует, что все они равны нулю: $\alpha_1=\ldots=\alpha_n=0$. Это равносильно тому, что ни один из векторов $a_1,\ldots,a_n$ не может быть выражен через остальные при помощи линейных операций (т. е. сложения и умножения на числа из **F**) (см. п. 3.3.2).

Если дано некоторое множество векторов $a_1,\ldots,a_n\in L_{\mathbf{F}}$, то ***ранговой*** (***базисной***) ***подсистемой*** такого множества называются линейно независимые векторы

$$a_{i_1},\ldots,a_{i_r},\ \{i_1,\ldots,i_r\}\subseteq\{1,\ldots,n\},$$

через которые можно выразить любой из оставшихся векторов a_i (п. 3.3.5). Иными словами, подсистема $a_{i_1},\ldots,a_{i_r}$ является максимальной линейно независимой подсистемой множества $\{a_1,\ldots,a_n\}$

и добавление еще одного вектора из этого множества сделает ее линейно зависимой. ***Рангом*** множества векторов называется минимальное число векторов базисной подсистемы этого множества. В конце концов мы докажем, что все базисные подсистемы множества векторов (не обязательно конечного) содержат одинаковое количество элементов (см. п. 7.3.1); пока же в определении стоит слово «минимальное», иными словами, ранг приравнивается к количеству векторов в самой компактной базовой подсистеме.

Чтобы проиллюстрировать это понятие, укажем, что по теореме 3.4.1:

• одноэлементное множество, состоящее из нулевого вектора, имеет ранг 0: ведь этот вектор линейно зависим ($1 \cdot \underline{0} = \underline{0}$, а $1 \neq 0$, т. е. линейно независимым нуль-вектор не является), и базисных подсистем это множество не имеет;

• любое множество коллинеарных векторов, содержащее ненулевые элементы, имеет ранг 1: ведь все коллинеарные векторы выражаются через любой из них, не равный нулю, с помощью умножения на число, т. е. базисные подсистемы одноэлементны;

• множество компланарных, но не коллинеарных векторов имеет ранг 2, так как любой вектор такого множества раскладывается по неколлинеарной паре векторов этого же множества, – такие неколлинеарные пары и будут базисными;

• наконец, множество некомпланарных геометрических векторов имеет ранг 3.

6.2.4. Теорема Кронекера–Капелли. Назовем ***столбцовым рангом матрицы*** ранг множества ее столбцов, рассматриваемых как векторы с обычными операциями.

Матрицу системы линейных уравнений мы определили в п. 6.2.2, а ***расширенной матрицей*** системы называется та же матрица с добавлением столбца свободных членов в качестве (n + 1)-го столбца:

$$(A \mid B)_{m\times(n+1)} = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right).$$

Аналогично столбцовому можно определить ***строчный ранг матрицы***; впоследствии (см. 7.3.2) окажется, что эти ранги равны между собой и указание, какой именно ранг имеется в виду, не обязательно. Пока же мы будем считать рангом матрицы столбцовый, но в формулировку теоремы это слово включать не будем.

Напомним (раздел 3.3), что любая линейная комбинация некоторого множества векторов может быть записана через линейную комбинацию ее базисной подсистемы: для этого достаточно выразить все векторы, не входящие в базисную подсистему, через векторы базисной подсистемы, и подставить эти выражения в исходную комбинацию.

Теорема (Кронекера–Капелли). ***Система линейных алгебраических уравнений совместна тогда и только тогда, когда ранг матрицы системы равен рангу ее расширенной матрицы:*** $r(A) = r(A|B)$.

Доказательство состоит из цепочки равносильных утверждений.

Пусть система совместна. Тогда существуют такие числа $x_1^0, \ldots, x_n^0$, что удовлетворяется векторно-столбцовая форма системы:

$$x_1^0 \overset{|}{a}_1 + \ldots + x_n^0 \overset{|}{a}_m = \overset{|}{b}.$$

Тогда столбец правых частей $\overset{|}{b}$, выражаясь через все столбцы матрицы A, тем самым выражается через столбцы минимальной ранговой подсистемы.

Следовательно, приписывание к матрице A столбца $B = \overset{|}{b}$ не расширяет минимальной ранговой подсистемы и не приводит к увеличению ранга матрицы (разумеется, к уменьшению тоже).

Значит, $r(A) = r(A|B)$.

Просматривая эти утверждения в обратном порядке, мы увидим, что из каждого следующего вытекает предыдущее и тем самым получается обратное утверждение теоремы: из равенства рангов следует, что система совместна.

Это дает критерий совместности системы линейных уравнений, хотя пока и непонятно, как вычислять ранги матриц.

6.3. ЛИНЕЙНОЕ МНОГООБРАЗИЕ

Мы попросим читателя освежить в памяти понятия линейного пространства и подпространства (пп. 3.2.2, 3.3.17).

6.3.1. Более естественным, с геометрической точки зрения, чем линейное подпространство, является понятие ***линейное многообразие***. Пусть W – подпространство линейного пространства L, а m – вектор, не лежащий в этом подпространстве: $m \in L \setminus W$. Тогда линейным многообразием M называется такое подмножество пространства L, каждый элемент которого представляется в виде

$$u = m + w,$$

где $w \in W$. Символически это записывается так:

$$M = m + W.$$

Если попытаться превратить данное определение в понятие, то можно сказать, что линейное многообразие – это подпространство W, сдвинутое на «посторонний» элемент m.

Обратившись к пространству геометрических векторов, «выходящих из нуля», и уяснив, что подпространства там – это прямые и плоскости, проходящие через начало координат (ср. с п. 3.3.18), можно сдвигать их на «посторонние» им (т. е. неколлинеарные прямой или некомпланарные плоскости) векторы и получить всевозможные прямые и плоскости.

И потому с виду многообразия и подпространства почти неотличимы, зато алгебраическая разница весьма велика. В многообразиях в общем случае нет не только нулевого вектора, но и противоположных к его элементам; нельзя остаться в пределах многообразия, складывая пары его элементов. Однако если взять три элемента m_1, m_2, m_3, то окажется, что $m_1 + m_2 - m_3 \in M$.

6.3.2. Чтобы проиллюстрировать понятие линейного многообразия, приведем еще одно их свойство. Нам известно, что пересечение плоскостей, если оно непусто, – это либо плоскость, либо прямая; непустое пересечение прямых – это прямая либо точка. Поскольку точка – тоже линейное многообразие: сумма оканчивающегося в ней вектора с нуль-подпространством, с геометрической точки зрения мы можем быть уверены, что непустое пересечение многообразий окажется снова линейным многообразием. Единственное отличие от подпространств здесь заключается в оговорке «непустое» – ведь пересечение подпространств всегда непусто, так как обязательно содержит нулевой элемент.

Рассмотрим вопрос о пересечении многообразий P и M в общем случае. Пусть многообразия $P = p + U$ и $M = m + W$ имеют непустое пересечение, в котором содержится элемент v. Тогда он представляется, с одной стороны, в виде $v = p + u$, с другой – в виде $v = m + w$.

Из равенства $p + u = m + w$ следует, что

$$p - m = -u + w \in U + W.$$

Наоборот, если «элементы сдвигов» двух многообразий P и M удовлетворяют свойству $p - m \in U + W$, т. е. $p - m = -u + w$ для некоторых $u \in U$ и $w \in W$, то равные элементы $p + u = m + w$ попадают одновременно и в многообразие $P = p + U$, и в $M = m + W$, т. е. пересечение их непусто.

Таким образом, ***включение*** $p - m \in U + W$ ***оказывается критерием непустоты пересечения многообразий***.

Остается понять, какой элемент сдвига и какое подпространство ему соответствуют. Элемент $p + u = m + w$, не лежащий ни в подпространстве U, ни в W, но содержащийся в пересечении многообразий, очевидно обеспечивает требуемый сдвиг для подмножества $P \cap M$. А если пара элементов a и b содержится в каждом из многообразий, то их разность $a - b$ попадает в каждое из подпространств U и W, а тем самым – в их пересечение $U \cap W$.

Окончательный ответ таков:

$$P \cap M = p + u + U \cap W = m + w + U \cap W,$$

где элемент $p + u$ определяется из критерия непустоты:

$$p - m = -u + w \in U + W.$$

6.4. СТРУКТУРА ОБЩЕГО РЕШЕНИЯ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Геометрические гипотезы у нас готовы: если пересечение нескольких плоскостей непусто (т. е. система уравнений совместна), то получится прямая либо плоскость (для двух плоскостей) и как третья возможность – точка (для трех и более). Как мы поняли в п. 6.3.3, все это линейные многообразия, даже точка – это вектор сдвига, приводящий в нее, плюс нулевое подпространство.

Для однородной системы, где в правых частях стоят нули, значения $x = y = z = 0$ обязательно будут решением. Значит, в этом случае плоскость и прямая пройдут через ноль, а единственная точка и сама будет нулем, т. е. решение будет линейным подпространством.

Чтобы убедиться в этом для произвольной системы уравнений, не сводящихся к плоскостям в геометрическом пространстве, будем пользоваться матричной формой записи системы (см. п. 6.2.2).

6.4.1. Решение однородной системы. ***Однородная система уравнений*** соответствует матричному равенству

$$AX = (0),$$

где в правой части стоит одностолбцовая нулевая матрица. Как и в п. 3.3.17, нам достаточно проверить замкнутость множества решений такой системы относительно операций сложения и умножения на число, чтобы убедиться, что это линейное пространство.

Но это прямо следует из свойств матричных операций:

- если $AX_1 = (0)$ и $AX_2 = (0)$, то

$$A(X_1 + X_2) = AX_1 + AX_2 = (0);$$

- если $AX_1 = (0)$, то для любого числа α

$$A(\alpha X_1) = \alpha AX_1 = \alpha(0) = (0).$$

Вот и все. Линейное подпространство, определяемое однородной системой уравнений $AX = (0)$, называется общим решением однородной системы и обозначается $X_{\text{о.о}}$.

6.4.2. Решение неоднородной системы. Обозначим $X_{\text{о.н}}$ все множество решений совместной неоднородной системы $AX = B$, т. е. *общее* решение неоднородной системы, а $X_{\text{ч.н}}$ – некоторое конкретное, частное решение неоднородной системы. Тогда представление линейного многообразия $X_{\text{о.н}}$ с помощью подпространства $X_{\text{о.о}}$ будет выглядеть так:

$$X_{\text{о.н}} = X_{\text{ч.н}} + X_{\text{о.о}}.$$

Проверим совпадение этих множеств. Сначала проверим, что сумма $X_{\text{ч.н}} + X_1$ для всякого $X_1 \in X_{\text{о.о}}$ попадает во множество $X_{\text{о.н}}$, а для этого подставим сумму $X_{\text{ч.н}} + X_1$ в систему:

$$A(X_{\text{ч.н}} + X_1) = AX_{\text{ч.н}} + AX_1 = B + (0) = B.$$

Этим доказано включение $X_{\text{ч.н}} + X_{\text{о.о}} \subseteq X_{\text{о.н}}$. Проверим обратное включение, т. е. убедимся, что для всякого $X^* \in X_{\text{о.н}}$ найдется $X_1 \in X_{\text{о.о}}$ такой, что $X^* = X_{\text{ч.н}} + X_1$. Или, что то же самое, проверим просто то, что

$$X_1 = X^* - X_{\text{ч.н}} \in X_{\text{о.о}}.$$

Но $AX_1 = A(X^* - X_{\text{ч.н}}) = AX^* - AX_{\text{ч.н}} = B - B = (0)$.

Этим доказано, что $X^* \in X_{\text{ч.н}} + X_{\text{о.о}}$, что $X_{\text{о.н}} \subseteq X_{\text{ч.н}} + X_{\text{о.о}}$ и нужное нам равенство подтверждено.

Остается упомянуть один термин. В подпространстве $X_{\text{о.о}}$ можно ввести базис, и по традиции он называется ***фундаментальной системой решений***, сокращенно – ф. с. р.

ГЛАВА 7

МЕТОД ГАУССА

7.1. ЭЛЕМЕНТАРНЫЕ ПРЕОБРАЗОВАНИЯ СИСТЕМЫ УРАВНЕНИЙ

Целью этого раздела является построение решения произвольной системы линейных уравнений или своевременное обнаружение ее несовместности. Сама система при нахождении решения изменяется – но ее решение сохраняется. Собственно говоря, процесс нахождения состоит в получении цепочки систем с одним и тем же множеством общего решения (такие системы, как и их матрицы, называются ***эквивалентными***); первое звено цепочки – это данная система, а последнее – система, позволяющая выписать ответ уже без вычислений.

7.1.1. Пример. Решим систему, исключая последовательно неизвестные x и y:

$$\begin{cases} 2x - y + 3z = -3, \\ -x + 2y + z = 2, \\ 3x - 3y - z = -2, \end{cases} \Rightarrow \begin{cases} x = -1,5 + 0,5y - 1,5z, \\ (1,5 - 0,5y + 1,5z) + 2y + z = 2, \\ 3(-1,5 + 0,5y - 1,5z) - 3y - z = -2, \end{cases} \Rightarrow$$

$$\begin{cases} x = -1,5 + 0,5y - 1,5z, \\ 1,5y + 2,5z = 0,5, \\ -1,5y - 5,5z = 2,5, \end{cases} \Rightarrow \begin{cases} x = -1,5 + 0,5y - 1,5z, \\ 1,5y = 0,5 - 2,5z, \\ (-0,5 + 2,5z) - 5,5z = 2,5, \end{cases} \Rightarrow$$

$$\begin{cases} x = -1,5 + 0,5y - 1,5z, \\ y = \dfrac{2}{3}(0,5 - 2,5z), \\ -3z = 3, \end{cases} \Rightarrow \begin{cases} x = -1,5 + 0,5 \cdot 2 + 1,5 = 1, \\ y = \dfrac{2}{3}(0,5 + 2,5) = 2, \\ z = -1. \end{cases}$$

Ясно, что система из трех уравнений

$$\begin{cases} x = 1, \\ y = 2, \\ z = -1 \end{cases}$$

имеет такое же решение, как и исходная: ведь, выражая неизвестное x на первом шаге, мы получили равенство с другой расстановкой слагаемых, но и только; оно удовлетворяется при одних и тех же значениях x, y и z, что и исходное. Опять же, подставляя x, выраженный через y и z, в два других уравнения, мы меняем величину x на равную ей сумму – при условии, что первое уравнение удовлетворено. То же и с переменной y. Все преобразования, сделанные нами, не нарушают равенств, зато для нас они были целесообразны: от трех уравнений с участием всех трех неизвестных мы перешли к двум уравнениям с двумя неизвестными, а потом и к одному – с одним. Последнее равенство $z = -1$, на наш взгляд, уже не уравнение, а ответ для неизвестного z; с его помощью получено еще два ответа для неизвестных x и y.

Сделаем еще два наблюдения. Во-первых, если вместо того, чтобы выражать x из первого уравнения и подставлять его во второе и третье, мы вычтем соответственные части уравнений (из левой – левую, из правой – правую), поделив обе части первого уравнения на такое число (–2), чтобы неизвестное x исчезло:

$$\begin{array}{r} -\begin{array}{r} -x + 2y + z = 2 \\ 2x - y + 3z = -3 \; |:(-2) \end{array} \\ \hline 0x + (2 - 0,5)y + (1 + 1,5)z = (2 - 1,5) \end{array}$$

$$\begin{array}{r} -\begin{array}{r} 3x - 3y - z = -2 \\ 2x - y + 3z = -3 \; |\cdot 1,5 \end{array} \\ \hline 0x - 1,5y - 5,5z = 2,5 \end{array},$$

то в результате получится

$$\begin{cases} 1,5y + 2,5z = 0,5, \\ -1,5y - 5,5z = 2,5. \end{cases}$$

Тем самым вычитание уравнений приводит в точности к тому же самому, что и подстановка (т. е. исключение) одного из неизвестных.

Во-вторых, если мы составим расширенную матрицу системы (см. п. 6.2.2)

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & -3 \\ -1 & 2 & 1 & 2 \\ 3 & -3 & -1 & -2 \end{array}\right)$$

и вместо вычитания уравнений будем проводить действия с ее строками (ведь все равно действия с уравнениями превращаются в действия с коэффициентами при одинаковых неизвестных: коэффициенты при неизвестных y и z складываются с коэффициентами при них же, соответственно и правые части), то результат опять будет прежний:

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & -3 \\ -1 & 2 & 1 & 2 \\ 3 & -3 & -1 & -2 \end{array}\right) \begin{array}{l} \\ +0,5\ \mathrm{I} \Rightarrow \\ -1,5\ \mathrm{I} \end{array} \left(\begin{array}{ccc|c} 2 & -1 & 3 & -3 \\ 0 & 1,5 & 2,5 & 0,5 \\ 0 & -1,5 & -5,5 & 2,5 \end{array}\right)$$

(римская цифра I в записи «+0,5 I» означает половину 1-й строки; впредь римские цифры будут использоваться в качестве номеров строк, чтобы не путать их с арабскими цифрами коэффициентов).

7.1.2. Глядя на пример из п. 7.1.1, сделаем следующие выводы:

- все действия ради более компактных записей можно проделывать с расширенной матрицей системы;
- идея решения состоит в последовательном исключении неизвестных, т. е. в приведении матрицы системы к такому виду, где в левой части строки стоит тем больше нулей, чем ниже стоит сама строка;
- способом получения этих нулей может быть вычитание из одних строк, расположенных ниже, других, расположенных выше, с подходящими множителями.

Далее мы ознакомимся с точными формулировками.

7.1.3. Элементарные преобразования расширенной матрицы системы. Так как везде в дальнейшем неизвестными будет служить одна и та же буква с разными номерами – с 1-го до n-го, нам легко не только записать расширенную матрицу по системе уравнений, но и восстановить систему по ее матрице: достаточно подписать к элементам самой матрицы системы эти нумерованные буквы, добавить недостающие «плюсы» («минусы» у коэффициентов уже есть) и приравнять к элементам правого столбца расширенной матрицы:

$$\left(\begin{array}{cc|c} 1 & -2 & 1 \\ -3 & 2 & 0 \end{array}\right) \to \begin{cases} x_1 - 2x_2 = 1 \\ -3x_1 + 2x_2 = 0 \end{cases} \quad (\text{здесь } n = 2).$$

Поэтому везде далее мы будем рассматривать только расширенную матрицу системы вместо самой системы.

Следующие четыре типа преобразований матрицы будем называть ***элементарными***.

1. Перестановка строк расширенной матрицы – это соответствует перестановке уравнений и очевидно не сказывается на решении системы.

2. Умножение или деление строки расширенной матрицы на ненулевое число – это соответствует умножению обеих частей уравнения на одно и то же число, что опять-таки не влияет на решение.

3. Прибавление к элементам одной (k-й) строки расширенной матрицы соответственных элементов другой (i-й) строки, умноженной на некоторое число α.

Это значит, что и к левой, и к правой части одного (k-го) уравнения прибавляются левая и правая части другого (i-го), умноженные на одно и то же число α:

$$\left(\begin{array}{ccc|c} a_{i1} & \dots & a_{in} & b_i \\ a_{k1} & \dots & a_{kn} & b_k \end{array}\right) \begin{array}{c} \\ +\alpha < i > \end{array} \Rightarrow \left(\begin{array}{ccc|c} a_{i1} & \dots & a_{in} & b_i \\ a_{k1} + \alpha a_{i1} & \dots & a_{kn} + \alpha a_{in} & b_k + \alpha b_i \end{array}\right).$$

Почему такое преобразование не меняет решения? Если i-е и k-е уравнения оба удовлетворены какими-то значениями неизвестных, то и новое, k'-е, уравнение тоже будет удовлетворено: ведь к обеим частям выполнявшегося k-го уравнения прибавлены равные величины, полученные из i-го уравнения.

Если же, наоборот, i-е и k'-е уравнения удовлетворены, то, вычитая из обеих частей k'-го уравнения обе части i-го, умноженные на α, мы вернемся к исходному k-му уравнению, которое тоже должно удовлетворяться:

$$\left(\begin{array}{ccc|c} a_{i1} & \ldots & a_{in} & b_i \\ a_{k1}+\alpha a_{i1} & \ldots & a_{kn}+\alpha a_{in} & b_k+\alpha b_i \end{array}\right) -\alpha < i > \Rightarrow \left(\begin{array}{ccc|c} a_{i1} & \ldots & a_{in} & b_i \\ a_{k1} & \ldots & a_{kn} & b_k \end{array}\right)$$

Поэтому пара уравнений (i, k) равносильна паре (i, k') и преобразование этого типа тоже сохраняет решение.

4. Перестановка столбцов матрицы системы (нерасширенной) с запоминанием их исходных номеров – это соответствует временной перенумерации неизвестных, но благодаря тому, что мы помним первоначальные номера, решение не пострадает: нам известно, какое значение неизвестного к какому коэффициенту подставлять.

Сделаем два важных замечания: во-первых, множества решений систем, соответствующих преобразуемым матрицам, совпадают, поэтому эти системы эквивалентны, и эквивалентными называются также матрицы и сами преобразования. Во-вторых, преобразования всех четырех типов обратимы, т. е. позволяют вернуться от преобразованной матрицы к исходной аналогичными преобразованиями.

7.2. ПРЯМОЙ И ОБРАТНЫЙ ХОДЫ МЕТОДА ГАУССА

7.2.1. Для начала определим специальные виды матриц, которых потребует выполнение шагов ***алгоритма Гаусса***. Именно матрица называется ***верхнетреугольной до k-го столбца***, если в ней $a_{ij} = 0$ при всех $i > j < k$, но $a_{jj} \neq 0$ для всех $j < k$.

Это означает, что в столбцах, стоящих левее k-го ($j < k$), ниже главной диагонали стоят нули: $a_{ij} = 0$ при $i > j$, а на главной диагонали стоят отличные от нуля элементы: $a_{jj} \neq 0$. Чтó стоит выше главной диагонали – безразлично.

Например, верхнетреугольными являются такие матрицы:

$\left(\begin{array}{c|ccc}1&2&3&0\\0&0&0&0\end{array}\right)$ – до 2-го столбца,

$\left(\begin{array}{cc|cc}1&2&3&0\\0&2&0&0\end{array}\right)$ – до 3-го столбца,

$\left(\begin{array}{c|cc}1&2&1\\0&1&2\\0&3&1\end{array}\right)$ – до 2-го столбца,

$\left(\begin{array}{ccc|c}1&2&1&\\0&1&2&\\0&0&2&\end{array}\right)$ – до 4-го столбца (которого на самом деле нет);

левее черты стоят столбцы, в которых выполняются условия верхнетреугольности; а их число определяет «показатель верхнетреугольности», т. е. число k, фигурирующее в определении).

7.2.2. Окончательной же целью прямого хода алгоритма Гаусса является матрица, ***трапециевидная с заполнением*** k. Это матрица, во-первых, верхнетреугольная до $(k + 1)$-го столбца, у которой, во-вторых, строки $(k + 1)$-я и ниже заполнены нулями; на языке формул это выглядит так:

$a_{i\,j} = 0$ при всех $i > j$ или $i > k$,

$a_{j\,j} \neq 0$ при $j < k + 1$.

Например, матрицы

$\begin{pmatrix}1&2&3&0\\0&0&0&0\end{pmatrix}$ – трапециевидная с заполнением 1,

$\begin{pmatrix}1&2&3&0\\0&1&0&0\end{pmatrix}$ – трапециевидная с заполнением 2,

$\begin{pmatrix}1&2&1\\0&1&2\\0&1&2\end{pmatrix}$ – нетрапециевидная ($a_{32} \neq 0$),

$\begin{pmatrix}1&2&1\\0&1&2\\0&0&2\end{pmatrix}$ – трапециевидная с заполнением 3,

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$ – нетрапециевидная ($a_{22} = 0$, но $a_{23} \neq 0$).

7.2.3. Шаг прямого хода алгоритма Гаусса. Предположим, что нам дана матрица, верхнетреугольная до k-го столбца:

$$\left(\begin{array}{cccc|ccc} a_{11} & & & & a_{1k} & \dots & a_{1n} \\ 0 & a_{22} & & & & & \\ 0 & 0 & \ddots & & & & \\ \vdots & & & & & & \\ - & - & - & & - & - & - \\ & & & & a_{kk} & \dots & a_{kn} \\ & 0 & & & \vdots & & \vdots \\ & & & & a_{mk} & \dots & a_{mn} \end{array}\right).$$

С помощью элементарных преобразований ее надо сделать верхнетреугольной до (k + 1)-го столбца.

Сосредоточим внимание на той части матрицы, в которой элемент a_{kk} является левым верхним, т. е. на прямоугольнике «сокращенной матрицы»:

$$\begin{pmatrix} a_{kk} & \cdots & a_{kn} \\ \vdots & & \vdots \\ a_{mk} & \cdots & a_{mn} \end{pmatrix}.$$

Если здесь нам удастся добиться равенства

$$a_{k+1\,k} = a_{k+2\,k} = \ldots = a_{mk} = 0 \text{ при } a_{kk} \neq 0,$$

то, повторяя те же элементарные преобразования в полной матрице, мы добьемся повышения показателя верхнетреугольности k.

Если в «сокращенной матрице» $a_{k\,k} = 0$, но есть какой-то другой элемент $a_{i\,j} \neq 0$, то, переставляя k-ю и i-ю строки, затем k-й и j-й столбцы, мы добъемся того, что ненулевой элемент выйдет на место (k k).

Если в сокращенной матрице все элементы нулевые, то можно переходить к следующему пункту.

Если же $a_{k\,k} \neq 0$, можно вычесть k-ю строку поочередно из всех лежащих ниже, умножив ее соответственно на коэффициенты $a_{l\,k} / a_{k\,k}$, $l = k + 1, \ldots, m$. Тогда во вновь получившейся матрице элемент, стоящий в k-м столбце на любом месте, будет равен нулю:

$$a_{l\,k} - \frac{a_{l\,k}}{a_{k\,k}} a_{k\,k} = a_{l\,k} - a_{l\,k} = 0.$$

$$\begin{pmatrix} a_{1\,1} & & a_{1\,k} & \cdots & a_{1\,n} \\ 0 & \ddots & & & \\ \vdots & a_{k-1\,k-1} & a_{k-1\,k} & & a_{n\,k} \\ 0 & 0 & a_{k\,k} & a_{k\,k+1} & a_{k\,n} \\ \vdots & & 0 & \vdots & \vdots \\ 0 & 0 & 0 & a_{m\,k+1} & a_{m\,n} \end{pmatrix}.$$

Тем самым матрица стала верхнетреугольной до (k + 1)-го столбца.

Пример. Дана матрица, верхнетреугольная до 2-го столбца:

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 1 \end{pmatrix}.$$

Надо сделать ее верхнетреугольной до 3-го столбца. Поскольку $a_{22} = 0$, переставим 2-й и 4-й столбцы, запоминая их номера (точнее, записывая над ними их прежние номера):

$$\begin{array}{c} \begin{matrix} & \mathrm{IV} & & \mathrm{II} \end{matrix} \\ \begin{pmatrix} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 0 \end{pmatrix} \end{array}.$$

Так как теперь $a_{22} \neq 0$, то можно вычесть из третьей строки вторую, домножив ее на коэффициент $\dfrac{a_{32}}{a_{22}} = \dfrac{1}{2}$:

$$\begin{array}{cc} & \begin{matrix} \quad & \mathrm{IV} & & \mathrm{II} \end{matrix} \\ & \begin{pmatrix} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 0 \end{pmatrix} \begin{matrix} \\ \\ -1/2\ \mathrm{II} \end{matrix} \end{array} \Rightarrow \begin{array}{c} \begin{matrix} \quad & \mathrm{IV} & & \mathrm{II} \end{matrix} \\ \begin{pmatrix} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} \end{array}.$$

Получилась матрица, верхнетреугольная до 3-го столбца (а на самом деле, и до 4-го столбца, т. е. это трапециевидная матрица с заполнением 3).

7.2.4. Начало и окончание прямого гауссова хода. Поскольку любую матрицу (без нулевых элементов или с беспорядочно расположенными нулями) можно считать верхнетреугольной до 1-го столбца, первый шаг гауссова хода будет состоять в том, чтобы добиться неравенства $a_{11} \neq 0$ и равенства $a_{21} = a_{31} = \ldots = a_{m1} = 0$; второй шаг приведет к ситуации $a_{22} \neq 0,\ a_{32} = \ldots = a_{m2} = 0$ и так далее (это описано в предыдущем пункте). За счет чего этот процесс может остановиться?

Во-первых, закончатся строки: матрица станет верхнетреугольной до $(k + 1)$-го столбца, но $(k + 1)$-й строки просто нет: $k = m < n$:

$$\begin{pmatrix} a_{11} & \cdots & & a_{1\,k-1} & a_{1k} & \cdots & a_{1n} \\ 0 & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{k-1\,k-1} & a_{k-1\,k} & \cdots & a_{k-1\,n} \\ 0 & \cdots & 0 & 0 & a_{kk} & \cdots & a_{kn} \end{pmatrix}.$$

Но это будет означать, что матрица трапециевидна с заполнением $k = m$, следовательно, цель прямого хода достигнута.

Во-вторых, кончатся столбцы: матрица станет верхнетреугольной до $(k + 1)$-го столбца при $k = n < m$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1\,n-1} & a_{1n} \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & \cdots & 0 & a_{nn} \\ \vdots & & & \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Но и эта матрица трапециевидна с заполнением $k = n$, так что ход закончен.

В-третьих, в блоке (так принято называть прямоугольные фрагменты матриц)

$$\begin{pmatrix} a_{kk} & \cdots & a_{kn} \\ \vdots & & \vdots \\ a_{mk} & \cdots & a_{mn} \end{pmatrix}$$

может не быть элементов, отличных от нуля (см. п. 7.2.3), – но это будет означать, что везде ниже главной диагонали стоят нули и что нижние строки, начиная с k-й, целиком нулевые:

$$\left(\begin{array}{cccc|ccc} a_{11} & a_{12} & \cdots & a_{1\,k-1} & a_{1k} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2\,k-1} & & \cdots & \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & & & a_{k-1\,k-1} & a_{k-1\,k} & \cdots & a_{k-1\,n} \\ - & - & - & - & - & - & - \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{array}\right).$$

Но это трапециевидная матрица с заполнением $k - 1$. Цель достигнута.

7.2.5. После того как матрица стала трапециевидной с заполнением k, начинается новый этап. Во-первых, можно точно сказать, совместна ли система. Здесь необходимо подчеркнуть, что элементарными преобразованиями расширенной матрицы системы к трапециевидной форме приводится *нерасширенная* матрица системы. И если в столбце свободных членов есть ненулевые элементы ниже k-го места, система несовместна.

В самом деле, по матрице

$$\left(\begin{array}{ccc|ccc|c} a_{11} & \cdots & a_{1k} & a_{1\,k+1} & \cdots & a_{1\,n} & b_1 \\ & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & & a_{kk} & a_{k\,k+1} & \cdots & a_{kn} & b_k \\ - & - & - & - & - & - & - \\ & & & & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 & b_j \end{array}\right) \begin{array}{l} \\ \\ \\ \\ b_j \neq 0 \\ (j > k) \end{array}$$

восстанавливается система:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1k}x_k + a_{1\,k+1}x_{k+1} + \ldots + a_{1n}x_n = b_1, \\ \qquad\qquad \vdots \\ 0x_1 + \ldots + a_{kk}x_k + a_{k\,k+1}x_{k+1} + \ldots + a_{kn}x_n = b_k, \\ 0x_1 + \ldots + 0x_k + 0x_{k+1} + \ldots + 0x_n = b_j. \end{cases}$$

Здесь нижнее уравнение не может быть удовлетворено ни при каких значениях неизвестных $x_1, \ldots, x_n$, т. е. решений нет.

Надо заметить, что мы продолжаем нумеровать неизвестные слева направо в возрастающей последовательности, хотя из-за перестановок столбцов эта исходная нумерация могла нарушиться, однако благодаря тому, что исходные номера надписаны над переставлявшимися столбцами, точно известно, какое по номеру неизвестное приписывать к коэффициентам, стоящим в любом столбце матрицы.

Если же $b_{k+1} = \ldots = b_m = 0$, то система совместна. В самом деле, мы можем придать любые значения правым $n - k$ неизвестным:

$$x_{k+1} = x_{k+1}^0, \ldots, x_n = x_n^0$$

и перенести их в правую часть:

$$a_{11}x_1 + \ldots + a_{1k-1}x_{k-1} \quad + a_{1k}x_k = b_1 - a_{1k+1}x_{k+1}^0 - \ldots - a_{1n}x_n^0,$$

$$\vdots$$

$$a_{k-1\,k-1}x_{k-1} + a_{k-1\,k}x_k = b_{k-1} - a_{k-1\,k}x_{k+1}^0 - \ldots - a_{k-1\,n}x_n^0,$$

$$a_{kk}x_k = b_k - a_{kk+1}x_{k+1}^0 - \ldots - a_{kn}x_n^0.$$

Обозначив правые части $c_1, \ldots, c_k$, получим возможность поочередно выражать остальные неизвестные из уравнений снизу вверх:

$$x_k = c_k / a_{kk}.$$

$$x_{k-1} = \left(c_{k-1} - a_{k-1\,k} \frac{c_k}{a_{kk}} \right) \Big/ (a_{k-1\,k-1})$$

и т. д. Благодаря тому что все диагональные коэффициенты $a_{ii} \neq 0$ отличны от нуля, никаких препятствий не встретится; некоторое решение, зависящее от того, какие значения приданы неизвестным $x_{k+1}, \ldots, x_n$, будет получено. Поэтому выражаемые неизвестные $x_1, \ldots, x_k$ называются ***связанными*** или ***базисными***, а те, которым придаются произвольные значения, ***свободными***, – через них выражаются связанные неизвестные (опять-таки надо сделать оговорку насчет того, что при переставлявшихся столбцах свободными оказываются *правые n – k* номеров, а связанными – *левые k* из них).

7.2.6. Обратный ход метода Гаусса. Как видно из формул для неизвестных x_k, $x_{k-1}, \ldots, x_1$, пункт 7.2.5, уже из системы с трапециевидной матрицей можно было бы получить выражения связанных неизвестных через свободные, которые естественно было бы считать ответом. Но и это можно упростить благодаря матричной форме записи. Если в трапециевидной матрице с заполнением *k* мы начнем вычитать из 1-й, 2-й и далее до (*k* – 1)-й строки – *k*-ю строку, умножаемую соответственно на коэффициенты

$$\frac{a_{1k}}{a_{kk}},\ \frac{a_{2k}}{a_{kk}},\ \ldots,\ \frac{a_{k-1\,k}}{a_{kk}},$$

то у нас обратятся в нуль элементы, стоящие в k-м столбце выше главной диагонали:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & | & a_{1\,k+1} & \cdots & | & b_1 \\ 0 & a_{22} & \cdots & a_{2k} & | & a_{2\,k+1} & \cdots & | & b_2 \\ \vdots & \vdots & & \vdots & | & \vdots & & | & \vdots \\ 0 & 0 & \cdots & a_{kk} & | & a_{k\,k+1} & \cdots & | & b_k \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1\,k+1} & 0 & | & a'_{1\,k+1} & \cdots & | & b'_1 \\ 0 & a_{22} & \cdots & a_{2\,k-1} & 0 & | & a'_{2\,k+1} & \cdots & | & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & | & \vdots & & | & \vdots \\ 0 & 0 & & a_{k-1\,k-1} & 0 & | & a'_{k-1\,k+1} & \cdots & | & b'_{k-1} \\ 0 & 0 & & 0 & a_{kk} & | & a_{k\,k+1} & \cdots & | & b_k \end{pmatrix}.$$

То же самое можно проделать и с $(k-1)$-м столбцом, и так далее вплоть до момента, когда левый квадратный блок матрицы

$$\begin{pmatrix} a_{11} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & a_{kk} \end{pmatrix}$$

станет просто диагональным.

Последнее, что останется сделать, – поделить строки матрицы на числа $a_{11}, a_{22}, \ldots, a_{kk}$ соответственно, чтобы левые k столбцов образовали просто единичную матрицу (оставшиеся элементы мы переобозначим):

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & d_{1\,r+1} & \ldots & d_{1\,n} & | & l_1 \\ 0 & 1 & & \vdots & d_{2\,r+1} & & d_{2\,n} & | & l_2 \\ \vdots & & \ddots & 0 & \vdots & & & | & \vdots \\ 0 & \cdots & 0 & 1 & d_{r\,r+1} & & d_{r\,n} & | & l_r \end{pmatrix}.$$

Получившаяся матрица по-прежнему эквивалентна исходной, поскольку ничем, кроме элементарных преобразований, мы не пользовались; соответствующая ей система имеет то же решение, что и данная. Но теперь записать связанные неизвестные через свободные совсем легко:

$$\begin{gathered} x_1 = l_1 - d_{1\,k+1}x_{k+1} - \ldots - d_{1n}x_n, \\ \vdots \\ x_k = l_k - d_{k\,k+1}x_{k+1} - \ldots - d_{kn}x_n. \end{gathered}$$

Иногда эту систему равенств называют ***общим решением системы в параметрической форме***.

7.2.7. Если же мы хотим построить решение в векторной форме, включающее все неизвестные (как оно определялось в п. 6.2.1), то получится:

$$X_{о.н.} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} l_1 - d_{1\,r+1} \cdot x_{r+1} - \ldots - d_{1\,n} \cdot x_n \\ \vdots \\ l_r - d_{r\,r+1} \cdot x_{r+1} - \ldots - d_{r\,n} \cdot x_n \\ x_{r+1} \\ \ddots \\ x_n \end{pmatrix} =$$

$$= \begin{pmatrix} l_1 \\ \vdots \\ l_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} -d_{1\,r+1} \\ \vdots \\ -d_{r\,r+1} \\ 1 \\ \vdots \\ 0 \end{pmatrix} \cdot x_{r+1} + \ldots + \begin{pmatrix} -d_{1\,n} \\ \vdots \\ -d_{r\,n} \\ 0 \\ \vdots \\ 1 \end{pmatrix} \cdot x_n .$$

Это более ясная форма решения, по сравнению с той общей, которую мы получили в п. 6.4.2. Поскольку неизвестные $x_{k+1}, \ldots, x_n$ могут принимать любые значения, в том числе нулевые, вектор

$$X_{\text{ч.н.}} = \begin{pmatrix} l_1 \\ \vdots \\ l_k \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

сам по себе является частным решением неоднородной системы.

Остальные слагаемые – линейная оболочка векторов фундаментальной системы решений (см. п. 6.4.2):

$$\vec{d}_1 = \begin{pmatrix} -d_{1\,k+1} \\ \vdots \\ -d_{k\,k+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \; \vec{d}_{n-k} = \begin{pmatrix} -d_{1n} \\ \vdots \\ -d_{kn} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

(это линейная оболочка, а не линейная комбинация, именно потому, что $x_{k+1}, \ldots, x_n$ – переменные, которые принимают всевозможные числовые значения).

Такое линейное подпространство и является общим решением однородной системы с той же матрицей A:

$$X_{\text{о.о}} = \left\{ \vec{d}_1\, x_{k+1} + \ldots + \vec{d}_{n-k}\, x_n \mid x_{k+1}, \ldots, x_n \in \mathbf{R} \right\}.$$

Линейную независимость векторов $\vec{d}_1, \ldots, \vec{d}_{n-k}$ легко видеть из того, что нижние $n - k$ координат их линейной комбинации равны коэффициентам линейной комбинации и равны числовым значениям переменных $x_{k+1}, \ldots, x_n$; поэтому если линейная комбинация этих векторов обратилась в нуль, то и все коэффициенты равны нулю.

7.2.8. Пример решения системы.

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = -1, \\ 2x_1 + 4x_2 - 4x_3 + 3x_4 = -3, \\ 3x_1 + 6x_2 - x_3 + 2x_4 = -2, \\ x_1 + 2x_2 + x_3 = 0. \end{cases}$$

Запишем матрицу и элементарные преобразования первого шага прямого гауссова хода:

$$\begin{array}{cc} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 1 & -1 \\ 2 & 4 & -4 & 3 & -3 \\ 3 & 6 & -1 & 2 & -2 \\ 1 & 2 & 1 & 0 & 0 \end{array}\right) & \begin{array}{c} \\ -2\mathrm{I} \\ -3\mathrm{I} \\ -\mathrm{I} \end{array} \end{array} \Rightarrow \left(\begin{array}{cccc|c} 1 & 2 & -1 & 1 & -1 \\ 0 & 0 & -2 & 1 & -1 \\ 0 & 0 & 2 & -1 & 1 \\ 0 & 0 & 2 & -1 & 1 \end{array}\right).$$

Поскольку нижние три строки очевидно дублируют друг друга, две из них можно сделать сплошь нулевыми и просто устранить из матрицы (содержащаяся в них информация о неизвестных одинакова, и они представляют собой просто «балласт»). Кроме того, переставим второй и четвертый столбцы:

$$\begin{array}{c} \begin{array}{cccccc} & \mathrm{IV} & & \mathrm{II} & & \end{array} \\ \left(\begin{array}{cccc|c} 1 & 1 & -1 & 2 & -1 \\ 0 & 1 & -2 & 0 & -1 \end{array}\right) \end{array}.$$

Это трапециевидная матрица с заполнением 2, и свободных (x_2 и x_3) и связанных (x_1 и x_4) неизвестных будет по два. Поскольку система совместна ($b_3 = b_4 = 0$), можно отбросить нижние строки и сделать шаг обратного гауссова хода, вычитая из первой строки вторую:

$$\begin{array}{c} \begin{array}{cccccc} & \mathrm{IV} & & \mathrm{II} & & \end{array} \\ \left(\begin{array}{cccc|c} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & -2 & 0 & -1 \end{array}\right) \end{array}.$$

Отсюда выражаются связанные неизвестные:

$$x_1 = -2x_2 - x_3,$$
$$x_4 = -1 + 2x_3.$$

В векторной форме решение примет вид:

$$X_{\text{о.н}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2x_2 - x_3 \\ x_2 \\ x_3 \\ -1 + 2x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} x_2 + \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix} x_3,$$

$$X_{\text{ч.н}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}; \text{ ф.с.р.: } \vec{d}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ \vec{d}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

7.3. НЕКОТОРЫЕ ПРИМЕНЕНИЯ МЕТОДА ГАУССА

7.3.1. Теорема о размерности. Число базисных векторов линейного пространства $L_{\mathbf{F}}$ в п. 3.3.9 мы называли его размерностью. Корректность этого определения неочевидна: не может ли пространство $L_{\mathbf{F}}$ иметь два базиса с ***разным*** числом элементов: $v_1, \ldots, v_m$ и $u_1, \ldots, u_n$?

Лемма. Пусть $v_1, \ldots, v_m$ – базис пространства $L_{\mathbf{F}}$ и $u_1, \ldots, u_n$ – некоторый набор векторов, причем $n > m$. Тогда векторы $u_1, \ldots, u_n$ линейно зависимы.

Доказательство. Разложим элементы $u_1, \ldots, u_n$ по базису $v_1, \ldots, v_m$:

$$u_1 = a_{11}v_1 + \ldots + a_{m\,1}v_m;$$

$$\vdots$$

$$u_n = a_{1n}v_1 + \ldots + a_{m\,n}v_m;$$

тогда столбцы координат этих элементов окажутся такими:

$$\vec{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m\,1} \end{pmatrix}, \ldots, \ \vec{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{m\,n} \end{pmatrix}.$$

Приравняем к нулю их линейную комбинацию: $x_1u_1 + ... + x_nu_n = \underline{0}$, или, что то же самое,

$$x_1 \cdot \begin{pmatrix} a_{11} \\ \vdots \\ a_{m\,1} \end{pmatrix} + \ldots + x_n \cdot \begin{pmatrix} a_{1n} \\ \vdots \\ a_{m\,n} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow A_{m \times n} X = (0).$$

Матрица коэффициентов последней однородной системы прямоугольна, и приведенная к трапециевидной форме, она будет иметь заполнение не выше числа строк m. Тогда $n - m > 0$ неизвестных (ведь $n > m$) будут свободными и смогут принимать любые, в том числе и ненулевые значения. Значит нулевая линейная комбинация векторов $u_1, \ldots, u_n$ может иметь ненулевые коэффициенты:

$$x_1u_1 + ... + x_nu_n = \underline{0}, \quad x_k \neq 0.$$

Следовательно, векторы линейно зависимы.

Следствие: **теорема о размерности**. Все базисы конечномерного линейного пространства содержат одинаковое число элементов – однозначно определенную *размерность пространства.*

7.3.2. Похожие вопросы можно задавать и в связи с любой матрицей: каковы наибольшие (нерасширимые) линейно независимые наборы ее строк и столбцов, одинаково ли число элементов в тех и других?

Поскольку конечные множества строк и столбцов не образуют линейного пространства над полем **F**, то вместо слова *размерность* используется слово *ранг*. А именно, ***строчным (столбцовым) рангом*** матрицы row-rank (A) (соответственно col-rank (A)) называется число строк в максимальном линейно независимом (или *ранговом*) наборе ее строк (столбцов). Однозначно ли определены эти числа и связаны ли они друг с другом?

Теорема о ранге. Строчный и столбцовый ранги матрицы равны между собой и совпадают с заполнением ее трапециевидной формы.

Доказательство состоит из двух частей: обоснование равенства столбцового ранга заполнению трапециевидной формы матрицы и равенства ему же строчного ранга.

(1) Предварительно установим, что оба ранга определены однозначно. В п. 3.3.20 мы отмечали, что линейная оболочка любого векторного набора представляет собой линейное подпространство, в том числе и линейная оболочка столбцов матрицы $A_{m\times n}$:

$$L_{col} = \left\{ \beta_1 \overset{|}{a}_1 + \beta_2 \overset{|}{a}_2 + \ldots + \beta_n \overset{|}{a}_n \mid \beta_1, \ldots, \beta_n \in \mathrm{F} \right\},$$

где

$$\overset{|}{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, \overset{|}{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

По теореме о размерности мы знаем, что подпространство L_{col} имеет определенную размерность. Она не превышает числа n, поскольку в случае, когда все столбцы матрицы линейно независимы, они сами и являются базисом подпространства L_{col}. Если же столбцы линейно зависимы, то одни из них выразимы через другие – например,

$$\overset{|}{a}_1 = \gamma_2 \overset{|}{a}_2 + \ldots + \gamma_n \overset{|}{a}_n .$$

Тогда выражение первого столбца можно подставить в любую линейную комбинацию столбцов, так что все подпространство L_{col} окажется линейной оболочкой меньшего набора векторов:

$$\beta_1 \overset{|}{a}_1 + \beta_2 \overset{|}{a}_2 + \ldots + \beta_n \overset{|}{a}_n = \beta_1 (\gamma_2 \overset{|}{a}_2 + \ldots + \gamma_n \overset{|}{a}_n) + \beta_2 \overset{|}{a}_2 + \ldots + \beta_n \overset{|}{a}_n =$$

$$= (\beta_1\gamma_2 + \beta_2) \overset{|}{a}_2 + \ldots + (\beta_1\gamma_n + \beta_n) \overset{|}{a}_n = \delta_2 \overset{|}{a}_2 + \ldots + \delta_n \overset{|}{a}_n .$$

Так можно исключать линейно выразимые столбцы, пока не останется линейно независимого набора – он и будет базисом в L_{col}, а его численность – размерностью подпространства L_{col} и рангом исходного столбцового множества. То есть *столбцовый ранг определен однозначно*.

Точно так же это устанавливается и для множества строк.

(2) Докажем, что ***столбцовый ранг равен заполнению трапециевидной формы матрицы***.

Как мы видели в п. 6.2.2, любая система линейных уравнений может быть переписана в векторно-стобцовой форме:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2; \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m; \end{cases} \Leftrightarrow$$

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} \cdot x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \cdot x_2 + \ldots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \cdot x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Для однородной системы это принимает вид исследования на линейную зависимость, или, что то же самое, взаимную выразимость:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0; \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0; \end{cases} \Leftrightarrow \quad \overset{|}{a}_1 x_1 + \overset{|}{a}_2 x_2 + \ldots + \overset{|}{a}_n x_n = \overset{|}{0}.$$

Если последнее равенство возможно при некоторых $x_k \neq 0$, мы назовем столбцы линейно зависимыми и сможем выразить один из них через остальные. А в общем виде решение такой однородной системы запишется так же, как и решение неоднородной при $X_{\text{ч.н}} = (0)$:

$$X_{\text{о.о}} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -d_{1\,r+1} \cdot x_{r+1} - \ldots - d_{1\,n} \cdot x_n \\ \vdots \\ -d_{r\,r+1} \cdot x_{r+1} - \ldots - d_{r\,n} \cdot x_n \\ x_{r+1} \\ \ddots \\ x_n \end{pmatrix} =$$

$$= \begin{pmatrix} -d_{1\,r+1} \\ \vdots \\ -d_{r\,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdot x_{r+1} + \ldots + \begin{pmatrix} -d_{1\,n} \\ \vdots \\ -d_{r\,n} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \cdot x_n$$

Свободным неизвестным можно придавать любые значения, в том числе поочередно придавать $x_{r+1}, \ldots, x_n$ значение 1, и значение 0 остальным: сначала $x_{r+1} = 1,\ x_{r+2} = 0,\ \ldots,\ x_n = 0,$ и так далее до $x_{r+1} = 0,\ \ldots,\ x_{n-1} = 0,\ x_n = 1.$

Тогда получатся $n - r$ частных решений:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix}_{r+1} = \begin{pmatrix} -d_{1\,r+1} \\ \vdots \\ -d_{r\,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \quad \ldots \quad ; \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix}_{n} = \begin{pmatrix} -d_{1\,n} \\ \vdots \\ -d_{r\,n} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Каждое из них задает выражение одного из «свободных» столбцов через «связанные»:

$$\overset{|}{0} = \overset{|}{a}_1 x_1 + \ldots + \overset{|}{a}_r x_r + \overset{|}{a}_{r+1} x_{r+1} + \ldots + \overset{|}{a}_n x_n =$$

$$=\begin{cases} \overset{(r+1)}{=} \; -\overset{|}{a}_1 d_{1\,r+1} - \ldots - \overset{|}{a}_r d_{r\,r+1} + \overset{|}{a}_{r+1} = \overset{|}{0}; \\ \vdots \\ \overset{(n)}{=} \; -\overset{|}{a}_1 d_{1\,n} - \ldots - \overset{|}{a}_r d_{r\,n} + \overset{|}{a}_n = \overset{|}{0}, \end{cases}$$

т. е. $n-r$ «свободных» столбцов выражаются через r «связанных» так:

$$\begin{cases} (r+1): \overset{|}{a}_{r+1} = -\overset{|}{a}_1 d_{1\,r+1} - \ldots - \overset{|}{a}_r d_{r\,r+1}; \\ \qquad\qquad \vdots \\ \quad (n): \quad \overset{|}{a}_n = \overset{|}{a}_1 d_{1\,n} + \ldots + \overset{|}{a}_r d_{r\,n}. \end{cases}$$

А «связанные» столбцы друг через друга не выражаются и линейно независимы, так как приравнивание к нулю их линейной комбинации означает запись той же однородной системы при нулевых значениях свободных неизвестных:

$$\overset{|}{0} = \overset{|}{a}_1 x_1 + \ldots + \overset{|}{a}_r x_r =$$

$$= \overset{|}{a}_1 x_1 + \ldots + \overset{|}{a}_r x_r + \overset{|}{a}_{r+1} \cdot 0 + \ldots + \overset{|}{a}_n \cdot 0 .$$

Но в этом случае решение системы полностью нулевое:

$$X_{\text{о.о}} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} -d_{1\,r+1} \cdot x_{r+1} - \ldots - d_{1\,n} \cdot x_n \\ \vdots \\ -d_{r\,r+1} \cdot x_{r+1} - \ldots - d_{r\,n} \cdot x_n \\ 0 \\ \ddots \\ 0 \end{pmatrix} =$$

$$= \begin{pmatrix} -d_{1\,r+1} \\ \vdots \\ -d_{r\,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdot 0 + \ldots + \begin{pmatrix} -d_{1\,n} \\ \vdots \\ -d_{r\,n} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \cdot 0.$$

Итак, через r линейно независимых «связанных» столбцов выражается $n-r$ «свободных»; значит, сами они образуют ранговую систему столбцов матрицы и, что то же самое, базис линейного подпространства L_{col}. Остается вспомнить, что возникающее в общем решении число r – это заполнение трапециевидной формы матрицы, и сделать **вывод**, что само *заполнение* также *определено однозначно*:

$$\text{col-rank}(A) = r.$$

Иными словами, приводя матрицу к трапециевидной форме разными способами, нельзя повлиять на заполнение r.

(3) Докажем, что и ***строчный ранг матрицы равен заполнению ее трапециевидной формы.*** Прежде всего, отметим линейную независимость ненулевых строк единичной слева формы матрицы.

Приравняем к нулю их линейную комбинацию:

$$\begin{array}{r} \alpha_1 \cdot \\ + \\ \alpha_2 \cdot \\ + \\ \vdots \\ + \\ \alpha_r \cdot \end{array} \begin{pmatrix} 1 & 0 & & 0 & d_{1\,r+1} & \cdots & d_{1\,n} \\ 0 & 1 & & 0 & d_{2\,r+1} & & \\ \vdots & & \ddots & & \vdots & & \\ 0 & 0 & & 1 & d_{r\,r+1} & \cdots & d_{r\,n} \end{pmatrix}$$

$$\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}$$

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_r & \Sigma & \ldots & \Sigma \end{pmatrix} = (0).$$

Отсюда видно, что все коэффициенты должны быть нулевыми.

Далее, сохранив сплошь нулевые строки в трапециевидной форме матрицы и выполнив в единичной слева форме обратные гауссовы преобразования в обратном порядке, мы восстановим строки исходной матрицы в их первоначальном виде, т. е. все они выражаются в виде линейных комбинаций строк единичной слева матрицы. Значит, строки единичной слева формы могут служить базисом линейной оболочки строк исходной матрицы – подпространства L_{row}. И если приведение матрицы к трапециевидной форме обошлось без столбцовых перестановок, то все доказано: размерность подпространства L_{row} равна r; таков же и строчный ранг матрицы.

Но и вмешательство перестановок столбцов никак не отражается на числе линейно независимых строк во всех гауссовых формах. Если восстановить исходный порядок столбцов в единичной слева форме, то левые r столбцов единичного блока могут смешаться с правыми $n-r$ столбцами – но и после этого строки останутся линейно независимыми, так как указанное выше обоснование равенства нулю всех коэффициентов линейной комбинации от порядка *записи* столбцов в матрицу не меняется.

Тем самым, восстановив в единичной слева форме матрицы исходный порядок столбцов, мы получим базис строчного подпространства L_{row}.

Отсюда $\text{row-rank}(A) = r$, и теорема полностью доказана.

Как **следствие**, можно говорить о ранге матрицы, не уточняя, столбцовый или строчный ранг имеется в виду; кроме того, приведение к трапециевидной форме оказывается *эффективным* с вычислительной точки зрения *способом нахождения значения* $\text{rank}(A)$.

ГЛАВА 8
ЛИНЕЙНЫЕ ОПЕРАТОРЫ

8.1. ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ

8.1.1. Понятие линейного отображения линейных пространств. Пусть V и W – линейные пространства над полем **F**. (Для простоты можно себе представлять, что **F** – это **R**, поле действительных чисел.) Если определено некоторое правило φ, которое каждому элементу $x \in V$ ставит в соответствие единственный элемент $y \in W$, в этом случае говорят, что φ является отображением пространства V в пространство W, и пишут $\varphi : V \to W$ или $\varphi(x) = y$, $x \in V$, $y \in W$, при этом элемент y называется ***образом x***, а x – ***прообразом y*** относительно отображения φ.

Определение. Отображение φ линейного пространства V в линейное пространство W называется ***линейным***, если:

$$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2), \quad x_1,\ x_2 \in V. \tag{Л1}$$

$$\varphi(\alpha\, x) = \alpha\varphi(x), \quad \alpha \in \mathbf{F},\ x \in V. \tag{Л2}$$

Сразу заметим, что если φ – линейное отображение, то $\varphi(\underline{0}) = \underline{0}$. Это легко следует из условия (Л2) и 3.2.6:

$$\varphi(\underline{0}) = \varphi(0x) = 0\varphi(x) = \underline{0}.$$

Приведем примеры линейных отображений.

Пример 1. Поворот плоскости на угол α вокруг фиксированной точки есть линейное отображение плоскости самой на себя. Свойство (Л1) ясно из правила параллелограмма, (Л2) – из сохранения длин при повороте.

Пример 2. В произвольном линейном пространстве V определим отображение φ: $V \to \{\underline{0}\}$, которое произвольному элементу x пространства V ставит в соответствие нулевой элемент $\underline{0}$. Такое отображение, очевидно, является линейным.

Пример 3. Все многочлены с действительными коэффициентами степени не более n с операциями сложения многочленов и умножения их на число образуют линейное пространство. Обозначим его L_n. Размерность этого пространства равна $n + 1$, в качестве базиса можно взять систему многочленов $1, x, \ldots, x^n$. Определим на этом пространстве отображение D, ставящее в соответствие многочлену $f[x]$ его первую производную $f'[x]$. Тогда нетрудно заметить, что D отображает L_n в L_{n-1}, при этом

$$D(f[x] + g[x]) = D(f[x]) + D(g[x]);$$

$$D(\alpha f[x]) = \alpha D(f[x]),$$

т. е. выполняются (Л1) и (Л2) и D является линейным отображением пространства L_n в пространство L_{n-1}.

8.1.2. Матрица линейного отображения в заданных базисах. Зафиксируем в пространстве V базис $\overline{v} = (v_1, \ldots, v_n)$, а в пространстве W базис $\overline{w} = (w_1, \ldots, w_k)$. Пусть $\varphi: V \to W$ – линейное отображение V в W. Поскольку $\varphi(v_1), \ldots, \varphi(v_n) \in W$, имеем:

$$\varphi(v_1) = a_{11}w_1 + a_{21}w_2 + \ldots + a_{k1}w_k,$$

$$\varphi(v_2) = a_{12}w_1 + a_{22}w_2 + \ldots + a_{k2}w_k,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\varphi(v_n) = a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{kn}w_k.$$

Из коэффициентов полученных равенств составим матрицу

$$A = \begin{pmatrix} a_{11}\ a_{12}\ \ldots\ a_{1n} \\ a_{21}\ a_{22}\ \ldots\ a_{2n} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{k1}\ a_{k2}\ \ldots\ a_{kn} \end{pmatrix}.$$

Матрица A называется ***матрицей линейного отображения*** φ в базисах $\overline{v}$ и $\overline{w}$ (см. 3.3.13). ***При этом столбцы матрицы*** A ***представляют собой коэффициенты разложения образов элементов*** $\varphi(v_1)$, …, $\varphi(v_n)$ ***базиса пространства*** V ***по базису пространства*** W.

Теперь убедимся в том, что, зная матрицу A, можно найти образ $\varphi(x)$ любого элемента $x \in V$. Действительно, пусть $x = x_1v_1 + \ldots + x_nv_n$. Тогда, в силу линейности отображения φ имеем

$$\varphi(x) = \varphi(x_1v_1 + \ldots + x_nv_n) = x_1\varphi(v_1) + \ldots + x_n\varphi(v_n) =$$
$$= x_1(a_{11}w_1 + a_{21}w_2 + \ldots + a_{k1}w_k) + \ldots + x_n(a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{kn}w_k).$$

Если раскрыть скобки и собрать коэффициенты при базисных элементах $w_1, \ldots, w_k$, получим

$$\varphi(x) = (x_1a_{11} + x_2a_{12} + \ldots + x_na_{1n})w_1 + \ldots +$$
$$+ (x_1a_{k1} + x_2a_{k2} + \ldots + x_na_{kn})w_k.$$

Обозначим

$$y_1 = x_1a_{11} + \ldots + x_na_{1n};$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$y_k = x_1a_{k1} + \ldots + x_na_{kn}.$$

Таким образом, если $\overset{|}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ – столбец координат элемента x, то y_1 есть произведение первой строки матрицы A на столбец $\overset{|}{x}$, соответственно y_k есть произведение k-й строки матрицы A на столбец $\overset{|}{x}$. Поэтому последние равенства очень удобно записывать в матричной форме. Обозначим через $\overset{|}{y}$ вектор-столбец координат элемента $y = \varphi(x)$ в базисе $\overline{w}$:

$$\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix},$$

тогда можно записать

$$\vec{y} = A\vec{x},$$

где произведение $A\vec{x}$ по правилам умножения матриц есть вектор-столбец длины k. Таким образом, задание линейного отображения φ на базисных элементах $v_1, \ldots, v_n$ однозначно определяет матрицу A, зная которую, в свою очередь, можно найти образ $\varphi(x)$ любого элемента $x \in V$, т. е. при заданных базисах в пространствах V и W линейному отображению φ можно поставить в соответствие прямоугольную матрицу размера $k \times n$.

С другой стороны, пусть дана

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{pmatrix}$$

– произвольная прямоугольная матрица размера $k \times n$. Покажем, что при заданных базисах она определяет линейное отображение φ пространства V в пространство W. Пусть $x \in V$, $x = x_1 v_1 + \ldots + x_n v_n$,

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

По определению полагаем

$$\varphi(x) = y_1 w_1 + \ldots + y_k w_k,$$

где $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}$ находится из записанного выше матричного равенства. Линейность определенного таким способом отображения φ следует из свойств операций над матрицами.

Итак, мы получили, что при фиксированных базисах пространств V и W любому линейному отображению $\varphi\colon V \to W$ соответствует прямоугольная матрица A и, наоборот, любая прямоугольная матрица A размера $k \times n$ определяет линейное отображение пространства V в пространство W. В связи с этим введем следующие обозначения:

$[\varphi]$ – матрица отображения φ;

φ_A – отображение, определяемое матрицей A.

Непосредственно из определений ясно, что отображение восстанавливается по матрице и наоборот: $\varphi = \varphi_{[\varphi]}$ и $A = [\varphi_A]$. Тем самым соответствие «отображение» $\leftrightarrow$ «матрица» взаимно однозначно. Кроме того, справедливы формулы $\varphi(x) = \overline{w}[\varphi]\overset{|}{x}$ и $\varphi_A(x) = \overline{w}A\overset{|}{x}$.

8.1.3. Пример 4. Найдем вид матриц линейных операторов из примеров 1–3 (8.1.1).

1. Пусть базисом будут векторы $\overline{i}, \overline{j}$. Тогда (см. рис. 39)

$$\varphi(\overline{i}) = (\cos\alpha,\quad \sin\alpha),\ \ \varphi(\overline{j}) = (-\sin\alpha,\quad \cos\alpha),\ \ [\varphi] = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

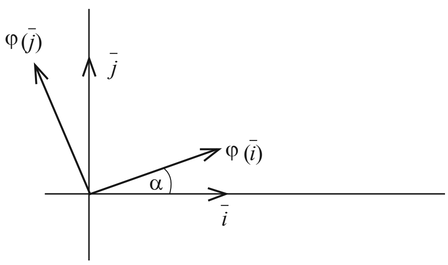


Рис. 39

(сравните с тем, о чем говорилось в п. 3.4.5).

2. Пусть $v_1, \ldots, v_n$ – базис пространства V. Тогда

$$\varphi(v_1) = \underline{0},\ \ \ldots,\ \ \varphi(v_n) = \underline{0}$$

и, следовательно,

$$[\varphi]=\begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

3. Так как $D : L_n \to L_{n-1}$, то для нахождения матрицы $[D]$ нужно выбрать базисы в пространствах L_{n-1} и L_n. Пусть это будут системы многочленов $1, x, \dots, x^{n-1}$ для пространства L_{n-1} и $1, \dots, x^n$ – для L_n. В этих базисах получаем такие столбцы координат для образов базисных векторов:

$$D(1)=\begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad D(x)=\begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad D(x^2)=\begin{pmatrix} 0 \\ 2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad D(x^n)=\begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ n \end{pmatrix}.$$

Следовательно,

$$[D]=\left(\underbrace{\begin{matrix} 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 2 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & n \end{matrix}}_{n+1}\right).$$

8.1.4. Ядро, образ, ранг и дефект линейного отображения. Очень часто возникают ситуации, когда удобно изучать не сам объект, а его образ при некотором отображении. В таких случаях нужно понимать, насколько «точно» можно судить по образу объекта о нем самом. В примере 1 п. 8.1.1 ясно, что образ «хорошо» сохраняет прообраз, так как есть взаимно однозначное соответствие между образом и прообразом. Ясно и то, что пример 2 представляет собой обратную ситуацию: по образу ничего нельзя сказать о прообразе:

всевозможные прообразы сливаются в один элемент $\underline{0}$. В примере 3 присутствует неоднозначность прообразов, так как отображение D все многочлены нулевой степени переводит в $\underline{0}$. Из рассмотренных примеров можно понять, что «точность» или «неточность» отображения определяется тем, как много элементов линейного пространства переводится этим отображением в элемент $\underline{0}$.

8.1.5. Определение. ***Ядром*** линейного отображения φ называется совокупность всех элементов пространства V, переводящихся отображением φ в нулевой элемент, а ***образом*** φ называется совокупность образов всех элементов из V.

Ядро принято обозначать Ker φ, а образ – Im φ.

Теорема. ***Ядро и образ линейного отображения являются подпространствами.***

Напомним (п. 3.3.17), что подмножество линейного пространства является подпространством тогда и только тогда, когда оно замкнуто относительно сложения элементов и умножения их на число. Поэтому для доказательства теоремы достаточно доказать эту замкнутость обоих подмножеств.

а) Пусть $x_1,\ x_2 \in \operatorname{Ker}\varphi$. Тогда очевидно, что

$$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) = \underline{0} + \underline{0} = \underline{0}, \quad \text{следовательно, } x_1 + x_2 \in \operatorname{Ker}\varphi.$$

Для любого числа α имеем

$$\varphi(\alpha x_1) = \alpha\varphi(x_1) = \alpha\underline{0} = \underline{0}, \text{ следовательно, } \alpha x_1 \in \operatorname{Ker}\varphi..$$

б) Пусть взяты произвольные элементы $y_1, y_2 \in \operatorname{Im}\varphi$. Тогда существуют такие x_1 и x_2, что $y_1 = \varphi(x_1)$, $y_2 = \varphi(x_2)$. Так как

$$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) = y_1 + y_2, \text{ то } y_1 + y_2 \in \operatorname{Im}\varphi.$$

Аналогично, из того, что $\varphi(\alpha x_1) = \alpha\varphi(x_1) = \alpha y_1$, следует, что $\alpha y_1 \in \operatorname{Im}\varphi$. Теорема доказана.

Линейное отображение пространства V в себя будем называть ***линейным преобразованием*** пространства V или ***линейным оператором***.

8.1.6. Определение. Размерность подпространства Im φ называется ***рангом*** преобразования φ, а размерность Ker φ – его ***дефектом***.

Теорема. ***Сумма ранга и дефекта линейного оператора*** φ ***равна размерности пространства V.***

Обозначим $\dim \operatorname{Ker} \varphi = d$, $\dim \operatorname{Im} \varphi = r$. Пусть $y_1,..., y_r$ – базис $\operatorname{Im} \varphi$ и $y_1 = \varphi(x_1)$, $\ldots$, $y_r = \varphi(x_r)$. Заметим, что элементы $x_1,\ldots, x_r$ линейно независимы. Действительно, пусть, напротив, $x_1,\ldots, x_r$ линейно зависимы. Тогда найдутся числа $\alpha_1,\ldots, \alpha_r$, не все равные нулю, и такие, что $\alpha_1 x_1 + \ldots + \alpha_r x_r = \underline{0}$. Отсюда

$$\varphi(\alpha_1 x_1 + \ldots + \alpha_r x_r) = \alpha_1 \varphi(x_1) + \ldots + \alpha_r \varphi(x_r) = \alpha_1 y_1 + \ldots + \alpha_r y_r = \underline{0}.$$

Последнее равенство при ненулевых $\alpha_1,\ldots, \alpha_r$ невозможно, поскольку $y_1,\ldots, y_r$ образуют базис $\operatorname{Im} \varphi$. Из полученного противоречия следует, что элементы $x_1,\ldots, x_r$ линейно независимы. Пусть далее $v_1,\ldots, v_d$ – базис подпространства $\operatorname{Ker} \varphi$. Если мы убедимся, что элементы $v_1,\ldots, v_d$, $x_1,\ldots, x_r$ образуют базис пространства V, теорема будет доказана. Для начала заметим линейную независимость этого набора. Пусть $\beta_1 x_1 + \ldots + \beta_r x_r + \gamma_1 v_1 + \ldots \gamma_d v_d = \underline{0}$ – некоторая линейная комбинация, обращающаяся в ноль. Тогда

$$\varphi(\beta_1 x_1 + \ldots + \beta_r x_r + \gamma_1 v_1 + \ldots + \gamma_d v_d) =$$
$$= \beta_1 \varphi(x_1) + \ldots + \beta_r \varphi(x_r) = \beta_1 y_1 + \ldots + \beta_2 y_2 = \underline{0}.$$

Отсюда видно, что $\beta_1 = \beta_2 = \ldots = \beta_r = 0$, так как $y_1,\ldots, y_r$ линейно независимы. Следовательно, $\gamma_1 v_1 + \ldots + \gamma_d v_d = \underline{0}$, откуда также $\gamma_1 = \gamma_2 = \ldots = \gamma_d = 0$, так как и $v_1,\ldots, v_d$ линейно независимы. Тогда и элементы $x_1,\ldots, x_r$, $v_1,\ldots, v_d$ линейно независимы. Осталось заметить, что любой элемент пространства V есть их линейная комбинация.

Пусть $b \in V$ – произвольный элемент. Тогда

$$\varphi(b) = \alpha_1 y_1 + \ldots + \alpha_r y_r = \alpha_1 \varphi(x_1) + \ldots + \alpha_r \varphi(x_r).$$

Понятно, что

$$\varphi(b - \alpha_1 x_1 - \ldots - \alpha_r x_r) = \underline{0}, \quad \text{и} \quad b - \alpha_1 x_1 - \ldots - \alpha_r x_r \in \operatorname{Ker} \varphi,$$

а значит, $b-\alpha_1 x_1-...-\alpha_r x_r=\alpha_{r+1}v_1+...+\alpha_{r+d}v_d$, откуда и следует, что $b=\alpha_1 x_1+...+\alpha_r x_r+\alpha_{r+1}v_1+...+\alpha_{r+d}v_d$.

Итак, $x_1,\ldots,\ x_r,\ \ v_1,\ldots,\ v_d$ образуют базис V и $\dim V=r+d$, или $\dim\ V=\dim\operatorname{Im}\varphi+\dim\operatorname{Ker}\varphi$, что и требовалось доказать.

8.1.7. Гомоморфизм и изоморфизм линейных пространств. Вернемся к рассуждениям о «точности» и «неточности» образа. Рассмотрим два случая: $\operatorname{Ker}\varphi=\{\underline{0}\}$ и $\operatorname{Ker}\varphi\neq\{\underline{0}\}$. В первом случае образы двух разных элементов $x_1,\ \ x_2\in V$ будут тоже различными, так как $x_1-x_2\neq\underline{0}$ и, следовательно, $\varphi(x_1-x_2)=\varphi(x_1)-\varphi(x_2)\neq\underline{0}$, т. е. $\varphi(x_1)\neq\varphi(x_2)$. Во втором случае может оказаться, что $x_1\neq x_2$, но при этом $x_1-x_2\in\operatorname{Ker}\varphi$, откуда

$$\varphi(x_1-x_2)=\varphi(x_1)-\varphi(x_2)=\underline{0}\ \text{ и }\ \varphi(x_1)=\varphi(x_2).$$

Схематически это выглядит так, как показано на рис. 40.

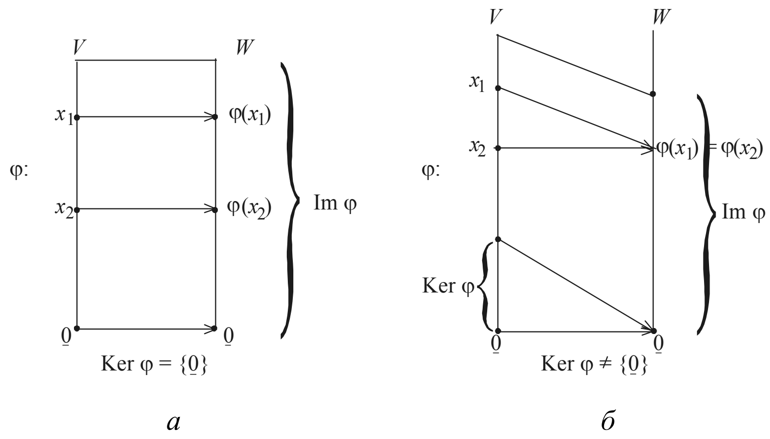


Рис. 40

Очевидно, что в первом случае образ «точен», а во втором «неточен». Эти два случая различаются с помощью понятий изоморфизма и гомоморфизма.

Дадим строгие определения.

8.1.8. Напомним, что понятие изоморфизма линейных пространств было введено в 3.3.8. Очевидно, имеет место

Утверждение. ***Линейные пространства V и W над полем F являются изоморфными, если существует линейное отображение $\varphi: V \to W$, устанавливающее взаимно однозначное соответствие между этими пространствами.***

В этом случае отображение φ является изоморфизмом пространства V на пространство W. Из взаимной однозначности следует, что

$$\operatorname{Ker} \varphi = \{\underline{0}\},$$

$$\varphi(V) = \operatorname{Im} \varphi = W.$$

Заметим, что, обратно, из этих равенств очевидно следует взаимная однозначность. Тогда по теореме 8.1.6 получается, что дефект изоморфизма равен нулю, а его ранг – размерности пространства V, т. е. пространства V и W имеют одинаковую размерность. Следовательно, матрица $[\varphi]$ квадратная.

8.1.9. Теперь рассмотрим вопрос о существовании обратного отображения.

Определение. Говорят, что линейное отображение $\varphi: V \to W$ ***обратимо***, если существует линейное отображение $\varphi^{-1}: \operatorname{Im}\varphi \to V$ такое, что для всякого $x \in V$ выполняется равенство $\varphi^{-1}(\varphi(x)) = x$.

Понятно, что не все линейные отображения обратимы. Скажем, отображение из примера 2(8.1.1) необратимо, если $V \neq \{\underline{0}\}$, так как для $x \neq \underline{0}$ $\varphi(x) = \underline{0}$, и если бы отображение φ^{-1} существовало, то в соответствии с замечанием в п. 8.1.1 мы получили бы $\varphi^{-1}(\varphi(x)) = \varphi^{-1}(\underline{0}) = \underline{0} \neq x$. Поэтому интересно описать класс обратимых линейных отображений.

Теорема. ***Линейное отображение $\varphi: V \to W$ обратимо тогда и только тогда, когда $\operatorname{Ker} \varphi = \{\underline{0}\}$.***

Пусть $\varphi: V \to W$ обратимо. Тогда существует линейное отображение $\varphi^{-1}: \operatorname{Im}\varphi \to V$, такое, что $\varphi^{-1}(\varphi(x)) = x$ для любого $x \in V$. Предположим, что $\operatorname{Ker} \varphi \neq \{\underline{0}\}$ и возьмем ненулевой $x \in \operatorname{Ker} \varphi$. Тогда $\varphi^{-1}(\varphi(x)) = \varphi^{-1}(\underline{0}) = \underline{0} \neq x$. Полученное противоречие доказывает утверждение теоремы в одну сторону.

Пусть, наоборот, $\operatorname{Ker}\varphi = \{\underline{0}\}$. Тогда для всякого $y \in \operatorname{Im}\ \varphi$ существует единственный $x \in V$ такой, что $\varphi(x) = y$. Единственность такого элемента x следует из того, что если для некоторого $y \in \operatorname{Im}\varphi$ существуют $x_1 \neq x_2$, для которых $\varphi(x_1) = \varphi(x_2) = y$, то

$$\varphi(x_1 - x_2) = \varphi(x_1) - \varphi(x_2) = \underline{0} \quad \text{и} \quad \underline{0} \neq (x_1 - x_2) \in \operatorname{Ker}\varphi,$$

что противоречит равенству $\operatorname{Ker}\varphi = \{\underline{0}\}$. Полученное противоречие доказывает утверждение о существовании единственного прообраза для всякого $y \in \operatorname{Im}\varphi$.

Теперь положим $\varphi^{-1}(y) = x$. Докажем, что определенное таким образом отображение является линейным. По определению отображение φ^{-1} задано на всем $\operatorname{Im}\varphi$. Осталось заметить, что выполняются свойства (Л1) и (Л2):

$$\varphi^{-1}(y_1 + y_2) = \varphi^{-1}(y_1) + \varphi^{-1}(y_2); \tag{Л1}$$

$$\varphi^{-1}(\alpha y) = \alpha\varphi^{-1}(y), \quad \alpha \in \mathbf{F}. \tag{Л2}$$

Л1. Пусть $y_1 = \varphi(x_1), \quad y_2 = \varphi(x_2)$, тогда

$$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) = y_1 + y_2.$$

Отсюда $\qquad \varphi^{-1}(y_1 + y_2) = x_1 + x_2 = \varphi^{-1}(y_1) + \varphi^{-1}(y_2).$

Л2. Пусть $y = \varphi(x)$, тогда

$$\alpha y = \alpha\varphi(x) = \varphi(\alpha x), \quad \varphi^{-1}(\alpha y) = \alpha x = \alpha\varphi^{-1}(y).$$

Тем самым теорема доказана.

8.1.10. В общем случае, когда мы говорим о линейном отображении $\varphi: V \to W$, то не предполагаем, что $\operatorname{Im}\varphi = W$, поэтому последний случай надо выделить особым названием. Иногда про такой случай говорят, что отображение φ является ***отображением «на»***, само отображение называется ***сюръективным***, а также ***эпиморфизмом***. Теперь можно сформулировать следствие из теоремы 8.1.8.

Следствие. ***Эпиморфизм*** $\varphi : V \to W$ ***обратим тогда и только тогда, когда он является изоморфизмом.***

Из обратимости φ следует, что $\operatorname{Ker} \varphi = \{\underline{0}\}$, а из того что φ – это эпиморфизм, ясно, что $\operatorname{Im} \varphi = W$. Из полученных условий вытекает взаимная однозначность отображения φ, т. е. по определению φ есть изоморфизм.

Если, наоборот, отображение φ является изоморфмизмом, то $\operatorname{Ker} \varphi = \{\underline{0}\}$, и изоморфизм φ обратим. Следствие доказано.

Отметим, что если φ – это изоморфизм, то обратное отображение φ^{-1} – тоже изоморфизм.

8.1.11. Теорема. ***Линейное отображение является изоморфизмом тогда и только тогда, когда его матрица*** $[\varphi]$ ***в любых базисах пространств*** *V* ***и*** *W* ***является квадратной обратимой матрицей.***

Пусть отображение φ – изоморфизм. В п. 8.1.8 мы уже замечали, что матрица изоморфизма [φ] является квадратной. Пусть $\overline{v} = (v_1, \ldots, v_n)$ и $\overline{w} = (w_1, \ldots, w_n)$ – фиксированные базисы пространств *V* и *W*, [φ] – матрица изоморфизма φ в этих базисах, $\overset{|}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ – координаты элемента $x \in V$ в базисе $\overline{v}$, $\overset{|}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ – координаты элемента $y = \varphi(x)$ в базисе $\overline{w}$, так что $x = (v_1, \ldots, v_n)\overset{|}{x}$, $y = (w_1, \ldots, w_n)\overset{|}{y}$. Тогда по формуле п. 8.1.2 имеем $\overset{|}{y} = [\varphi]\overset{|}{x}$. Так как φ^{-1} – тоже изоморфизм, то по той же формуле $\overset{|}{x} = [\varphi^{-1}]\overset{|}{y}$. Отсюда получаем равенство $\overset{|}{x} = [\varphi^{-1}][\varphi]\overset{|}{x}$. Так как $\overset{|}{x}$ – это произвольный вектор-столбец из пространства $\mathbf{F}^n$, получаем, что $[\varphi^{-1}][\varphi] = E$, откуда следует, что $[\varphi^{-1}] = [\varphi]^{-1}$ и матрица [φ] обратима.

Пусть, наоборот, *A* – обратимая квадратная матрица порядка *n*, φ_A – линейное отображение пространств $V \to W$, определяемое по формуле $y = \varphi_A(x) = (w_1,\ldots, w_n)A\overset{|}{x}$. Тогда отображение $\varphi_{A^{-1}}$, определяемое формулой $\varphi_{A^{-1}}(y) = (v_1,\ldots, v_n)A^{-1}\overset{|}{y}$, является линейным отображением пространств $W \to V$, обратным к φ_A, так как из равенства $AA^{-1}\overset{|}{y} = \overset{|}{y}$ вытекает

$$\varphi_A\varphi_{A^{-1}}(y) = (w_1,\ldots, w_n)AA^{-1}\overset{|}{y} = (w_1,\ldots, w_n)\overset{|}{y} = y,$$

равно как из равенства $A^{-1}A\overset{|}{x} = \overset{|}{x}$ вытекает $\varphi_{A^{-1}}\varphi_A(x) = x$, т. е. φ_A – обратимое отображение. Теорема доказана.

8.1.12. Можно обобщить понятие изоморфизма**.**

Определение. Линейное отображение $\varphi: V \to W$ называется ***гомоморфизмом*** пространства V на пространство W, если для любого элемента $y \in W$ найдется $x \in V$ такой, что $y = \varphi(x)$.

По сравнению с изоморфизмом здесь отсутствует требование взаимной однозначности, однако сохраняется требование $\operatorname{Im}\varphi = W$. Можно сказать, что изоморфизм – это гомоморфизм с нулевым ядром. Образно говоря, изоморфный образ – это «точная» копия объекта, а гомоморфный – «сокращенная». В зависимости от обстоятельств мы отдаем предпочтение либо тому, либо другому отображению. Вспомним наши примеры п. 8.1.1. Теперь понятно, что отображение в примере 1 – это изоморфизм, так как здесь $\operatorname{Ker}\varphi = \{\underline{0}\}$. В примере 2 – это гомоморфизм, но не изоморфизм (в случае $V \ne \{\underline{0}\}$), так как $\operatorname{Ker}\varphi = V$. Отображение в примере 3 – тоже гомоморфизм, так как $\operatorname{Ker} D = L_0 \ne \{\underline{0}\}$ и оно не является изоморфизмом.

8.1.13. Рассмотрим частный случай линейных отображений, когда $W = V$, т. е. когда гомоморфизм φ отображает пространство V в себя. Здесь $\operatorname{Im}\varphi \subset V$ и отображение φ в п. 8.1.5 мы назвали ***линейным оператором, действующим на векторном пространстве*** V. Если при этом $\operatorname{Ker}\varphi = \{\underline{0}\}$, то говорят об изоморфном отображении пространства V на себя, если же $\operatorname{Ker}\varphi \ne \{\underline{0}\}$, то такой оператор называется ***эндоморфизмом***. Поскольку в случае $W = V$ размерности пространств одинаковы, матрицы эндоморфизмов являются квадратными. Ниже мы остановимся на этом подробнее.

8.2. МАТРИЦЫ ЛИНЕЙНЫХ ОПЕРАТОРОВ

8.2.1. Преобразование матрицы линейного оператора при замене базиса. Выше мы уже отмечали, что можно говорить о матрице линейного оператора в данном базисе. Пусть по-прежнему $\varphi: V \to V$ – линейный оператор в пространстве V. Так как в пространстве имеется бесчисленное множество базисов, а матрица $[\varphi]$ зависит от базиса, то естественно попытаться понять, как меняется матрица линейного оператора при переходе от одного базиса к другому. Пусть $\overline{e} = (e_1, e_2, \ldots, e_n)$ – базис пространства V, назовем его «старым». Другой базис $\overline{e'} = (e'_1, e'_2, \ldots, e'_n)$ назовем «новым». Разложим элементы старого базиса по новому:

$$e_1 = \tau_{11} e'_1 + \tau_{21} e'_2 + \ldots + \tau_{n1} e'_n,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$e_n = \tau_{1n} e'_1 + \tau_2 e'_2 + \ldots + \tau_{nn} e'_n.$$

Тогда по определению п. 3.3.13 матрица

$$S = \begin{pmatrix} \tau_{11} \ \ldots \ \tau_{1n} \\ \ldots\ldots\ldots\ldots \\ \tau_{n1} \ \ldots \ \tau_{nn} \end{pmatrix}$$

является матрицей перехода от старого базиса $\overline{e}$ к новому базису $\overline{e'}$.

Посмотрим, как меняется матрица $[\varphi]$ линейного оператора φ при переходе от базиса $\overline{e}$ к базису $\overline{e'}$. Удобно обозначить матрицу φ в базисах $\overline{e}$ и $\overline{e'}$ через $[\varphi]_e$ и $[\varphi]_{e'}$. Пусть $x \in V$ – произвольный элемент, $\overset{|}{x}$ и $\overset{|}{x}{}'$, как и выше, – столбцы координат элемента x в базисах $\overline{e}$ и $\overline{e'}$ соответственно. Пусть $\varphi(x) = y$, и $\overset{|}{y}$ и $\overset{|}{y}{}'$ – столбцы координат элемента y в базисах $\overline{e}$ и $\overline{e'}$. В силу формул в конце п. 8.1.2 $\overset{|}{y} = [\varphi]_e \overset{|}{x}$; $\overset{|}{y'} = [\varphi]_{e'} \overset{|}{x'}$. Так как по формулам из п. 3.3.14 $\overset{|}{y'} = S\overset{|}{y}$, $\overset{|}{x} = S^{-1}\overset{|}{x'}$, то

$$[\varphi]_{e'} \overset{|}{x'} = \overset{|}{y'} = S\overset{|}{y} = S[\varphi]_e \overset{|}{x} = S[\varphi]_e S^{-1} \overset{|}{x'}.$$

Сравнивая начало и конец, получаем

$$[\varphi]_{e'} \overset{|}{x'} = S[\varphi]_e S^{-1} \overset{|}{x'}.$$

Поскольку это матричное равенство справедливо при любом столбце $\overset{|}{x'} \in \mathbf{F}^n$, из него получаем равенство для матриц

$$[\varphi]_{e'} = S[\varphi]_e S^{-1}.$$

Эта формула и дает ***правило преобразования матрицы линейного оператора при переходе к новому базису***.

Заметим, что матрицы A и B, связанные равенством $B = TAT^{-1}$ для некоторой матрицы T, называются ***подобными.*** Поэтому ***матрицы линейного оператора*** φ ***в разных базисах подобны***.

8.2.2. Изоморфизм алгебры линейных операторов и алгебры матриц. В п. 2.1.9 было дано определение алгебры над полем. Там же было замечено, что множество квадратных матриц с элементами из поля $\mathbf{F}$ образует алгебру над $\mathbf{F}$.

В п. 8.1.12 мы показали, что любому линейному оператору n-мерного пространства V можно поставить в соответствие квадратную матрицу порядка n. Введем на множестве линейных операторов сложение, умножение и умножение на элементы поля $\mathbf{F}$. Пусть φ, ψ – линейные операторы, $\alpha \in \mathbf{F}$, $x \in V$. Положим

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x), \tag{O1}$$

$$(\varphi\psi)(x) = \varphi(\psi(x)), \tag{O2}$$

$$(\alpha\varphi)(x) = \alpha\varphi(x). \tag{O3}$$

После такого определения операций над операторами понятно, что множество всех линейных операторов пространства V тоже образует алгебру над $\mathbf{F}$. Легко понять, что ***алгебра матриц и алгебра операторов изоморфны***. Для этого нужно установить между ними взаимно однозначное соответствие и показать, что это соответствие сохраняет все операции. Такое соответствие было установлено в п. 8.1.2. Убедимся в сохранении операций. Пусть φ, ψ – линейные операторы, $x \in V$ – некоторый элемент, $\overset{|}{x}$ – вектор-столбец координат элемента x в базисе $\overline{e}$ пространства V. Тогда, как и в п. 8.1.2,

$$\varphi(x) = \overline{e}[\varphi]_e \overset{|}{x}.$$

Поэтому $(\varphi+\psi)(x) = \overline{e}[\varphi+\psi]_e \overset{|}{x}$.

С другой стороны, согласно (O1)

$$(\varphi+\psi)(x) = \varphi(x) + \psi(x) = \overline{e}[\varphi]_e \overset{|}{x} + \overline{e}[\psi]_e \overset{|}{x} = \overline{e}([\varphi]_e + [\psi]_e) \overset{|}{x}.$$

Сравнивая оба равенства, получаем соответствие для операторной и матричной операций сложения:

$$[\varphi+\psi]_e = [\varphi]_e + [\psi]_e.$$

Аналогично доказывается соответствие операций умножения операторов друг на друга и умножения на элементы из поля $\mathbf{F}$:

$$[\varphi\psi]_e = [\varphi]_e [\psi]_e,$$

$$[\alpha\varphi]_e = \alpha[\varphi]_e.$$

Тем самым доказан изоморфизм алгебры линейных операторов и алгебры квадратных матриц.

8.3. СОБСТВЕННЫЕ ЭЛЕМЕНТЫ

8.3.1. Собственные числа и собственные элементы. Выше мы видели, что вид матрицы линейного оператора зависит от базиса пространства. Поскольку базисов существует бесконечное множество, можно искать такие базисы, в которых матрица оператора имеет наиболее простой вид. Решению этой задачи будут посвящены ближайшие подразделы. В них будем считать, что **F = R** – поле действительных чисел. Перейдем к определениям.

Определение. Ненулевой элемент $x \in V$ называется ***собственным элементом***, а $\lambda \in \mathbf{R}$ – ***собственным числом*** оператора φ, если $\varphi(x) = \lambda x$. В этом случае говорят, что собственный элемент x соответствует собственному числу λ.

8.3.2. Попытаемся понять, как искать собственные числа и собственные элементы операторов. Во-первых, заметим, что не все операторы имеют собственные элементы. Например, оператор из примера 1 (см. п. 8.1.1) таких элементов не имеет, поскольку при повороте плоскости на угол, не кратный 180°, ни один ненулевой вектор не сохраняет своего направления (при повороте на 180° все элементы – собственные, а каково их собственное число, подумайте сами).

Зафиксируем в пространстве V базис $\overline{e}$. Тогда, как и выше, $[\varphi]_e = A$ – матрица оператора в этом базисе, а $\overset{|}{x}$ – столбец координат элемента $x \in V$.

Предположим, что оператор имеет собственный элемент x, относящийся к собственному числу λ_0. Тогда из определения следует, что выполняется равенство

$$\varphi(x) = \lambda_0 x \quad \text{или} \quad \overline{e}[\varphi]_e \overset{|}{x} = \overline{e}(\lambda_0 \overset{|}{x}) \quad \text{и} \quad [\varphi]\overset{|}{x} = \lambda_0 \overset{|}{x}.$$

Отсюда, учитывая, что $[\varphi]_e = A$, получаем

$$A\overset{|}{x} - \lambda_0 \overset{|}{x} = (0) \quad \text{или} \quad (A - \lambda_0 E)\overset{|}{x} = (0).$$

Требование $\overset{|}{x} \neq (0)$ на языке систем линейных уравнений означает, что данная однородная система линейных уравнений имеет ненулевое решение. Как мы заметили в конце п. 7.3.1, это выполняется тогда и только тогда, когда матрица системы вырожденная (определитель матрицы системы равен нулю), т. е. $|A - \lambda_0 E| = 0$.

Определитель $|A - \lambda E| = f[\lambda]$ представляет собой многочлен n-й степени относительно λ, он называется ***характеристическим***

многочленом матрицы A (или, что то же самое, ***характеристическим многочленом оператора*** φ[1]). Равенство нулю определителя $|A-\lambda_0 E|$ означает, что λ_0 является корнем характеристического многочлена. Так как по определению $\lambda_0 \in \mathbf{R}$, то собственное число матрицы A является *действительным* корнем ее характеристического многочлена. О неприводимых квадратных трехчленах читателю известно из школьного курса, а мы отмечали в пп. 1.3.2–1.3.4, что не все корни многочлена с действительными коэффициентами обязательно действительны, поэтому и ***не все корни характеристического многочлена будут собственными числами***.

Однако ***если*** λ_0 ***– действительный корень характеристического многочлена, то он является собственным числом***. В самом деле, пусть $f[\lambda_0]=0$, $\lambda_0 \in \mathbf{R}$. Тогда однородная система уравнений с вырожденной матрицей $(A-\lambda_0 E)$ по замечанию 7.3.1 имеет ненулевое решение $\overset{|}{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, т. е. $(A-\lambda_0 E)\overset{|}{x}=(0)$ или $A\overset{|}{x}-\lambda_0 \overset{|}{x}=(0)$, или $A\overset{|}{x}=\lambda_0 \overset{|}{x}$, откуда и следует, что x – собственный элемент оператора φ. Таким образом, справедлива следующая

Теорема. ***Действительные характеристические корни оператора*** φ***, и только они являются собственными числами этого оператора.***

8.3.3. Из наших рассуждений также следует, что ***множество собственных элементов оператора*** φ, ***соответствующих одному и тому же собственному значению*** λ_0, ***после присоединения нулевого элемента образует линейное подпространство***. Это следует из того, что все эти элементы служат решениями однородной системы линейных уравнений (см. п. 6.4.1.). Обозначим это подпространство $L^{(\lambda_0)}$.

[1] Хотя в разных базисах оператор φ имеет разные матрицы, все они подобны (п. 8.2.1), а подобные матрицы имеют одинаковые характеристические многочлены (п. 8.3.5).

Что же можно сказать о собственных элементах, относящихся к разным собственным значениям? Оказывается, справедлива следующая

Теорема. ***Собственные элементы, относящиеся к различным собственным значениям, линейно независимы.***

Пусть элементы $b_1, \ldots, b_k$ являются собственными элементами оператора φ, относящимися к различным собственным значениям $\lambda_1, \ldots, \lambda_k$. Предположим, что они линейно зависимы. Тогда существует их нетривиальная линейная комбинация, равная нулевому элементу:

$$\beta_1 b_1 + \ldots + \beta_k b_k = \underline{0}.$$

Подействуем на эту комбинацию оператором φ и учтем, что $\varphi(b_i) = \lambda_i b_i$:

$$\beta_1 \varphi(b_1) + \ldots + \beta_k \varphi(b_k) = \beta_1 \lambda_1 b_1 + \ldots + \beta_k \lambda_k b_k = \underline{0}.$$

Умножим предыдущее равенство на λ_1 и вычтем из него последнее:

$$\beta_2 (\lambda_1 - \lambda_2) b_2 + \ldots + \beta_k (\lambda_1 - \lambda_k) b_k = \underline{0}$$

или

$$\beta'_2 b_2 + \ldots + \beta'_k b_k = \underline{0}.$$

Если $\beta_2 = \ldots = \beta_k = 0$, то $\beta_1 \neq 0$ и тогда с необходимостью $b_1 = \underline{0}$, что невозможно по определению собственного элемента. Значит, не все числа $\beta_2, \ldots, \beta_k$ равны нулю.

Так как все числа $\lambda_1, \ldots, \lambda_k$ разные, то $(\lambda_2 - \lambda_1) \neq 0$, $(\lambda_3 - \lambda_1) \neq 0, \ \ldots, \ (\lambda_k - \lambda_1) \neq 0$ и среди чисел $\beta'_2, \ldots, \beta'_k$ не все равны нулю. Левая часть равенства является нетривиальной линейной комбинацией элементов $b_2, \ldots, b_k$. Это означает, что элементы $b_2, \ldots, b_k$ линейно зависимы. Проведя такие рассуждения $(k-1)$ раз, в конце концов получим $b_k = \underline{0}$, что противоречит определению собственного элемента. Это противоречие доказывает утверждение теоремы.

8.3.4. Спектр линейного оператора. Множество всех ***характеристических чисел*** (корней характеристического уравнения) оператора φ принято называть ***спектром*** этого оператора.

Определение. Говорят, что линейный оператор φ, действующий на пространстве $V \cong \mathbf{R}^n$, имеет ***простой*** спектр, если все его характеристические числа действительны и различны.

В силу теоремы 8.3.3 все собственные элементы линейного оператора с простым спектром линейно независимы. Поскольку число этих элементов n, они образуют базис пространства $\mathbf{R}^n$. Так как для этих элементов $\varphi(b_1) = \lambda_1 b_1,\ \ldots,\ \varphi(b_n) = \lambda_n b_n$, то в базисе из собственных элементов матрица оператора φ имеет вид

$$A = \begin{pmatrix} \lambda_1 & 0 & \vdots & 0 \\ 0 & \lambda_2 & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \lambda_n \end{pmatrix}.$$

Такой вид матрицы называется ***диагональным***.

Справедливо и обратное, а именно: если в некотором базисе $e_1,\ldots,\ e_n$ матрица A линейного оператора φ имеет диагональный вид, то $\varphi(e_1) = \lambda_1 e_1,\ \ldots,\ \varphi(e_n) = \lambda_n e_n$, так как для всех $i = 1,\ldots,\ n$

$$\varphi[e_i] = \overline{e} \begin{pmatrix} \lambda_1 & 0 & & & 0 \\ 0 & \ddots & & & \\ & & \lambda_i & & \\ & & & \ddots & 0 \\ 0 & & & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \_ i = \overline{e} \begin{pmatrix} 0 \\ \vdots \\ \lambda_{\mathrm{i}} \\ \vdots \\ 0 \end{pmatrix} = \lambda_i e_i .$$

Таким образом, для линейных операторов нами получена

Теорема. ***Линейный оператор*** φ ***тогда и только тогда задается в базисе*** $e_1,\ldots,\ e_n$ ***диагональной матрицей, когда все элементы этого базиса являются собственными элементами оператора*** **φ.**

Так как при переходе от базиса к базису матрица линейного оператора подвергается преобразованию подобия (8.2.1), то можно сказать, что ***всякая матрица оператора с простым спектром подобна диагональной матрице, т. е. приводится к диагональному виду переходом к базису из собственных элементов.***

Обратите внимание, что теорема дает *критерий* диагонализуемости матрицы оператора с помощью перехода к подходящему (собственному) базису, а последнее замечание – *достаточное условие* диагонализуемости.

8.3.5. Инварианты характеристического многочлена. Мы видели в 8.1.2, что линейному оператору φ в фиксированном базисе можно поставить в соответствие матрицу [φ]. Вид этой матрицы зависит от базиса и меняется при переходе к другому базису. В п. 8.2.1 было замечено, что матрицы одного и того же линейного оператора в разных базисах подобны, т. е. связаны соотношением $B = TAT^{-1}$. Рассмотрим характеристический многочлен матрицы *B:* $|B-\lambda E| = g[\lambda]$. Докажем, что $g[\lambda] = f[\lambda] = |A-\lambda E|$. Здесь $f[\lambda]$ – характеристический многочлен матрицы A.

Действительно, по свойству 2.5.6

$$g[\lambda] = |B-\lambda E| = |TAT^{-1} - \lambda E| = |TAT^{-1} - \lambda TET^{-1}| =$$

$$= |T(A-\lambda E)T^{-1}| = |T| \cdot |A-\lambda E| \cdot |T^{-1}| = |A-\lambda E| = f[\lambda]$$

(мы воспользовались тем, что $|T| \cdot |T^{-1}| = |E| = 1$).

Отсюда следует, что ***характеристические числа линейного оператора не зависят от базиса, в котором рассматривается матрица линейного оператора***.

В заключение поговорим о коэффициентах характеристического многочлена матрицы A. Рассмотрим частный случай матрицы третьего порядка. В этом случае имеем

$$f[\lambda] = |A-\lambda E| = \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}.$$

Чтобы уловить закономерность в выражении коэффициентов характеристического многочлена через элементы матрицы, представим полученный определитель в виде суммы нескольких определителей, пользуясь линейностью определителя по каждому столбцу:

$$f[\lambda]=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}+\begin{vmatrix} -\lambda & a_{12} & a_{13} \\ 0 & a_{22}-\lambda & a_{23} \\ 0 & a_{32} & a_{33}-\lambda \end{vmatrix}=$$

$$=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}+\begin{vmatrix} a_{11} & 0 & a_{13} \\ a_{21} & -\lambda & a_{23} \\ a_{31} & 0 & a_{33}-\lambda \end{vmatrix}+$$

$$+\begin{vmatrix} -\lambda & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33}-\lambda \end{vmatrix}+\begin{vmatrix} -\lambda & 0 & a_{13} \\ 0 & -\lambda & a_{23} \\ 0 & 0 & a_{33}-\lambda \end{vmatrix}=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+$$

$$+\begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & -\lambda \end{vmatrix}+\begin{vmatrix} a_{11} & 0 & a_{13} \\ a_{21} & -\lambda & a_{23} \\ a_{31} & 0 & a_{33} \end{vmatrix}+\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & -\lambda & 0 \\ a_{31} & 0 & -\lambda \end{vmatrix}+$$

$$+\begin{vmatrix} -\lambda & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} -\lambda & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & a_{32} & -\lambda \end{vmatrix}+\begin{vmatrix} -\lambda & 0 & a_{13} \\ 0 & -\lambda & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix}+$$

$$+\begin{vmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix}=(-\lambda)^3+(-\lambda)^2(a_{11}+a_{22}+a_{33})+$$

$$+(-\lambda)\left(\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}\right)+\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Анализируя полученный результат, замечаем:

1) коэффициент при старшем члене $(-\lambda)^3$ равен единице;

2) коэффициент при $(-\lambda)^2$ равен сумме диагональных элементов матрицы A;

3) коэффициент при $(-\lambda)$ равен сумме диагональных миноров второго порядка, т. е. *алгебраических дополнений* к элементам главной диагонали (см. п.2.5.1);

4) свободный член равен определителю матрицы A:

$$|A-\lambda E| = -\lambda^3 + \lambda^2 \sum_{i=1,2,3} a_{ii} - \lambda \sum_{i=1,2,3} A_{ii} + |A|.$$

Обобщенная формула для определителя n-го порядка выглядит так:

$$|A-\lambda E| = (-\lambda)^n + (-\lambda)^{n-1} \sum_{i=1}^{n} a_{ii} + (-\lambda)^{n-2} \sum_{1 \le i_1 < i_2 \le n} d_{i_1 i_2} + \ldots +$$

$$+(-\lambda)^{n-k} \sum_{1 \le i_1 < \ldots < i_k \le n} d_{i_1 \ldots i_k} + \ldots + |A|.$$

Здесь $d_{i_1 \ldots i_k}$ – диагональный минор k-го порядка матрицы A, стоящий на пересечении строк и столбцов с одинаковыми номерами $i_1, \ldots, i_k$.

8.3.6. Рассмотрим пример, поясняющий применение этой формулы. Найдем характеристический многочлен матрицы

$$A = \begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix}.$$

$$f[\lambda] = (-\lambda)^4 + (-\lambda)^3 (3+1+5-1) +$$

$$+(-\lambda)^2 \left(\begin{vmatrix} 3 & -1 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 3 & 0 \\ 3 & 5 \end{vmatrix} + \begin{vmatrix} 3 & 0 \\ 4 & -1 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 0 & 5 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ -1 & -1 \end{vmatrix} + \begin{vmatrix} 5 & -3 \\ 3 & -1 \end{vmatrix} \right) +$$

$$+(-\lambda) \left(\begin{vmatrix} 3 & -1 & 0 \\ 1 & 1 & 0 \\ 3 & 0 & 5 \end{vmatrix} + \begin{vmatrix} 3 & -1 & 0 \\ 1 & 1 & 0 \\ 4 & -1 & -1 \end{vmatrix} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & 5 & -3 \\ -1 & 3 & -1 \end{vmatrix} + \begin{vmatrix} 3 & 0 & 0 \\ 3 & 5 & -3 \\ 4 & 3 & -1 \end{vmatrix} \right) +$$

$$+\begin{vmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{vmatrix} =$$

$$=\lambda^4-8\lambda^3+\lambda^2(4+15-3+5-1+4)-\lambda(20-4+4+12)+16=$$

$$=\lambda^4-8\lambda^3+24\lambda^2-32\lambda+16=(\lambda-2)^4.$$

8.3.7. Поскольку подобные матрицы имеют одинаковый характеристический многочлен (8.3.5), коэффициенты этого многочлена не меняются при преобразовании подобия, т. е. являются ***инвариантами*** при таком преобразовании. Поэтому ***коэффициенты характеристического многочлена линейного оператора не зависят от того, в каком базисе рассматривается матрица этого оператора.***

Особый интерес представляет второй коэффициент. Он носит специальное название.

Определение. Сумма диагональных элементов матрицы A называется ее ***следом*** и обозначается trA.

Как мы уже заметили, $tr(TAT^{-1})=trA$. Легко установить и равенство $tr(AB)=tr(BA)$.

8.3.8. Характеристические числа матриц A^{T} и A^{-1}. Заметим два простых факта.

1. Характеристические числа матрицы A^{T} являются характеристическими числами матрицы A. Это следует из равенства определителей исходной и транспонированной матриц (2.4.1):

$$|A-\lambda E|=|(A-\lambda E)^{\mathrm{T}}|=|A^{\mathrm{T}}-\lambda E|.$$

2. Если A – обратимая матрица, то ***характеристические числа матрицы A^{-1} обратны к характеристическим числам матрицы A.***

Заметим, что характеристические числа обратимой матрицы также обратимы, т. е. не могут равняться нулю: ведь их произведение равно свободному члену характеристического многочлена – определителю матрицы. Пусть теперь v – собственный вектор, а λ – собственное число матрицы A, т. е. $Av = \lambda v$. Тогда $| A - \lambda E \ | = \ 0$, следовательно,

$$| A | \left| A^{-1} - \frac{1}{\lambda} E \right| = \left| AA^{-1} - \frac{1}{\lambda} A \right| = \left| E - \frac{1}{\lambda} A \right| = -\frac{1}{\lambda^n} | A - \lambda E | = 0 .$$

Так как $| A | \neq 0$, отсюда получаем $\left| A^{-1} - \frac{1}{\lambda} E \ \right| = \ 0$. Здесь $\lambda \neq 0$, поскольку $| A | \neq 0$. Таким образом, λ является характеристическим числом матрицы A тогда и только тогда, когда $\frac{1}{\lambda}$ – характеристическое число матрицы A^{-1}.

8.4. СОБСТВЕННЫЕ ПОДПРОСТРАНСТВА

8.4.1. Сумма, пересечение и прямая сумма подпространств. С понятием подпространства мы уже знакомы (3.3.17). Поскольку подпространства являются подмножествами пространства, естественно рассмотреть на них операции, принятые над множествами: объединение и пересечение. С другой стороны, операции желательно определить так, чтобы результатом их применения были опять подпространства. Такими операциями являются сумма (а не объединение) и пересечение подпространств. Пусть L_1 и L_2 – подпространства пространства L.

Напомним (3.3.21), что ***суммой*** подпространств L_1 и L_2 называется множество

$$L_1 + L_2 = \{a_1 + a_2 \mid a_1 \in L_1, \quad a_2 \in L_2\} .$$

Прежде всего, заметим, что сумма подпространств тоже будет подпространством. Пусть взяты некоторые элементы $a, b \in L_1 + L_2$, $\alpha \in \mathbf{F}$. Как мы указали в п. 3.3.17, требуется доказать, что для них $a + b \in L_1 + L_2$ и $\alpha a \in L_1 + L_2$. Из определения следует, что возможно разложение:

$$a = a_1 + a_2, \quad b = b_1 + b_2 ,$$

где $a_1,\ b_1 \in L_1, \quad a_2,\ b_2 \in L_2$. Тогда

$$a+b=a_1+a_2+b_1+b_2=(a_1+b_1)+(a_2+b_2).$$

Так как теперь $(a_1+b_1)\in L_1$, $(a_2+b_2)\in L_2$, то по определению $a+b\in L_1+L_2$. Аналогично можно показать, что $\alpha a\in L_1+L_2$.

Определение. ***Пересечением*** подпространств L_1 и L_2 называется множество

$$L_1\bigcap L_2=\{a\,|\,a\in L_1 \text{ и } a\in L_2\}.$$

Очевидно, что $L_1\bigcap L_2$ является подпространством.

Если пересечение подпространств L_1 и L_2 состоит из нулевого элемента, их сумму называют прямой и обозначают $L_1\oplus L_2$ (см. п. 3.3.21).

8.4.2. Уточним вопрос о размерности суммы двух подпространств. Напомним, что размерностью подпространства называется число элементов в базисе этого подпространства (3.3.9). Справедлива

Теорема. ***Размерность суммы подпространств равна сумме их размерностей за вычетом размерности их пересечения***:

$$\dim(L_1+L_2)=\dim(L_1)+\dim(L_2)-\dim(L_1\bigcap L_2).$$

Для доказательства обозначим $\dim L_1=n_1$, $\dim L_2=n_2$. Пусть $\bar{a}=(a_1,\ldots,\ a_k)$ – базис пересечения $L_1\bigcap L_2$.

Очевидны неравенства $k\le n_1$, $k\le n_2$. Так как верно включение $L_1\bigcap L_2\subseteq L_1$, базис $\bar{a}$ можно достроить до базиса подпространства L_1 – пусть это будут элементы $b_{k+1},\ldots,\ b_{n_1}$. Аналогично достроим базис $\bar{a}$ до базиса L_2 – пусть это будут элементы $c_{k+1},\ldots,\ c_{n_2}$. Понятно, что любой элемент суммы L_1+L_2 выражается через элементы

$$a_1,\ldots,\ a_k,\ \ b_{k+1},\ldots,\ b_{n_1},\ \ c_{k+1},\ldots,\ c_{n_2}.$$

Заметим, что эти элементы линейно независимы. Пусть, напротив, существует их нетривиальная линейная комбинация, равная нулевому элементу:

$$\alpha_1a_1+\ldots+\alpha_ka_k+\beta_{k+1}b_{k+1}+\ldots+\beta_{n_1}b_{n_1}+\gamma_{k+1}c_{k+1}+\ldots+\gamma_{n_2}c_{n_2}=\underline{0}.$$

Среди чисел $\gamma_{k+1},\ldots,\gamma_{n_2}$ найдутся не равные нулю, так как иначе элементы $a_1,\ldots,a_k$, $b_{k+1},\ldots,b_{n_1}$ оказались бы линейно зависимыми и набор $(a_1,\ldots,a_k,\ b_{k+1},\ldots,b_{n_1})$ не был бы базисом. Обозначим $v=\gamma_{k+1}c_{k+1}+\ldots+\gamma_{n_2}c_{n_2}$. По построению $v\in L_2$. Из равенства линейной комбинации нулевому элементу следует, что $v\in L_1$, значит, $v\in L_1\bigcap L_2$. Тогда $v=\alpha'_1a_1+\ldots+\alpha'_ka_k$. Получаем равенство

$$\alpha'_1a_1+\ldots+\alpha'_ka_k-\gamma_{k+1}c_{k+1}-\ldots-\gamma_{n_2}c_{n_2}=\underline{0},$$

что невозможно, так как базисные элементы $a_1,\ldots,a_k$, $c_{k+1},\ldots,c_{n_2}$ линейно независимы. Полученное противоречие доказывает линейную независимость системы элементов

$$a_1,\ldots,a_k,\ b_{k+1},\ldots,b_{n_1},\ c_{k+1},\ldots,c_{n_2}.$$

Значит, эти элементы образуют базис, откуда получается требуемое:

$$\dim(L_1+L_2)=k+(n_1-k)+(n_2-k)=n_1+n_2-k=$$

$$=\dim L_1+\dim L_2-\dim(L_1\cap L_2).$$

Следствие. ***Размерность прямой суммы подпространств равна сумме размерностей слагаемых:***

$$\dim\,(L_1\oplus L_2)=\dim L_1+\dim L_2.$$

Этот результат является частью теоремы 3.3.26.

8.4.2. Собственные подпространства линейного оператора.

Определение. Подпространство $L\subset V$ называется ***инвариантным*** относительно линейного оператора φ, если для всякого $x\in L$ верно включение $\varphi(x)\in L$. Для любого линейного оператора φ существуют, по крайней мере, два инвариантных подпространства: $\{\underline{0}\}$ и V. Эти подпространства называются ***тривиальными инвариантными подпространствами оператора*** φ.

Рассмотрим примеры нетривиальных инвариантных подпространств. Прежде всего вспомним о собственных элементах линейных операторов. ***Одномерные подпространства, порожденные собственными элементами, являются инвариантными относительно оператора*** φ. В самом деле, пусть x – собственный элемент оператора φ, относящийся к собственному значению λ_0. Рассмотрим так называемую ***линейную оболочку***, натянутую на элемент x, т. е. одномерное подпространство $L_{\lambda_0} = \{\alpha\, x \mid \alpha \in \mathbf{R}\}$ (см. п. 3.3.20). Покажем, что оболочка L_{λ_0} инвариантна относительно действия φ. Пусть $a \in L_{\lambda_0}$, тогда $a = \alpha x$ и, значит,

$$\varphi(a) = \varphi(\alpha x) = \alpha\varphi(x) = \alpha\lambda_0 x \in L_{\lambda_0},$$

что и требовалось.

В п. 8.3.3 мы замечали, что множество собственных элементов оператора φ, соответствующих одному и тому же собственному значению λ_0, вместе с нулевым элементом тоже образует пространство, которое обозначается $L^{(\lambda_0)}$. Такие подпространства называются ***собственными подпространствами*** оператора φ. Они тоже инвариантны относительно этого оператора. Действительно, пусть $a \in L^{(\lambda_0)}$; покажем, что $\varphi(a) \in L^{(\lambda_0)}$: в самом деле, $\varphi(\varphi(a)) = \varphi(\lambda_0 a) = \lambda_0\varphi(a)$, т. е. $\varphi(a)$ является собственным элементом оператора φ с тем же собственным значением λ_0. Отличие подпространств L_{λ_0} и $L^{(\lambda_0)}$ состоит в том, что L_{λ_0} одномерно, а $L^{(\lambda_0)}$ может иметь и большую размерность.

8.4.4. Напомним, что мы пытаемся среди произвольных базисов пространства найти такие, в которых матрица линейного оператора имеет как можно более простой вид. В п. 8.3.4 был указан такой базис в том частном случае, когда оператор φ имеет простой спектр. Однако не все операторы имеют простой спектр. Но оказывается, что базисы, связанные с инвариантными относительно линейного оператора подпространствами, тоже упрощают вид матрицы оператора.

Пусть $L \subset V$ – инвариантное относительно оператора φ подпространство и $b_1, \ldots, b_k$ – базис подпространства L. В п. 3.3.19 показано, что базис любого подпространства можно достроить до базиса всего пространства. Сделаем это. Пусть $b_1,\ldots,\ b_k,\ e_{k+1},\ldots,\ e_n$ – базис пространства V. Нужно понять, какой вид имеет матрица $[\varphi]$ в этом базисе. Так как $\varphi(b_i)\in L$ для всех $i=1,\ldots,k,$ то $\varphi(b_i)=\alpha_{1i}b_1+\ldots+\alpha_{ki}b_k$. Поэтому имеем

$$[\varphi]=\left(\begin{array}{ccc|ccc}\alpha_{11} & \vdots & \alpha_{1k} & \alpha_{1k+1} & \vdots & \alpha_{1n}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \alpha_{k1} & \vdots & \alpha_{kk} & \vdots & \vdots & \vdots\\ \hline 0 & \vdots & 0 & \vdots & \vdots & \vdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & \vdots & 0 & \alpha_{nk+1} & \vdots & \alpha_{nn}\end{array}\right).$$

Такой вид называется ***клеточно-треугольным***.

Если окажется, что подпространство $L'=\{e_{k+1},\ldots,\ e_n\}$, порожденное элементами $e_{k+1},\ldots,\ e_n$, тоже инвариантно относительно оператора φ, т. е. $\varphi(L')\subset L',$ то вид матрицы $[\varphi]$ станет еще проще: $[\varphi]=\left(\begin{array}{c|c}* & 0\\ \hline 0 & *\end{array}\right)$. Такой вид называется ***клеточно-диагональным***. Заметим, что $L\cap L'=\{\underline{0}\}$, и в соответствии с п. 8.4.1 пространство V оказывается прямой суммой $V=L\oplus L'$.

Если теперь окажется, что $V=L_1\oplus L_2\oplus\ldots\oplus L_s$, где все подпространства L_i $(i=1,\ldots,\ s)$ инвариантны относительно оператора φ, то в базисе V, составленном из базисов подпространств $L_1,\ldots,\ L_s$, матрица $[\varphi]$ имеет клеточно-диагональный вид:

$$[\varphi]=\begin{pmatrix}A_1 & & & 0\\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_s\end{pmatrix}.$$

В том частном случае, о котором говорилось в п. 8.3.4, разложение пространства в прямую сумму инвариантных подпространств было следующим: $V=L_{\lambda_1}\oplus L_{\lambda_2}\oplus\ldots L_{\lambda_n}$, и соответственно матрица [φ] имела диагональный вид с одномерными клетками, так как все подпространства $L_{\lambda_1},\ldots, L_{\lambda_n}$ одномерны.

В п. 8.3.4 мы поняли, что если оператор φ обладает простым спектром, то его матрица приводится к диагональному виду. Но наличие простого спектра не является необходимым условием приведения к диагональному виду. Как известно, единичная матрица, являясь диагональной, не имеет простого спектра. Поэтому интерес представляет следующая теорема, которую мы приводим без доказательства

Теорема. ***Для того чтобы матрица линейного оператора приводилась к диагональному виду над полем* R *действительных чисел, необходимо и достаточно, чтобы все корни ее характеристического многочлена были действительными и размерности собственных подпространств были равны кратностям соответствующих собственных значений как корней характеристического многочлена***.

Таким образом, если характеристический многочлен оператора φ разлагается на линейные множители над полем действительных чисел:

$$f[\lambda] = (-1)^n (\lambda - \lambda_1)^{k_1} \ldots (\lambda - \lambda_m)^{k_m}$$

и размерности собственных подпространств $L^{(\lambda_1)},\ldots, L^{(\lambda_m)}$ оказались равны степеням скобок в разложении многочлена (т. е. кратностям корней $\lambda_1,\ldots, \lambda_m$!):

$$\dim L^{(\lambda_1)} = k_1,\ \ldots,\ \dim L^{(\lambda_m)} = k_m,$$

то матрица [φ] оператора φ приводится к диагональному виду в «сборном» базисе подпространств $L^{(\lambda_1)},\ldots, L^{(\lambda_m)}$.

ГЛАВА 9
ЕВКЛИДОВЫ ПРОСТРАНСТВА

9.1. ЛИНЕЙНЫЕ, БИЛИНЕЙНЫЕ И КВАДРАТИЧНЫЕ ФОРМЫ

9.1.1. Линейные формы. Почти во всех отношениях на n-мерное линейное пространство можно смотреть, как на обобщение понятия трехмерного пространства геометрических векторов. Однако до сих пор в наших рассмотрениях отсутствовали важнейшие геометрические величины, такие, как длина вектора и угол между векторами. В этом пункте мы введем аналоги этих понятий для линейных пространств произвольной размерности.

В разд. 8.1 мы рассматривали линейные отображения линейных пространств. При этом элементам пространства V ставились в соответствие элементы пространства W или, в случае линейных операторов, самого пространства V. Сейчас мы рассмотрим важный частный случай, а именно такие отображения, когда элементам пространства V над полем $\mathbf{F}$ ставятся в соответствие числа поля $\mathbf{F}$ (или, что то же самое, элементы одномерного пространства над этим полем), т. е. числовые функции, определенные на V.

Определение. Числовая функция $L(x)$ от одного аргумента x, определенная на линейном пространстве V над числовым полем $\mathbf{F}$, называется ***линейной формой***, если она удовлетворяет следующим условиям:

а) $L(x+y)=L(x)+L(y)$ для любых $x, y \in V$;

б) $L(\alpha x)=\alpha L(x)$ для любых $x \in V$, $\alpha \in \mathbf{F}$.

Из а) и б) легко получаем, что

$$L(\alpha_1 x_1+\ldots+\alpha_k x_k)=\alpha_1 L(x_1)+\ldots+\alpha_k L(x_k).$$

Примером линейной формы на $\mathbf{R}^3$ может служить скалярное произведение с одним фиксированным сомножителем, т. е. $L(\overline{x}) = (\overline{a},\ \overline{x})$, где $\overline{a} \in \mathbf{R}^3$ – фиксированный вектор, $\overline{x} \in \mathbf{R}^3$ – произвольный. Выполнение условий а) и б) очевидно следует из свойств скалярного произведения. Если при этом $\overline{a} = (a_1;\ a_2;\ a_3)$, $\overline{x} = (x_1;\ x_2;\ x_3)$ – координаты векторов в некотором ортонормированном базисе, то, как известно,

$$(\overline{a},\ \overline{x}) = a_1x_1 + a_2x_2 + a_3x_3 = \sum_{i=1}^{3} a_ix_i .$$

Найдем вид произвольной линейной формы в линейном пространстве V. Пусть $e_1,..., e_n$ – базис пространства V, $x \in V$ – элемент, раскладывающийся по базису так: $x = x_1e_1 + ... + x_ne_n$. Тогда

$$L(x) = L(x_1e_1 + ... + x_ne_n) = x_1L(e_1) + ... + x_nL(e_n).$$

Так как $e_1,..., e_n$ – фиксированные элементы, то числа $L(e_1) = \alpha_1,..., L(e_n) = \alpha_n$ тоже фиксированы. Следовательно,

$$L(x) = \alpha_1x_1 + ... + \alpha_nx_n = \sum_{i=1}^{n} \alpha_ix_i .$$

Это и есть общий вид линейной формы в n-мерном линейном пространстве.

9.1.2. Билинейные и квадратичные формы. Теперь рассмотрим числовые функции двух аргументов.

Определение. Числовая функция $A(x, y)$ от двух аргументов x, y, определенная на линейном пространстве V над числовым полем $\mathbf{F}$, называется ***билинейной формой***, если она является линейной функцией от x при каждом фиксированном y и линейной функцией от y при каждом фиксированном x, т. е. функция $A(x, y)$ удовлетворяет условиям:

а) $A(x + z, y) = A(x, y) + A(z, y)$;

б) $A(\alpha x, y) = \alpha A(x, y)$;

в) $A(x, y+z) = A(x, y) + A(x, z)$;

г) $A(x, \alpha y) = \alpha A(x, y)$

для всех $x, y, z \in V$, $\alpha \in \mathbf{F}$.

Для любых элементов $x_1, \ldots, x_k$, $y_1, \ldots, y_m \in V$ и любых чисел $\alpha_1, ..., \alpha_k$, $\beta_1, ..., \beta_m \in \mathbf{F}$ из условий а) – г) легко получаем

$$A\left(\sum_{i=1}^{k} \alpha_i x_i, \quad \sum_{j=1}^{m} \beta_j y_j\right) = \sum_{j=1}^{m} \sum_{i=1}^{k} \alpha_i \beta_j A(x_i, y_j).$$

Используя это равенство, найдем общий вид билинейной формы. Пусть $\overline{e} = (e_1, .., e_n)$ – базис пространства V**,**

$$x = \sum_{i=1}^{n} x_i e_i; \qquad y = \sum_{j=1}^{n} y_j e_j.$$

Тогда $A(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j A(e_i, e_j)$.

Обозначим

$$A(e_i, e_j) = a_{ij}.$$

В этих обозначениях получим

$$A(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j,$$

что и дает нам общий вид билинейной формы.

Матрица $A_e = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ называется матрицей билинейной формы в базисе $\overline{e}$.

Заметим, что последнее равенство легко записывается в матричной форме. Будем, как и раньше, считать, что

$$V = \mathbf{F}^n \quad \text{и} \quad \overset{|}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \overset{|}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

В свою очередь, $\overline{x}$ и $\overline{y}$ будут обозначать строки, т. е. $\overline{x} = (x_1, ..., x_n)$, $\overline{y} = (y_1, ..., y_n)$. После этих замечаний перепишем равенство в матричной форме:

$$A(x, y) = \overline{x} A_e \overset{|}{y}.$$

Действительно,

$$\overline{x} A_e \overset{|}{y} = (x_1, ..., x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} =$$

$$= \left(\sum_{i=1}^{n} a_{i1} x_i, ..., \sum_{i=1}^{n} a_{in} x_i \right) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}\, x_i y_j .$$

Точно так же $A(x, y) = \overline{y} A_e^{\mathrm{T}} \overset{|}{x}$.

Примером билинейной формы может служить скалярное произведение в $\mathbf{R}^3$. Если $x, y \in \mathbf{R}^3$, то, как известно, через координаты в ортонормированном базисе оно записывается так:

$$(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 .$$

Это согласуется с общим случаем при

$$a_{ij} = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

9.1.3. На скалярном произведении остановимся подробнее, так как, с одной стороны, оно является примером билинейной формы, а с другой – тесно связано с «метрикой» $\mathbf{R}^3$, потому что длину вектора можно определить как квадратный корень из его скалярного квадрата: $|\overline{a}| = \sqrt{(\overline{a}, \overline{a})}$, а угол между двумя векторами – как арккосинус отношения их скалярного произведения к произведению длин этих векторов: $\angle\, \overline{a}\overline{b} = \arccos \dfrac{(\overline{a}, \overline{b})}{|\overline{a}|\ |\overline{b}|}$.

Получается, что если в пространстве определено скалярное произведение, то с его помощью можно ввести длины векторов и углы между ними. Как мы только что заметили, скалярное произведение является билинейной формой, но не всякую билинейную форму можно принять за скалярное произведение. Чтобы это понять, вспомним свойства скалярного произведения:

$$(x, y) = (y, x); \tag{C1}$$

$$(\alpha x, y) = \alpha(x, y), \quad \alpha \in \mathbf{R}; \tag{C2}$$

$$(x + y, z) = (x, z) + (y, z); \tag{C3}$$

$$(x, x) > 0\text{, если } x \neq \underline{0} \text{ и } (x, x) = 0\text{, только если } x = \underline{0}. \tag{C4}$$

Из этих четырех свойств билинейная форма обладает только двумя: (C2) и (C3). А формы, обладающие свойствами (C1) и (C4), носят специальные названия.

9.1.4. Определение. Билинейная форма $A(x, y)$ называется ***симметрической***, если для всех элементов x и y $A(x, y) = A(y, x)$.

Определение. Билинейная форма $A(x, y)$ называется ***положительно определенной***, если для любого ненулевого элемента x выполняется неравенство $A(x, x) > 0$.

Теперь понятно, что именно симметрическая и положительно определенная билинейная форма может играть роль скалярного произведения.

От билинейной формы легко перейти к квадратичной форме, о геометрическом смысле которой мы поговорим несколько позже, а пока только введем определения.

9.1.5. Определение. ***Квадратичной формой*** в линейном пространстве над полем **F** называется числовая функция $Q(x) = A(x, x)$ от одного аргумента, которая получается из произвольной билинейной формы $A(x, y)$ заменой y на x.

Общий вид квадратичной формы получается из общего вида билинейной формы заменой y_j на x_j:

$$Q(x) = A(x, x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j .$$

Поскольку $x_i x_j = x_j x_i$, в последней записи удобно предполагать $a_{ij} = a_{ji}$, ведь если это не так, то этого нетрудно добиться, преобразовав матрицу билинейной формы:

$$a_{ij} x_i x_j + a_{ji} x_j x_i = (a_{ij} + a_{ji})\, x_i x_j = \frac{a_{ij} + a_{ji}}{2} x_i x_j + \frac{a_{ij} + a_{ji}}{2} x_j x_i .$$

Поэтому ***матрица квадратичной формы всегда считается симметрической, т. е. удовлетворяющей условию*** $a_{ij} = a_{ji}$ ***при всех*** i, j. ***Это означает, что*** $A^{\mathrm{T}} = A$.

9.2. ПОНЯТИЕ ЕВКЛИДОВА ПРОСТРАНСТВА

9.2.1. Определение метрики евклидова пространства. А теперь сделаем то, о чем говорили в начале п. 9.1.1. На основе понятия линейного пространства определим так называемое ***евклидово пространство***, в котором уже будут присутствовать привычные «метрические» характеристики.

Определение. Линейное пространство V, определенное над полем действительных чисел **R**, называется ***евклидовым***, если на нем определена числовая фукция двух аргументов (x, y), называемая ***скалярным произведением*** и удовлетворяющая условиям:

$$(x, y) = (y, x); \qquad \text{(E1)}$$

$$(x + y, z) = (x, z) + y, z); \qquad \text{(E2)}$$

$$(\alpha x, y) = \alpha(x, y); \qquad \text{(E3)}$$

$$(x, x) > 0 \text{ при } x \neq 0,\ (x, x) = 0 \text{ при } x = \underline{0} \qquad \text{(E4)}$$

для всех $x, y, z \in V$ и $\alpha \in \mathbf{R}$.

Как мы заметили выше, роль скалярного произведения может играть любая симметрическая и положительно определенная билинейная форма.

9.2.2. Приведем примеры евклидовых пространств.

1. Совокупности геометрических векторов на прямой, на плоскости, в пространстве, рассматриваемые в аналитической геомет-

рии. Их принято обозначать V_1, V_2, V_3. Во всех трех пространствах скалярное произведение вводится по формуле

$$(\overline{x}, \overline{y}) = |\overline{x}||\overline{y}|\cos\angle\,\overline{xy}.$$

2. Пространство $\mathbf{R}^n$ упорядоченных наборов действительных чисел со скалярным произведением, определенным по формуле

$$(x, y) = x_1y_1 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i,$$

где $x = (x_1; \ldots; x_n)$, $y = (y_1; \ldots; y_n)$.

3. Пространство $\mathbf{R}(a, b)$ вещественных функций, непрерывных на отрезке $[a, b]$ со скалярным произведением, определенным по формуле

$$(f, g) = \int_a^b f(x)g(x)dx.$$

Читателю предлагается убедиться в том, что скалярные произведения, введенные двумя последними формулами, удовлетворяют условиям (Е1)–(Е4) из определения евклидова пространства.

9.2.3. Теперь определим длину (норму) элемента и угол между двумя элементами.

Определение. ***Длиной*** элемента $x \in V$ называется величина $|x| = \sqrt{(x, x)}$.

Определение. ***Углом между двумя ненулевыми элементами*** $x, y \in V$ называется угол φ, такой, что $0 \le \varphi \le \pi$ и

$$\cos\varphi = \frac{(x, y)}{|x|\,|y|}.$$

Если один из элементов – нулевой, угол считаем произвольным.

9.2.4. Неравенства Коши–Буняковского и треугольника. Нужно убедиться в том, что только что определенные длина и угол не противоречат привычным геометрическим свойствам, т. е. что в евклидовом пространстве выполняются соотношения для сторон треугольника, а величина косинуса угла не может оказаться по модулю больше единицы. Начнем с последнего.

Рассмотрим элемент $\alpha x - y$, где $\alpha \in \mathbf{R}$, $x, y \in V$.

По условию (Е4) из 9.2.1 имеем

$$(\alpha x - y,\ \alpha x - y) \geq 0.$$

Используя (Е1)–(Е3), преобразуем последнее равенство

$$\alpha^2 (x, x) - 2\alpha (x, y) + (y, y) \geq 0.$$

Это квадратный трехчлен относительно переменной α и, поскольку для любых значений α он сохраняет знак, его дискриминант не может быть больше нуля, т. е.

$$D = 4(x, y)^2 - 4(x, x)\,(y, y) \leq 0$$

или

$$(x, y)^2 \leq (x, x)\,(y, y).$$

После извлечения квадратного корня получаем

$$-|x|\,|y| \leq (x, y) \leq |x|\,|y|.$$

Это неравенство называется ***неравенством Коши–Буняковского***. Из него следует, что дробь, определяющая косинус в п. 9.2.3, по модулю не превосходит единицы. Заметим, что если в качестве скалярного произведения взять такую же форму суммы произведений одинаковых координат, что и в $\mathbf{R}^3$, то получится неравенство

$$|x_1 y_1 + \ldots + x_n y_n| \leq \sqrt{x_1^2 + \ldots + x_n^2}\ \sqrt{y_1^2 + \ldots + y_n^2}\ .$$

Выясним, когда здесь достигается равенство. Если $|(x, y)| = |x|\,|y|$, то дискриминант квадратного трехчлена равен нулю и, следовательно, существует единственный корень, обращающий этот квадратный трехчлен в нуль. Пусть α_0 – этот корень. Тогда

$$\alpha_0^2 (x, x) - 2\alpha_0 (x, y) + (y, y) = 0$$

или

$$(\alpha_0 x - y,\ \alpha_0 x - y) = 0.$$

Отсюда по условию (E4) следует, что $\alpha_0 x - y = \underline{0}$ и $y = \alpha_0 x$, т. е. элементы x и y коллинеарны. Обратно, если элементы x и y коллинеарны, т. е. $y = \alpha x$, где $\alpha \in \mathbf{R}$, то достигается равенство

$$|(x,y)| = |(x,\alpha x)| = |\alpha|\,|(x,x)| = |\alpha|\;|x|^2 =$$
$$= |x|\;|\alpha|\;|x| = |x|\;|\alpha x| = |x|\;|y|.$$

Итак, ***в неравенстве Коши–Буняковского равенство достигается тогда и только тогда, когда элементы*** x ***и*** y ***коллинеарны.*** Следовательно, и угол φ (9.2.3) между элементами получится равным нулю или π только в этом случае.

Переходим к неравенствам треугольника. Из школьного курса геометрии известно, что длина любой стороны треугольника не больше суммы длин двух других его сторон и не меньше модуля разности длин этих сторон. Пусть $x, y \in V$. По аналогии с векторами на плоскости будем считать, что элементы x, y, $x+y$ образуют стороны треугольника. Значит, нужно убедиться в том, что

$$|x+y| \le |x|+|y| \quad \text{и} \quad |x+y| \ge \big||x|-|y|\big|.$$

По неравенству Коши–Буняковского:

$$|x+y|^2 = (x+y,\, x+y) = (x,x) + 2(x,y) + (y,y) \le$$
$$\le |x|^2 + 2|x|\;|y| + |y|^2 = (|x|+|y|)^2.$$

Извлекая квадратный корень, получаем требуемое:

$$|x+y| \le |x|+|y|.$$

Если же воспользоваться левым из двух неравенств Коши–Буняковского, то получим

$$|x+y|^2 \ge |x|^2 - 2|x|\;|y| + |y|^2 = (|x|-|y|)^2.$$

Извлекая квадратный корень, получаем второе неравенство: $|x+y| \ge \big||x|-|y|\big|$.

Эти вновь полученные неравенства называются ***неравенствами треугольника***.

9.3. ПРОЦЕСС ОРТОГОНАЛИЗАЦИИ

9.3.1. Поскольку мы определили угол между элементами, появилась возможность говорить об ортогональности элементов. Вспомним, что в аналитической геометрии удобным критерием ортогональности служит равенство нулю скалярного произведения (4.1.7).

Определение. Элементы $x, y \in V$ называются ***ортогональными,*** если $(x, y) = 0$. Если $x \neq \underline{0}$ и $y \neq \underline{0}$, то это означает, что косинус угла между ними равен нулю, т. е. что определяемый по формуле 9.2.3 угол между ними равен 90°. Если же один из элементов нулевой, то его можно считать ортогональным ко всем элементам из V.

С понятием ортогональности связаны не только метрические соотношения. Очень удобным в пространстве геометрических векторов был базис, составленный из векторов $\overline{i}, \overline{j}, \overline{k}$ – попарно ортогональных векторов единичной длины. Оказывается, аналогичные базисы существуют и в евклидовом пространстве.

9.3.2. Теорема. ***Система попарно ортогональных ненулевых элементов евклидова пространства является линейно независимой.***

Пусть элементы $b_1, ..., b_k$ попарно ортогональны, т. е. для любых $1 \leq i < j \leq k$ выполняется равенство $(b_i, b_j) = 0$. Предположим, что они линейно зависимы. Тогда найдутся числа $\alpha_1, \ldots, \alpha_k$, не все равные нулю и такие, что $\alpha_1 b_1 + ... + \alpha_k b_k = \underline{0}$. Будем считать, что $\alpha_1 \neq 0$. Домножим скалярно обе части равенства на b_1:

$$\alpha_1(b_1, b_1) + ... + \alpha_k(b_k, b_1) = 0.$$

Из попарной ортогональности элементов $b_1, ..., b_k$ следует равенство $\alpha_1 |b_1|^2 = 0$.

Так как по условию $|b_1| \neq 0$, это означает, что $\alpha_1 = 0$, что противоречит нашему предположению. Полученное противоречие и доказывает теорему.

Так как в евклидовом пространстве выполняются основные «метрические» соотношения из пространства геометрических векторов, не удивительно, что остаются справедливыми и теорема Пифагора и ее обобщение. Теорема Пифагора на языке векторов плоскости утверждает, что для ортогональных векторов (катетов!) $\overline{x}$, $\overline{y}$ выполняется равенство $|\overline{x}+\overline{y}|^2=|\overline{x}|^2+|y|^2$. Если x, y – элементы евклидова пространства с условием $(x, y)=0$, то легко получаем аналогичное равенство

$$|x+y|^2=(x+y,\ x+y)=(x,\ x)+(y,\ y)=|x|^2+|y|^2.$$

Это утверждение можно обобщить.

Теорема Пифагора. ***Если*** $x_1,\ldots,x_k$ ***– система попарно ортогональных элементов евклидова пространства, то***

$$|x_1+\ldots+x_k|^2=|x_1|^2+\ldots+|x_k|^2.$$

Доказательство – аналогично приведенному выше.

9.3.3. Процесс ортогонализации, матрица Грама–Шмидта. Сейчас мы опишем процесс, с помощью которого можно из произвольного базиса пространства V получить ортогональный базис. Этот процесс называется процессом ортогонализации. Вначале рассмотрим пример в трехмерном пространстве. Путь $\overline{a}$, $\overline{b}$, $\overline{c}$ – линейно независимые и, следовательно, некомпланарные векторы. Для удобства будем полагать, что все три вектора имеют одно и то же начало. Опишем процесс «переработки» этих векторов в попарно ортогональные векторы $\overline{a}'$, $\overline{b}'$, $\overline{c}'$. Через $l(\overline{a})$ обозначим прямую, содержащую вектор $\overline{a}$, через $\pi(\overline{a}, \overline{b})$ – плоскость, содержащую векторы $\overline{a}$ и $\overline{b}$, через $\text{пр}_l\ \overline{a}$ – проекцию вектора $\overline{a}$ на прямую l, которую можно определить аналогично проекции вектора на плоскость (см. п. 4.2.6).

1. На первом шаге в качестве вектора $\overline{a}'$ выбираем вектор $\overline{a}$, так что $l(\overline{a}')=l(\overline{a})$.

2. На втором шаге в плоскости векторов $\overline{a}'$ и $\overline{b}$ строим прямую l', перпендикулярную прямой $l(\overline{a}')$, и раскладываем вектор $\overline{b}$ в сумму двух проекций: $\overline{b} = \text{пр}_{l(\overline{a}')}\overline{b}+\text{пр}_{l'}\overline{b}$. В качестве вектора $\overline{b}'$ выбираем проекцию $\overline{b}$ на l', т. е. $\overline{b}' = \overline{b}-\text{пр}_{l(\overline{a}')}\overline{b}$. Очевидно, что при этом $\pi(\overline{a}', \overline{b}') = \pi(\overline{a}', \overline{b})$.

3. На третьем шаге строим прямую l'', перпендикулярную плоскости $\pi(\overline{a}', \overline{b}')$, и раскладываем вектор $\overline{c}$ в сумму двух проекций: $\overline{c} = \text{пр}_{\pi(\overline{a}', \overline{b}')}\overline{c} + \text{пр}_{l''}\overline{c}$. В качестве вектора $\overline{c}'$ берем проекцию $\overline{c}$ на l'', т. е. $\overline{c}' = \overline{c} - \text{пр}_{\pi(\overline{a}', \overline{b}')}\overline{c}$ (рис. 41).

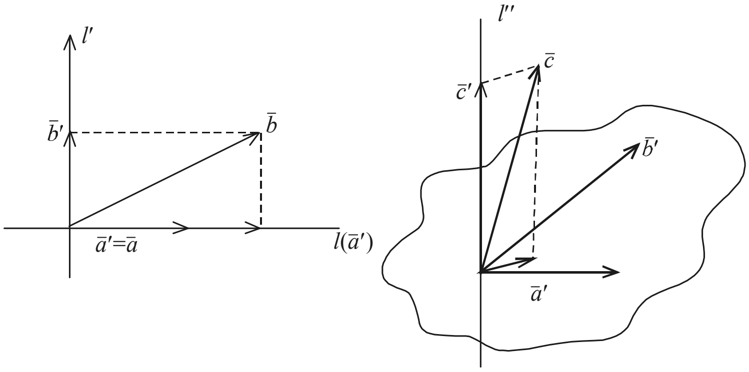


Рис. 41

Система векторов $\overline{a}'$, $\overline{b}'$, $\overline{c}'$ попарно ортогональна по построению.

9.3.4. Рассмотрим аналогичный процесс в произвольном евклидовом пространстве V. Пусть $b_1,..., b_n$ – произвольный базис. Будем «перерабатывать» этот базис так же, как в примере из V_3. На первом шаге в качестве b_1' возьмем элемент b_1. На втором шаге будем искать элемент b_2', удовлетворяющий условиям:

1) $b_2' = b_2 + \alpha_{12}b_1'$; здесь $\alpha_{12} \in \mathbf{R}$ – подходящее число;

2) $(b_2',\ b_1') = 0$.

Условие 2 означает следующее:

$$(b_2 + \alpha_{12}b_1',\ b_1') = \alpha_{12}(b_1',\ b_1') + (b_2,\ b_1') = 0,$$

откуда

$$\alpha_{12} = -\frac{(b_2,\ b_1')}{(b_1',\ b_1')}.$$

Таким образом, искомый элемент b_2' имеет вид

$$b_2' = b_2 - \frac{(b_2,\ b_1')}{(b_1',\ b_1')} b_1'.$$

Заметим, что при этом $L(b_1,\ b_2) = L(b_1',\ b_2')$.

Пусть мы уже построили элементы $b_1', ...,\ b_k'$, удовлетворяющие условиям:

а) $b_i' = b_i + \alpha_{1i} b_1' + ... + \alpha_{i-1\,i} b_{i-1}',\ \ i = 1, ..., k$;

б) $\left(b_i',\ b_j'\right) = 0$ при $i \neq j,\ \ i, j = 1, ..., k.$

И при этом $L(b_1, \ldots,\ b_i) = L(b_1', \ldots,\ b_i'),\ \ i = 1, ..., k.$

Построим элемент b_{k+1}', удовлетворяющий условиям

$$b_{k+1}' = b_{k+1} + \alpha_{1\,k+1} b_1' + ... + \alpha_{k\,k+1} b_k';$$

$$(b_{k+1}',\ b_i') = 0, \quad i = 1, ..., k.$$

Используя условие ортогональности, получаем

$$\begin{cases} \alpha_{1\,k+1}(b_1',\ b_1') + (b_{k+1},\ b_1') = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{k\,k+1}(b_k',\ b_k') + (b_{k+1},\ b_k') = 0. \end{cases}$$

Из полученной системы уравнений легко находим числа:

$$\alpha_{1\,k+1} = -\frac{(b_{k+1},\ b_1')}{(b_1',\ b_1')},\ \ \ldots,\ \ \alpha_{k\,k+1} = -\frac{(b_{k+1},\ b_k')}{(b_k',\ b_k')}$$

такие, что элемент b_{k+1}' удовлетворяет требуемым условиям а), б). Очевидно, что при этом $L(b_1, \ldots,\ b_{k+1}) = L(b_1', \ldots,\ b_{k+1}')$.

Проделав n шагов, получим систему элементов $b_1', ...,\ b_n'$ таких, что порождаемое ими подпространство совпадает с подпространством, порожденным векторами $b_1, ...,\ b_n$, т. е. со всем пространством V, и при этом они попарно ортогональны. Линейная независимость этой системы следует из п. 9.3.1. Таким образом, исходя из произвольного базиса $b_1, ...,\ b_n$ пространства V мы построили ортогональный базис $b_1',..., b_n'$. Описанный процесс называется ***процессом ортогонализации.***

Построенный базис можно еще немного «улучшить», если его элементы нормировать, т. е. сделать их длину равной единице. Это можно сделать, умножив каждый элемент на число, обратное его модулю, т. е. вместо $b_1',..., b_n'$ взять элементы $e_1,..., e_n$, где $e_i = \frac{1}{|b_i'|}\, b_i'$. Из приведенных рассуждений следует

Теорема. ***В*** n***-мерном евклидовом пространстве*** V ***существует ортонормированный базис*** $e_1,..., e_n$.

9.3.5. Матрица Грама–Шмидта. Пусть теперь $b_1,..., b_n$ – произвольный базис V. Так как скалярное произведение является билинейной формой, то можно говорить о матрице этой билинейной формы в базисе $b_1,..., b_n$. Элементами этой матрицы по определению служат числа $a_{ij} = (b_i,\ b_j)$, поэтому матрица выглядит так:

$$A = \begin{pmatrix} (b_1,b_1) & (b_1,b_2) & \cdots & (b_1,b_n) \\ (b_2,b_1) & (b_2,b_2) & \cdots & (b_2,b_n) \\ \dots & \dots & \dots & \dots \\ (b_n,b_1) & (b_n,b_2) & \cdots & (b_n,b_n) \end{pmatrix}.$$

Матрица такого вида называется ***матрицей Грама–Шмидта*** для базиса $b_1,..., b_n$. Заметим, что ***если базис ортогональный, то матрица Грама–Шмидта превращается в диагональную, а если базис ортонормированный, то – в единичную.***

Выясним геометрический смысл определителя матрицы A, также называемого ***определителем Грама–Шмидта.*** Обозначим его через G. Прежде всего заметим, что если от произвольного базиса $b_1,..., b_n$ перейти к ортогональному базису $b_1',..., b_n'$, построенному нами выше, определитель Грама–Шмидта не изменится. Действительно, так как $b_2' = b_2 + \alpha_1\, b_1$, то достаточно в определителе матрицы A ко второму столбцу прибавить первый, умноженный на α_1, и ко второй строке прибавить первую, умноженную на α_1, чтобы везде вместо b_2 появился b_2'. Это следует из свойств скалярного произведения (и определителя, см. свойство 8 в п. 2.2.2 и пп. 2.4.3–2.4.5), так как

$$\alpha_1 (b_i,\ b_1) + (b_i,\ b_2) = (b_i,\ \alpha_1 b_1 + b_2) = (b_i,\ b_2').$$

Продолжая этот процесс, заменим все векторы $b_1,..., b_n$ на $b_1',..., b_n'$. Тогда получим

$$G = \begin{vmatrix} (b_1',\ b_1') & 0 & \cdots & 0 \\ 0 & (b_2', b_2') & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & (b_n', b_n') \end{vmatrix} = |\, b_1' \,|^2 \,|\, b_2' \,|^2 \ldots |\, b_n' \,|^2 .$$

Что же собой представляет такое произведение? Чтобы это понять, обратимся к процессу ортогонализации.

Вернемся к примеру из V_3. Определитель Грама–Шмидта для векторов $\overline{a}$, $\overline{b}$, $\overline{c}$ будет равен

$$G = |\,\overline{a}\,|^2 |\,\overline{b}'|^2\, |\,\overline{c}'|^2 .$$

Каков геометрический смысл этого произведения? $|\,\overline{b}'\,|$ – это длина перпендикуляра, опущенного из конца вектора $\overline{b}$ на вектор $\overline{a}$ (см. рис. 41), другими словами, высота параллелограмма, построенного на векторах $\overline{a}$ и $\overline{b}$; $|\,\overline{c}'|$ – это длина перпендикуляра, опущенного из конца вектора $\overline{c}$ на плоскость векторов $\overline{a}$ и $\overline{b}$ (см. рис. 41), т. е. высота параллелепипеда, построенного на векторах $\overline{a}$, $\overline{b}$, $\overline{c}$. Теперь понятно, что $|\,\overline{a}\,|\,|\,\overline{b}'|$ – площадь параллелограмма, $|\,\overline{a}\,|\,|\,\overline{b}'|\,|\,\overline{c}'|$– объем параллелепипеда. Получили, что определитель Грама–Шмидта трех некомпланарных векторов пространства V_3 равен квадрату объема параллелепипеда, построенного на этих векторах.

Чтобы обобщить эту конструкцию на n-мерный случай, дадим индуктивное определение объема n-мерного параллелепипеда, т. е. параллелепипеда, построенного на элементах $b_1, \ldots, b_n$.

Обозначим его через v_n. Положим по определению $v_1 = |b_1|$ – «одномерный объем», т. е. длина вектора b_1;

$v_2 = v_1 |h_2|$ – «двумерный объем», т. е. площадь параллелограмма, построенного на элементах b_1, b_2;

$v_3 = v_2 |h_3|$ – «трехмерный объем», т. е. объем параллелепипеда, построенного на элементах b_1, b_2, b_3;

....

$v_n = v_{n-1} |h_n|$ – n-мерный объем, т. е. объем параллелеппеда, построенного на элементах $b_1, \ldots, b_n$.

Здесь $|h_i|$ – длина «перпендикуляра», опущенного из конца элемента b_i на подпространство, порожденное элементами $b_1, \ldots, b_{i-1}$. В процессе ортогонализации мы берем

$$b_i' = b_i + \alpha_{1i}\, b_1' + \ldots + \alpha_{i-1\,i}\, b_{i-1}',$$

где $b_i' \perp b_1, \ldots, b_i' \perp b_{i-1}$, т. е. $b_i' \perp L(b_1', \ldots, b_{i-1}') = L_{i-1}$ – элемент b_i' перпендикулярен подпространству, порожденному элементами $b_1', \ldots, b_{i-1}'$. Обозначим $\alpha_{1i}\, b_1' + \ldots + \alpha_{i-1\,i}\, b_{i-1}' = a_{i-1}$. Тогда $b_i = b_i' - a_{i-1}$. В этом разложении элемента b_i слагаемое a_{i-1} принадлежит пространству L_{i-1}, а слагаемое b_i' ему перпендикулярно. Вот почему b_i' играет роль перпендикуляра, опущенного из конца элемента b_i на подпространство L_{i-1}, т. е. $|h_i| = |b_i'|$.

Теперь понятно, что произведение $|b_1'|^2 |b_2'|^2 \ldots |b_n'|^2$ и тем самым определитель Грама–Шмидта есть квадрат объема n-мерного параллелепипеда, построенного на ребрах $b_1, \ldots, b_n$.

9.3.6. Отметим некоторые очевидные свойства определителя Грама–Шмидта.

1. Справедливо неравенство

$$0 \le G \le |b_1|^2 \ldots |b_n|^2.$$

Оно следует из разложений п. 9.3.5, так как по теореме Пифагора имеем $|b_i|^2 = |b_i'|^2 + |a_{i-1}|^2$ и $|b_i'|^2 \le |b_i|^2$.

2. Равенство $G = |b_1|^2 \ldots |b_n|^2$ выполняется тогда и только тогда, когда элементы $b_1, \ldots, b_n$ попарно ортогональны. Это следует из того, что в нашем случае $|b_i'| = |b_i|$ и $|a_{i-1}| = 0$ для всех $i = 2, \ldots, n$, а это и означает, что $b_1, \ldots, b_n$ попарно ортогональны.

3. Обращение в ноль определителя $G = 0$ равносильно тому, что элементы $b_1, \ldots, b_n$ линейно зависимы. Это вытекает из того, что равенство $|b_i'| = 0$ верно только при $b_i + \alpha_{1i} b_1' + \ldots + \alpha_{i-1\,i} b_{i-1}' = \underline{0}$ и, следовательно, элементы $b_1, \ldots, b_n$ линейно зависимы.

9.3.7. Изоморфизм евклидовых пространств. В заключение этого раздела ознакомимся с изоморфизмом евклидовых пространств. По сравнению с определением изоморфизма линейных пространств добавляется требование о сохранении скалярного произведения.

Определение. Два евклидовых пространства *V* и *W* называются ***изоморфными***, если существует изоморфизм $\varphi: V \to W$ линейных пространств *V* и *W*, который удовлетворяет условию: для всякой пары элементов $x, y \in V$

$$(x, y) = (\varphi(x), \ \varphi(y)).$$

Образно говоря, изоморфизм евклидовых пространств означает одинаковость «масштаба измерений» в них.

Мы знаем, что два линейных пространства одинаковой размерности *n* над полем **F** изоморфны хотя бы потому, что каждое из них изоморфно $\mathbf{F}^n$ (п. 3.3.8). Оказывается, что два евклидовых пространства одинаковой размерности тоже изоморфны.

Теорема*. Если V и W – евклидовы пространства одинаковой размерности n, то они изоморфны.*

Для доказательства воспользуемся теоремой, сформулированной в конце п. 9.3.4. Пусть $e_1, \ldots, e_n$ – ортонормированный базис пространства *V*, $e_1', \ldots, e_n'$ – ортонормированный базис пространства *W*. Тогда положим

$$\varphi(x) = \varphi(x_1 e_1 + \ldots + x_n e_n) = x_1 e_1' + \ldots + x_n e_n'.$$

При таком определении отображения имеем:

1) $\varphi(x+y)=\varphi(x)+\varphi(y)$;

2) $\varphi(\alpha x)=\alpha\varphi(x), \quad \alpha\in\mathbf{R}$;

3) $(\varphi(x),\ \varphi(y))=\left(\sum\limits_{i=1}^{n}x_i e'_i,\ \sum\limits_{j=1}^{n}y_j e'_j\right)=$

$$=\sum_{i=1}^{n}\ \sum_{j=1}^{n}x_i y_j\left(e'_i,\ e'_j\right)=\sum_{i=1}^{n}x_i y_i=(x,\ y);$$

4) $\operatorname{Ker}\varphi=\{x\in V \mid \varphi(x)=\underline{0}\}=\{\underline{0}\}$.

Из этих условий следует, что φ – изоморфизм евклидовых пространств V и W.

Теорема доказана.

9.4. ОРТОГОНАЛЬНЫЕ И САМОСОПРЯЖЕННЫЕ ОПЕРАТОРЫ

9.4.1. Матрицы ортогональных и самосопряженных операторов. В главе 8 мы рассмотрели линейные операторы, действующие на линейных пространствах. Понятно, что можно рассматривать линейные операторы, действующие на евклидовых пространствах. В связи с тем, что в евклидовых пространствах определено скалярное произведение, появляются новые свойства линейных операторов.

Пусть V – евклидово пространство, $e_1,\dots,\ e_n$ – ортонормированный базис, $\varphi\colon\ V\to V$ – линейный оператор, действующий на нем. В п. 8.1.8 мы заметили, что линейный оператор φ на пространстве V задает изоморфное отображение тогда и только тогда, когда $\operatorname{Ker}\varphi=\{\underline{0}\}$. В случае евклидовых пространств, для того чтобы оператор φ задавал изоморфное отображение линейного пространства, этого мало. Необходимо еще сохранять метрику пространства, т. е. не изменять скалярного произведения.

Определение. Линейный оператор $\varphi\colon\ V\to V$ называется ***ортогональным***, если следующее условие выполняется для всех $x,\ y\in V$:

$$(\varphi(x),\ \varphi(y))=(x,\ y).$$

Посмотрим, как устроена матрица ортогонального оператора. Прежде всего понятно, что ортогональный оператор переводит ортонормированный базис в ортонормированный. Пусть $[\varphi]$ – матрица оператора φ в базисе $e_1,..., e_n$. Столбцы этой матрицы суть координаты элементов $\varphi(e_1),..., \varphi(e_n)$ в базисе $\overline{e}=(e_1,..., e_n)$ (см. пп. 3.3.12, 3.3.13). Так как оба базиса $\overline{e}$ и $\varphi(\overline{e})$ ортонормированы, то скалярное произведение равно сумме произведений соответствующих координат, т. е. в этом случае в матрице $[\varphi]$ сумма произведений элементов *i*-го столбца на элементы *j*-го столбца равна нулю при $i \neq j$ и единице при $i = j$. Это означает, что выполняется замечательное равенство

$$[\varphi]^{\mathrm{T}}[\varphi]=E.$$

Определение. Матрица *A,* удовлетворяющая условию $AA^{\mathrm{T}}=E$, называется ***ортогональной***.

И мы только что доказали, что матрица ортогонального оператора в ортонормированном базисе является ортогональной.

Определение. Оператор $\varphi^*\colon V \to V$ называется ***сопряженным*** к оператору $\varphi\colon V \to V$, если для всех $x, y \in V$ выполняется равенство

$$(\varphi(x),\ y)=(x,\ \varphi^*(y)).$$

Перепишем это равенство на языке матриц. Вспомним, что в обозначениях 3.3.12

$$\overset{|}{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \overset{|}{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \overset{|}{x}^{\mathrm{T}} = \overline{x},\ (x,\ y)=\overline{x}\,\overset{|}{y}.$$

$$(\varphi(x),\ y)=([\varphi]\overset{|}{x})^{\mathrm{T}}\,\overset{|}{y}=\overline{x}\ [\varphi]^{\mathrm{T}}\,\overset{|}{y}, \quad (\overset{|}{x},\ \varphi^*(y))=\overline{x}[\varphi^*]\overset{|}{y}.$$

Так как $(\varphi(x),\ y)-(x,\ \varphi^*(y))=0$, то получаем

$$\overline{x}[\varphi]^{\mathrm{T}}\,\overset{|}{y}-\overline{x}[\varphi^*]\overset{|}{y}=\overline{x}([\varphi]^{\mathrm{T}}-[\varphi^*])\overset{|}{y}=0.$$

Последнее равенство справедливо для любых x и y, поэтому из него следует, что $[\varphi]^{\mathrm{T}} - [\varphi^*] = (0)$, т. е.

$$[\varphi^*] = [\varphi]^{\mathrm{T}}.$$

Это равенство означает, что ***в ортонормированном базисе матрица сопряженного оператора*** φ^* ***получается из матрицы оператора*** φ ***транспонировани*ем.

Определение. Оператор φ называется ***самосопряженным***, если $\varphi = \varphi^*$.

Теперь понятно, что ***матрица самосопряженного оператора в ортонормированном базисе является симметрической***, так как для самосопряженного оператора $[\varphi] = [\varphi]^{\mathrm{T}}$.

9.4.4. Спектральные свойства матриц ортогональных и самосопряженных операторов. Остановимся на свойствах характеристических чисел ортогональных и симметрических матриц.

Пусть Q – ортогональная матрица, тогда по определению $QQ^{\mathrm{T}} = E$, откуда $Q^{-1} = Q^{\mathrm{T}}$. Кроме того, так как для определителей верно равенство $|Q^{\mathrm{T}}| = |Q|$ (см. 2.4.1), то по теореме 2.5.6 $|Q|^2 = 1$, т. е. $|Q| = \pm 1$.

Матрица Q^{-1} ***тоже является ортогональной***, так как в соответствии с задачей 2 п. 2.1.8 выполняется цепочка равенств

$$(Q^{-1})(Q^{-1})^{\mathrm{T}} = (Q^{-1})(Q^{\mathrm{T}})^{\mathrm{T}} = Q^{-1}Q = E.$$

Если Q и S – ортогональные матрицы, то матрица QS ***– тоже ортогональная***, так как

$$(QS)(QS)^{\mathrm{T}} = QSS^{\mathrm{T}}Q^{\mathrm{T}} = E.$$

Рассмотрим спектральные свойства ортогональных и симметрических матриц.

Теорема. ***Характеристические числа ортогональной матрицы по модулю равны единице.***

Пусть λ – характеристическое число ортогональной матрицы Q, которое в соответствии с п. 8.3.8 будет характеристическим числом матрицы $Q^{\mathrm{T}} = Q^{-1}$, и тогда $\lambda = 1/\lambda$, $\lambda^2 = 1$ и $|\lambda| = 1$.

$$(x,\ x) = (Q^{-1}Qx,\ x) = (Q^{\mathrm{T}}(Qx),\ x) = (Qx,\ Qx) = \lambda^2(x,\ x),$$

откуда $\lambda^2 = 1$ и $|\lambda| = 1$, что и требовалось доказать.

9.4.5. Перейдем к симметрическим матрицам. Прежде всего, для них справедлива

Теорема. ***Все характеристические числа симметрических матриц с действительными элементами действительны.***

Пусть A – симметрическая матрица и λ – ее характеристическое число (может быть, комплексное). Тогда однородная система уравнений $(A - \lambda_0 E)\overset{|}{x} = (0)$ имеет ненулевое решение (которое может оказаться комплексным, см. 8.3.2).

Пусть $(\beta_1, ..., \beta_n)$ – такое решение. Тогда

$$\begin{cases} a_{11}\beta_1 + a_{12}\beta_2 + ... + a_{1n}\beta_n = \lambda\beta_1, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\beta_1 + a_{n2}\beta_2 + ... + a_{nn}\beta_n = \lambda\beta_n. \end{cases}$$

Умножим i-е уравнение этой системы на $\overline{\beta}_i$ – комплексно-сопряженное к β_i – и сложим все уравнения системы после такого домножения:

$$\sum_{i=1}^{n} a_{i1}\,\beta_1\,\overline{\beta}_i + \sum_{i=1}^{n} a_{i2}\,\beta_2\,\overline{\beta}_i + ... + \sum_{i=1}^{n} a_{in}\,\beta_n\,\overline{\beta}_i = \lambda\sum_{i=1}^{n}\beta_i\,\overline{\beta}_i\ .$$

Докажем, что левая часть равенства – действительное число. Это легко сделать, если заметить, что комплексно-сопряженное к этому числу равно ему самому. Посчитаем:

$$\overline{\sum_{i=1}^{n} a_{i1}\,\beta_1\,\overline{\beta}_i + ... + \sum_{i=1}^{n} a_{in}\,\beta_n\,\overline{\beta}_i} = \sum_{i=1}^{n} a_{i1}\,\overline{\beta}_1\,\overline{\overline{\beta}}_i + ... + \sum_{i=1}^{n} a_{in}\,\overline{\beta}_n\overline{\overline{\beta}}_i =$$

$$=\sum_{i=1}^{n} a_{i1}\,\overline{\beta}_1\,\beta_i+\ldots+\sum_{i=1}^{n} a_{in}\,\overline{\beta}_n\beta_i=\sum_{i=1}^{n} a_{1i}\,\overline{\beta}_1\,\beta_i+\ldots+\sum_{i=1}^{n} a_{ni}\,\overline{\beta}_n\beta_i\,.$$

(Мы воспользовались равенствами $a_{ij}=a_{ji}$.) Перепишем и приведем подобные при $\beta_1,\ \beta_2,\ldots,\ \beta_n$. Получается такое выражение:

$$\sum_{i=1}^{n} a_{i1}\,\overline{\beta}_i\,\beta_1+\sum_{i=1}^{n} a_{i2}\,\overline{\beta}_i\,\beta_2+\ldots+\sum_{i=1}^{n} a_{in}\overline{\beta}_i\,\beta_n\,.$$

Сравнивая с исходным выражением, видим, что получилось оно же, только без черты, что и доказывает наше утверждение о действительности левой части приведенного выше равенства. Поскольку $(\beta_1,\ldots,\ \beta_n)$ – ненулевое решение,

$$\sum_{i=1}^{n}\beta_i\,\overline{\beta}_i\ =\ \sum_{i=1}^{n}|\,\beta_i\,|^2\neq 0\,.$$

Значит, мы можем выразить λ:

$$\lambda\ =\frac{\sum\limits_{i=1}^{n} a_{i1}\,\beta_1\,\overline{\beta}_i+\ldots+\sum\limits_{i=1}^{n} a_{in}\,\beta_n\,\overline{\beta}_i}{\sum\limits_{i=1}^{n}|\,\beta_i\,|^2}\,.$$

Так как числитель и знаменатель – действительные числа, то и число λ действительное, что и доказывает теорему.

9.4.6. Собственный базис симметрической матрицы. Для доказательства теоремы 9.4.7 нам понадобится

Лемма. ***Для любого ненулевого элемента*** $x\in V$ ***справедливо разложение пространства V в прямую сумму подпространств:***

$$V=L_x\oplus L_x^{\perp}\,,$$

где $L_x=\{\alpha x/\alpha\in R\}$, $L_x^{\perp}=\{y\in V\,|\,(x,\ y)=0\}$.

Вначале заметим, что $L_x^{\perp}$ – это подпространство.

1. Пусть $y,\ z\in L_x^{\perp}$, тогда $(x,\ y)=0$, $(x,\ z)=0$ и, следовательно, $(x,\ y+z)=0$, откуда $y+z\in L_x^{\perp}$.

2. Пусть $y \in L_x^{\perp}$, тогда $(x,\ y) = 0$ и для любого действительного числа α имеем: $\alpha\,(x,\ y) = (x,\ \alpha\, y) = 0$, откуда $\alpha\, y \in L_x^{\perp}$.

Очевидно, что $L_x \cap L_x^{\perp} = \{\underline{0}\}$, так как если $v \in L_x \cap L_x^{\perp}$, то $(v,\ v) = 0$ и $v = \underline{0}$. Наконец, из разд. 9.3.4, 9.3.5 понятно, что справедливо равенство $V = L_x + L_x^{\perp}$, так как для всякого $y \in V$ найдется действительное число γ такое, что $y = \gamma x + y_1$, где $y_1 \in L_x^{\perp}$. Тем самым мы доказали, что $V = L_x \oplus L_x^{\perp}$.

По следствию п. 8.4.2 получаем

$$\dim V = \dim L_x + \dim L_x^{\perp}, \text{ откуда } \dim L_x^{\perp} = n - 1.$$

9.4.7. Теперь мы можем доказать очень важную теорему о собственных элементах действительной симметрической матрицы.

Теорема. ***Для всякой действительной симметрической матрицы в $\mathbf{R}^n$ существует ортонормированный базис, состоящий из собственных элементов этой матрицы.***

Доказательство проведем индукцией по размерности матрицы. При n = 1 утверждение очевидно. Пусть теорема справедлива для матриц размерности не выше $n-1$. Докажем ее для матриц размерности n.

Пусть A – симметрическая матрица, λ_1 – ее характеристическое число. Так как по теореме 9.4.5 $\lambda_1 \in \mathbf{R}$, то в пространстве $\mathbf{R}^n$ найдется действительный собственный элемент $x \neq \underline{0}$, соответствующий числу λ_1. Тогда по лемме 9.4.6 получим равенства $\mathbf{R}^n = L_x \oplus L_x^{\perp}$ и $\dim L_x^{\perp} = n - 1$. Покажем, что $L_x^{\perp}$ инвариантно относительно линейного оператора φ_A, т. е. для любого элемента $y \in L_x^{\perp}$ выполняется включение $\varphi_A(y) \in L_x^{\perp}$. Если $y \in L_x^{\perp}$, то $(x,\ y) = 0$. Фиксируем в $\mathbf{R}^n$ ортонормированный базис. Тогда в силу симметричности матрицы A $\varphi_A = \varphi_A^*$, т. е. оператор φ_A – самосопряженный. Итак,

$$(x,\ \varphi_A(y)) = \left(\varphi_A^*(x),\ y\right) = (\varphi_A(x),\ y) = (\lambda_1 x,\ y) = \lambda_1(x,\ y) = 0.$$

Значит, выполнено $\varphi_A(y) \in L_x^{\perp}$, т. е. пространство $\mathbf{R}^n$ разложено в прямую сумму подпространств, инвариантных относительно оператора φ_A. В п. 8.4.4 показано, что в этом случае матрица $[\varphi_A]$ в базисе, составленном из базисов L_x и $L_x^{\perp}$, имеет клеточно-диагональный вид:

$$[\varphi_A] = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right),$$

где B является матрицей самосопряженного оператора φ_A, действующего на $L_x^{\perp}$ и «суженного» до этого подпространства. Поэтому в любом ортонормированном базисе пространства $L_x^{\perp}$ матрица B является симметрической матрицей размерности $n-1$.

По предположению индукции в подпространстве $L_x^{\perp} \cong \mathbf{R}^{n-1}$ существует ортонормированный базис $(e_2, ..., e_n)$ из собственных элементов матрицы B, в котором она имеет диагональный вид. Если к этому базису присоединить элемент $e_1 = \dfrac{1}{|x|}x$, то получим ортонормированный базис $\bar{e} = (e_1, ..., e_n)$ пространства $\mathbf{R}^n$, состоящий из собственных элементов матрицы A, в котором матрица A имеет диагональный вид:

$$A_e = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Теорема доказана.

9.5. ПРИВЕДЕНИЕ КВАДРАТИЧНОЙ ФОРМЫ К КАНОНИЧЕСКОМУ ВИДУ

9.5.1. Преобразование матрицы квадратичной формы при переходе к новому базису. Так как квадратичная форма получается из билинейной заменой элемента y на элемент x, то для квадратичной формы получаем:

$$Q(x) = A(x,\ x) = \overline{x} A \overset{|}{x},$$

где $\overset{|}{x}$ – столбец коэффициентов элемента x в некотором базисе, где форма $Q(x)$ имеет матрицу A.

Посмотрим, как изменится матрица A при переходе к новому базису.

Пусть S – матрица перехода от старого базиса к новому, тогда $\overset{|}{x} = S^{-1} \overset{|}{x}'$, и мы получаем, что

$$Q(x) = \overline{x} A \overset{|}{x} = (S^{-1} \overset{|}{x}')^T A S^{-1} \overset{|}{x}' = \overline{x}'(S^{-1})^T A S^{-1} \overset{|}{x}'.$$

Отсюда видно, что если

$$A' = (S^{-1})^T A S^{-1},$$

то A' – матрица формы $Q(x)$ в новом базисе.

9.5.2. Приведение квадратичной формы к каноническому виду. Прежде чем продолжить изучение квадратичных форм, поговорим об их геометрическом смысле. Рассмотрим квадратичную форму в $\mathbf{R}^2$ в некотором фиксированном базисе

$$A(x,\ x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2.$$

Если перейти к привычным обозначениям координат буквами x, y и заменить коэффициенты большими буквами латинского алфавита, получим $Ax^2 + Bxy + Cy^2$. Если это выражение приравнять к какому-либо числу, получим уравнение линии второго порядка на плоскости. В школьном курсе вы сталкивались с уравнением окружности

$$x^2 + y^2 = R^2.$$

Из школьного же курса можно было понять, что уравнение любой линии зависит от системы координат. Например, интуитивно понятно, что уравнения одного и того же эллипса в системе координат $Ox'y'$ может оказаться «лучше», чем в системе Oxy (рис. 42).

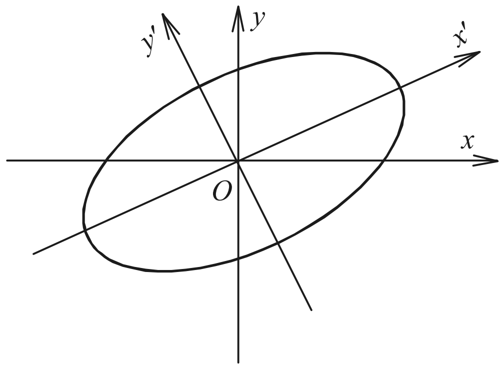


Рис. 42

А что такое смена системы координат при неизменном начале? Это замена базиса. Значит, могут найтись базисы, в которых уравнение той же самой линии выглядит проще. Конечно, нужно уметь находить такие базисы. С этой целью мы теперь вернемся к квадратичным формам и рассмотрим так называемую задачу о приведении квадратичной формы к каноническому виду.

Определение. Квадратичная форма $Q(x)$ имеет ***канонический вид*** в базисе $\overline{e}$, если

$$Q(x) = A(x, x) = a_{11}x_1^2 + a_{22}x_2^2 + \dots + a_{nn}x_n^2,$$

когда $x = x_1e_1 + \dots + x_ne_n$.

Таким образом, в каноническом виде формы $Q(x)$ все коэффициенты при произведениях различных неизвестных равны нулю. Ясно, что матрица квадратичной формы канонического вида – диагональная. Следовательно, ***привести форму к каноническому виду означает найти такой базис в пространстве*** $\mathbf{R}^n$***, в котором ее матрица примет диагональный вид.***

Для дальнейшего нам понадобится следующее утверждение.

Теорема. ***Матрица перехода от одного ортонормированного базиса к другому такому же базису является ортогональной***.

Пусть $\overline{e} = (e_1, \dots, e_n)$, $\overline{e}' = (e_1', \dots, e_n')$ – ортонормированные базисы, S – матрица перехода от $\overline{e}$ к $\overline{e}'$. Тогда

$$\overset{|}{e}\,\overline{e} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} (e_1, \dots, e_n) = \begin{pmatrix} (e_1, e_1) & 0 & \dots & 0 \\ 0 & (e_2, e_2) & & \\ \vdots & & \ddots & 0 \\ 0 & & 0 & (e_n, e_n) \end{pmatrix} = E\,.$$

Аналогично $\overset{|}{e}'\overline{e}' = E$. А поскольку $\overline{e} = \overline{e}'S$, получаем

$$\overset{|}{e}\,\overline{e} = S^{\mathrm{T}}\,\overset{|}{e}'\overline{e}'S = S^{\mathrm{T}}ES = S^{\mathrm{T}}S = E, \text{ что и требовалось.}$$

Мы договорились (см. п. 9.1.5), что будем считать матрицу A квадратичной формы в любом базисе симметрической. Тогда по теореме 9.4.7 существует ортонормированный базис, в котором матрица A имеет диагональный вид. Пусть S – матрица перехода от ортонормированного базиса $\overline{e} = (e_1, \dots, e_n)$ к ортонормированному базису $\overline{e}' = (e_1', \dots, e_n')$, составленному из собственных векторов матрицы A. Сделаем преобразование координат с матрицей S^{-1}: $\overset{|}{x} = S^{-1}\overset{|}{x}'$. Тогда по формуле п. 9.5.1 матрица A квадратичной формы будет иметь вид $A' = (S^{-1})^{\mathrm{T}} A S^{-1}$. Но вследствие ортонормированности обоих базисов матрица S ортогональна: $S^{\mathrm{T}} = S^{-1}$, и тогда

$$A' = SA\,S^{-1}\,.$$

Последнее равенство соответствует формуле п. 8.2.1 и означает, что матрица A' имеет диагональный вид как матрица линейного оператора в базисе из собственных элементов.

Таким образом, нами доказана следующая

Теорема. ***Всякая действительная квадратичная форма некоторым ортогональным преобразованием координат может быть приведена к каноническому виду.***

Можно заметить, что существует множество различных ортогональных матриц S, приводящих матрицу A квадратичной формы к диагональному виду, но сам этот вид при любом ортогональном преобразовании определен однозначно с точностью до порядка диагональных элементов, так как на диагонали стоят собственные числа матрицы A, которые и будут коэффициентами канонического вида квадратичной формы. Это следует из того, что подобные матрицы имеют одинаковые характеристические многочлены (8.3.5)

9.5.3. Пример. Привести к каноническому виду квадратичную форму, определенную в некотором базисе как

$$Q(x) = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4 .$$

Составим матрицу этой формы:

$$A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

Найдем характеристический многочлен этой матрицы:

$$|A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix} = (\lambda - 1)^3 (\lambda + 3) .$$

Характеристическими числами будут

$$\lambda_1 = \lambda_2 = \lambda_3 = 1 ; \quad \lambda_4 = -3 .$$

Следовательно, $Q(x)$ приводится к следующему каноническому виду:

$$Q(x) = y_1^2 + y_2^2 + y_3^2 - 3y_4^2 .$$

ГЛАВА 10

КРИВЫЕ И ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА

10.1. КАНОНИЧЕСКИЕ УРАВНЕНИЯ, ГЕОМЕТРИЧЕСКИЕ И ОПТИЧЕСКИЕ СВОЙСТВА КРИВЫХ ВТОРОГО ПОРЯДКА

10.1.1. Общее уравнение кривой второго порядка. В главе 5 вы ознакомились с линейными объектами: прямыми, плоскостями. Уравнения этих объектов содержали неизвестные только в первой степени, такие уравнения в алгебре принято называть линейными. В главе 10 мы ознакомимся с объектами, уравнения которых содержат неизвестные во второй степени, отсюда и общее название этих объектов. Рассмотрим вначале линии второго порядка на плоскости.

Нетрудно понять, что общее уравнение такой линии имеет вид

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Несмотря на кажущееся обилие возможностей (шесть различных действительных коэффициентов!), помимо линейных геометрических объектов и точек, такое уравнение может описывать только три вида ***кривых второго порядка: эллипс, гиперболу и параболу***. На ***окружность*** чаще всего смотрят как на частный случай эллипса. И в школе, кроме нее, знакомятся с частными случаями параболы и гиперболы. Существует много различных подходов к рассмотрению кривых второго порядка. Наш подход схематически выглядит следующим образом.

1. Определение линии как геометрического места точек, удовлетворяющих некоторому свойству.

2. Выбор канонической системы координат, где это свойство принимает наиболее простой вид соотношений между переменными.

3. Вывод канонического уравнения.

После такого краткого описания всех линий будет получено универсальное определение эллипса, гиперболы и параболы (исключая окружность!), а также будут рассмотрены их оптические свойства и очень сжато рассказано о теории конических сечений.

10.1.2. Окружность

Определение. ***Окружностью*** называется геометрическое место точек плоскости, равноудаленных от одной данной точки, именуемой ***центром***.

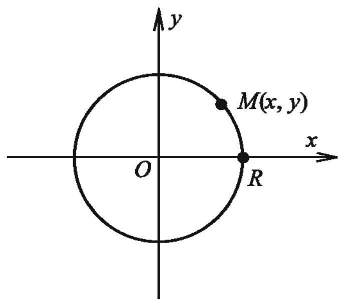


Рис. 43

Каноническую систему координат естественно выбрать так, чтобы начало координат совпадало с центром окружности (рис. 43).

Пусть все точки окружности удалены от центра на расстояние R и пусть $M(x;\, y)$ – произвольная ее точка. Тогда $|OM| = R$ или $\sqrt{x^2 + y^2} = R$, или $x^2 + y^2 = R^2$. Последнее уравнение и называется ***каноническим уравнением*** окружности.

10.1.3. Эллипс

Определение. ***Эллипсом*** называется геометрическое место точек, сумма расстояний которых до двух данных точек, именуемых ***фокусами***, постоянна и равна $2a$.

Фокусы эллипса принято обозначать как F_1 и F_2, а расстояние между ними равно $F_1F_2 = 2c$. Разумеется, мы будем считать, что $a > c > 0$.

Каноническую систему координат выберем так: за ось Ox примем прямую, проходящую через фокусы, а ось Oy проведем через середину отрезка F_1F_2 (рис. 44).

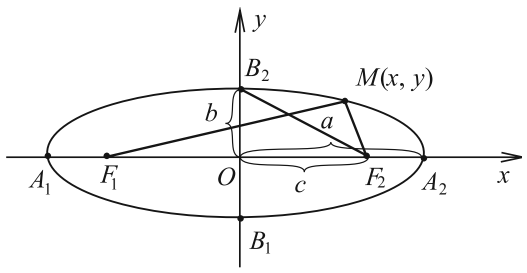


Рис. 44

10.1.4. Каноническое уравнение эллипса. Пусть точка $M(x;\, y)$ – произвольная точка эллипса. Тогда по определению $|\overline{F_1M}| + |\overline{F_2M}| = 2a$. Чтобы получить каноническое уравнение эллипса, нужно последнее равенство расписать через координаты и сделать некоторые алгебраические преобразования. Понятно, что координаты фокусов будут $F_1(-c;\ 0),\ \ F_2(c;\ 0)$, а координаты векторов $\overline{F_1M} = (x+c;\, y),\ \ \overline{F_2M} = (x-c;\, y)$. Тогда

$$\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2}=2a,$$

$$\sqrt{(x+c)^2+y^2}=2a-\sqrt{(x-c)^2+y^2}\,.$$

Возведем обе части равенства в квадрат и приведем подобные члены в обеих частях равенства:

$$(x+c)^2+y^2=4a^2-4a\sqrt{(x-c)^2+y^2}+(x-c)^2+y^2,$$

$$x^2+2cx+c^2=4a^2-4a\sqrt{(x-c)^2+y^2}+x^2-2cx+c^2.$$

Уединим квадратный корень и опять возведем в квадрат обе части равенства:

$$-4a\sqrt{(x-c)^2+y^2}=4cx-4a^2,$$

$$a^2((x-c)^2+y^2)=c^2x^2-2a^2cx+a^4,$$

$$a^2(x^2-2cx+c^2+y^2)=c^2x^2-2a^2cx+a^4,$$

$$a^2x^2-2a^2cx+a^2c^2+a^2y^2=c^2x^2-2a^2cx+a^4,$$

$$x^2(a^2-c^2)+a^2y^2=a^4-a^2c^2,$$

$$x^2(a^2-c^2)+a^2y^2=a^2(a^2-c^2).$$

Обозначим $a^2-c^2=b^2$ и разделим обе части последнего равенства на выражение a^2b^2:

$$\frac{x^2b^2}{a^2b^2}+\frac{a^2y^2}{a^2b^2}=1.$$

Окончательно получили

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1.$$

Это уравнение и называется ***каноническим уравнением эллипса***.

10.1.5. Геометрические свойства эллипса. Мы видим, что уравнение эллипса не меняется при замене x на $(-x)$ и y на $(-y)$. Это говорит о том, что ***оси координат являются осями симметрии эллипса, а начало координат – центром симметрии эллипса***. Прямоугольник, определяемый неравенствами $|x|\le a$, $|y|\le b$, называется ***основным прямоугольником*** эллипса. Эллипс целиком содержится в этом прямоугольнике.

Выясним геометрический смысл параметров a и b. Положив $y=0$ и подставив в каноническое уравнение эллипса, мы тем самым найдем координаты точек A_1 и A_2 пересечения эллипса с осью Ox (рис. 44):

$$\frac{x^2}{a^2}=1, \quad x^2=a^2, \quad x=\pm\ a,$$

т. е. $A_1(-a;0)$, $A_2(a;0)$.

Таким образом,

$$|\overline{A_1A_2}| = 2a.$$

Аналогично при $x = 0$ получаем координаты точек B_1 и B_2 пересечения эллипса с осью Oy:

$$\frac{y^2}{b^2} = 1; \quad y^2 = b^2; \quad y = \pm\ b,$$

т. е. $B_1(-b;\ 0)$, $B_2(b;\ 0)$, $|\overline{B_1 B_2}| = 2b$.

Теперь становятся понятными названия параметров эллипса:

a – ***большая полуось***;

b – ***малая полуось***;

c – ***половина фокусного расстояния.***

Точки A_1, A_2, B_1, B_2 называются ***вершинами эллипса***.

На рис. 44 видно, что введенное формально в 10.1.4 равенство $b^2 = a^2 - c^2$ есть не что иное, как теорема Пифагора, записанная для треугольника OB_2F_2, ведь в силу симметрии эллипса относительно оси Oy $B_2F_1 = B_2F_2$, а по определению $B_2F_1 + B_2F_2 = 2a$, откуда $B_2F_2 = a$.

Выше мы отмечали, что окружность является частным случаем эллипса. Теперь понятно, что эллипс превращается в окружность при $a = b$ или, что то же самое, при $c = 0$. Последнее означает, что оба фокуса «сливаются» в одну точку, которая и является центром окружности.

Определение. ***Эксцентриситетом*** эллипса называется число $\varepsilon = \dfrac{c}{a}$.

Сразу заметим, что для эллипса $0 < \varepsilon < 1$. На эксцентриситет удобно смотреть, как на меру «сплющенности» эллипса: чем больше эксцентриситет, тем больше «сплющен» эллипс, и наоборот.

Определение. Расстояния r_1 и r_2 от произвольной точки эллипса до его фокусов называются ***фокальными радиусами*** этой точки, здесь индексы соответствуют фокусам:

$$MF_1 = r_1; \quad MF_2 = r_2.$$

Получим удобные формулы для нахождения радиусов r_1 и r_1. Вначале заметим, что если точка $M(x;\, y)$ лежит на эллипсе, то из канонического уравнения $y^2 = b^2 - \frac{b^2}{a^2}x^2$. Теперь найдем r_1:

$$r_1 = MF_1 = \sqrt{(x+c)^2 + y^2} = \sqrt{x^2 + 2cx + c^2 + b^2 - \frac{b^2}{a^2}x^2} =$$

$$= \sqrt{\left(1 - \frac{b^2}{a^2}\right)x^2 + 2cx + a^2} = \sqrt{\frac{c^2}{a^2}x^2 + 2cx + a^2} =$$

$$= \sqrt{\left(\frac{c}{a}x + a\right)^2} = \frac{c}{a}x + a = \varepsilon x + a.$$

Объясним, почему при извлечении корня можно опустить знак модуля. Это следует из того, что $\varepsilon x\ < |x\ | \le a$, откуда $a \pm \varepsilon x \ge 0$. Аналогично получаем равенство $r_2 = MF_2 = a - \varepsilon x$. Очевидно, что выполняется соотношение $r_1 + r_2 = a + \varepsilon x + a - \varepsilon x = 2a$.

С эллипсом связаны еще две замечательные прямые.

Определение. ***Директрисами*** эллипса называются прямые, перпендикулярные его фокальной оси и обладающие тем свойством, что отношение расстояний от любой точки эллипса до фокуса и до соответствующей директрисы есть величина постоянная, равная эксцентриситету.

Найдем уравнения директрис. Сразу заметим, что директрисы не могут пересечь эллипс, так как в этом случае для точки пересечения отношение таких расстояний равняется нулю, а по определению $\varepsilon > 0$.

Будем искать правую директрису (левая ей симметрична). Всякая прямая, параллельная оси *Oy*, имеет уравнение $x = x_0$, причем $x_0 > 0$ из-за только что сделанного замечания. Теперь $d_2 = x_0 - x$ (рис. 45). По определению директрисы имеем

$$\frac{r_2}{d_2} = \frac{a - \varepsilon x}{x_0 - x} = \varepsilon \quad \text{или} \quad a - \varepsilon x = \varepsilon x_0 - \varepsilon x.$$

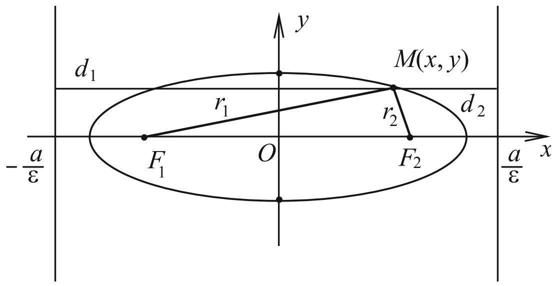


Рис. 45

Отсюда $x_0 = \dfrac{a}{\varepsilon}$. Таким образом, ***уравнения директрис*** имеют вид

$$x = \pm\frac{a}{\varepsilon}.$$

Пусть теперь $M(x;\ y)$ – произвольная точка и d – ее расстояние до директрисы $x = \dfrac{a}{\varepsilon}$, а r – ее расстояние до фокуса F_2. Докажем, что если верно равенство $\dfrac{r}{d} = \varepsilon$, то точка M лежит на эллипсе. Действительно, если $\dfrac{\sqrt{(x-c)^2 + y^2}}{\left| x - \dfrac{a}{\varepsilon} \right|} = \varepsilon$, то, возведя в квадрат обе части равенства, получим

$$\frac{(x-c)^2 + y^2}{\left(x - \dfrac{a}{\varepsilon}\right)^2} = \varepsilon^2,$$

или

$$x^2 - 2cx + c^2 + y^2 = \varepsilon^2 x^2 - 2\varepsilon^2 x\frac{a}{\varepsilon} + \varepsilon^2\frac{a^2}{\varepsilon^2} = \frac{c^2}{a^2}x^2 - 2\frac{c}{a}xa + a^2.$$

После сокращений и приведения подобных имеем

$$x^2\left(1 - \frac{c^2}{a^2}\right) + y^2 = a^2 - c^2.$$

Вспомним, что $a^2 - c^2 = b^2$, и разделим обе части последнего равенства на b^2:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Это и есть каноническое уравнение эллипса, значит, координаты точки M удовлетворяют уравнению эллипса, что и означает принадлежность точки M эллипсу. Проведенные рассуждения показывают, что в качестве определения эллипса можно взять следующее.

Эллипсом называется геометрическое место точек, отношение расстояний от которых до данной точки, именуемой фокусом, и до данной прямой, именуемой директрисой, есть постоянная величина, равная эксцентриситету ε***, где*** $0 < \varepsilon < 1$.

10.1.6. Оптические свойства эллипса. И, наконец, поговорим об оптических свойствах эллипса. Точки F_1 и F_2 не случайно называются фокусами. Представим себе, что внутренняя сторона эллипса зеркальна и луч света, дойдя до нее, отражается. Тогда можно показать, что если луч света выходит из точки F_1 в любом направлении, то после отражения он приходит в точку F_2, и наоборот. Образно говоря, в F_2 зажигаем свечу, а в F_1 от сфокусировавшихся лучей загорается спичка (рис. 46).

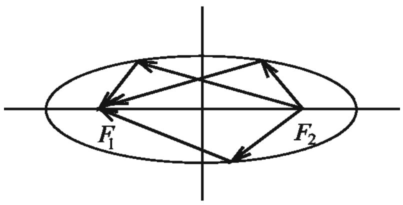


Рис. 46

10.1.7. Гипербола

Определение. ***Гиперболой*** называется геометрическое место точек, разность расстояний от которых до двух данных точек, именуемых ***фокусами***, постоянна и равна по модулю $2a$.

Как и в случае эллипса, фокусы гиперболы обозначим F_1 и F_2, а расстояния между ними $F_1F_2 = 2c$. Будем считать, что $c > 0$ и $0 < a < c$. (Это следует из того, что в треугольнике MF_1F_2 модуль разности сторон $|MF_1 - MF_2| = 2a$ меньше стороны $F_1F_2 = 2c$.)

Каноническую систему координат выбираем так же, как в случае эллипса, т. е. за ось Ox принимаем фокальную ось, а ось Oy проведем через середину отрезка F_1F_2 (рис. 47).

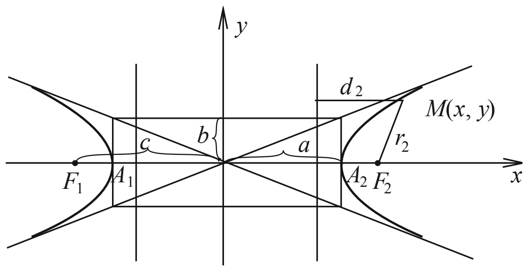


Рис. 47

Выведем каноническое уравнение гиперболы. По определению $|\overline{MF_1}| - |\overline{MF_2}| = 2a$, тогда

$$\sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2} = 2a.$$

Далее все выкладки проведем аналогично случаю эллипса.

$$\left(\sqrt{(x+c)^2+y^2}\right)^2 = \left(2a+\sqrt{(x-c)^2+y^2}\right)^2;$$

$$x^2+2cx+c^2+y^2 = 4a^2+4a\sqrt{(x-c)^2+y^2}+x^2-2cx+c^2+y^2;$$

$$4cx-4a^2 = 4a\sqrt{(x-c)^2+y^2};$$

$$(cx-a^2)^2 = a^2(x-c)^2+a^2y^2;$$

$$c^2x^2-2a^2cx+a^4 = a^2x^2-2a^2cx+a^2c^2+a^2y^2;$$

$$x^2(c^2-a^2)-a^2y^2 = a^2c^2-a^4 = a^2(c^2-a^2).$$

Обозначим $c^2-a^2=b^2$ и разделим обе части последнего равенства на a^2b^2:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1.$$

Это уравнение называется ***каноническим уравнением гиперболы***. Если выяснять геометрический смысл параметров ***a*** и ***b*** так же, как для эллипса, то получим, что при $y=0$; $x=\pm a$, т. е. точки $A_1(-a;\ 0)$, $A_2(a;\ 0)$ принадлежат линии и называются ***вершинами*** гиперболы, а расстояние между ними оказывается равным $2a$. При $x=0$ получим $y=\pm bi$, т. е. ***действительных*** точек пересечения гиперболы с осью Oy нет; однако же, по аналогии с эллипсом, параметры гиперболы имеют следующие названия:

a – ***действительная полуось***;

b – ***мнимая полуось***;

c – ***половина фокусного расстояния***.

Обратим внимание читателя на то, что в отличие от эллипса для гиперболы верно соотношение $c>a$, и связь между параметрами имеет вид

$$c^2=a^2+b^2.$$

Определение. ***Эксцентриситетом*** гиперболы называется число $\varepsilon=\dfrac{c}{a}$. Понятно, что для гиперболы $\varepsilon>1$.

Аналогично случаю эллипса расстояния от любой точки гиперболы до фокусов называются ***фокальными радиусами*** и обозначаются как r_1 и r_2. Но в отличие от эллипса здесь нужно рассмотреть отдельно точки левой и правой ветвей. Для левой ветви справедливо $\varepsilon x<x\le -a$, откуда $-\varepsilon x-a>0,\ -\varepsilon x+a>0$, а для правой ветви, напротив, верно $a\le x<\varepsilon x$, откуда $\varepsilon x-a>0,\ \varepsilon x+a>0$. А из канонического уравнения $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ для любой точки гиперболы выполняется $y^2=\dfrac{b^2}{a^2}x^2-b^2$.

Пусть точка $M(x;y)$ лежит на левой ветви гиперболы. Найдем для нее оба фокальных радиуса:

$$r_1=|\overline{MF_1}|=\sqrt{(x+c)^2+y^2}=\sqrt{x^2+2cx+c^2+\frac{b^2}{a^2}x^2-b^2}=$$

$$=\sqrt{x^2\left(1+\frac{b^2}{a^2}\right)+2cx+c^2-b^2}=\sqrt{x^2\frac{c^2}{a^2}+2cx+a^2}=$$

$$=\sqrt{\left(\frac{c}{a}x+a\right)^2}=-\varepsilon x-a\ ;$$

$$r_2=|\overline{MF_2}|=\sqrt{(x-c)^2+y^2}=\sqrt{x^2\left(1+\frac{b^2}{a^2}\right)-2cx+a^2}=$$

$$=\sqrt{\left(\frac{c}{a}x-a\right)^2}=-\varepsilon x+a\,.$$

Аналогично находим r_1 и r_2 для точек правой ветви. Полученные формулы удобно представить в форме таблицы.

Левая ветвь	Правая ветвь
$r_1=-\varepsilon x-a$ $r_2=-\varepsilon x+a$	$r_1=\varepsilon x+a$ $r_2=\varepsilon x-a$

Заметим, что для обеих ветвей $|r_1-r_2|=2a$. У гиперболы так же, как и у эллипса, существуют директрисы.

Определение. ***Директрисами*** гиперболы называются прямые, перпендикулярные ее фокальной оси и обладающие тем свойством, что отношение расстояний от любой точки гиперболы до фокуса и до соответствующей директрисы есть величина постоянная, равная эксцентриситету.

Так же, как и в случае эллипса, директрисы гиперболы не пересекаются с ней, поэтому если уравнение правой директрисы имеет вид $x=x_0$, то $x_0<a$. Тогда расстояние от точки $M(x;\,y)$, лежащей на правой ветви гиперболы, до правой директрисы,

$$d_2=x-x_0;\quad r_2=\varepsilon x-a\,.$$

Найдем значение x_0. Из равенства $\dfrac{\varepsilon x - a}{x - x_0} = \varepsilon$ получим

$$x_0 = \frac{a}{\varepsilon}.$$

Аналогично для левой ветви $x_0 = -\dfrac{a}{\varepsilon}$.

Как видим, уравнения директрис гиперболы имеют точно такой же вид, что для эллипса $x = \pm\dfrac{a}{\varepsilon}$. Совершенно так же, как и в случае эллипса (см. конец п. 10.1.5), можно доказать, что ***если для произвольной точки плоскости*** $M(x; y)$ ***отношение расстояний от нее до директрисы*** $x = \dfrac{a}{\varepsilon}$ ***и до фокуса*** F_2 ***равно*** ε***, то*** $M(x; y)$ ***лежит на гиперболе.*** Аналогично и для директрисы $x = -\dfrac{a}{\varepsilon}$ и точки F_1. Это позволяет дать другое определение гиперболы.

Гиперболой называется геометрическое место точек, отношение расстояний от которых до данной точки, именуемой фокусом, и до данной прямой, именуемой директрисой, есть постоянная величина, равная ε***, где*** $\varepsilon > 1$.

До сих пор рассмотрение гиперболы проводилось в точности так же, как эллипса. Но между ними есть существенное различие, обусловленное тем, что слово «сумма» в определении эллипса заменено словом «разность» в определении гиперболы. Дело в том, что ***эллипс – кривая ограниченная, а гипербола – неограниченная***. В школе гипербола рассматривалась как график обратно пропорциональной зависимости $y = \dfrac{k}{x}$. Оси координат служили для такой гиперболы ***асимптотами***, т. е. такими прямыми, к которым гипербола все более приближается при неограниченном удалении вдоль них (хотя и никогда их не пересекает). Такое «образное» определение асимптоты, конечно, не строго, но оно хорошо проясняет суть дела. Мы позволили себе ограничиться лишь таким определением, так как строгое дается в курсе математического анализа.

Вернемся к нашей гиперболе. У нее тоже есть асимптоты (чего нет у эллипса). Рассмотрим прямоугольник, определяемый неравенствами $|x| \le a$, $|y| \le b$. Этот прямоугольник называется ***основным*** для гиперболы. На сторонах этого прямоугольника, параллельных оси *Oy*, лежат вершины гиперболы A_1 и A_2; стороны, параллельные оси *Ox*, с гиперболой общих точек не имеют. ***Продолжения диагоналей этого прямоугольника и служат асимптотами гиперболы***. Это прямые с уравнениями $y = \pm\frac{b}{a}x$. Гипербола, для которой $b = a$, называется ***равнобочной***. В этом случае основной прямоугольник превращается в квадрат и асимптоты гиперболы, как диагонали квадрата, становятся взаимно перпендикулярными. Это и есть случай «школьной» гиперболы. Она приобретает привычный вид, если за оси координат принять ее асимптоты, т. е. повернуть систему координат на угол 45°. При этом уравнение гиперболы будет иметь вид

$$xy = k \text{ или } y = \frac{k}{x}.$$

10.1.8. Оптические свойства гиперболы. С точки зрения полуосей гипербола – линия наполовину «мнимая» по сравнению с эллипсом. И ее оптическое свойство также демонстрирует эту мнимость. Именно ***луч, выходя из одного фокуса гиперболы, после отражения от любой ее ветви кажется исходящим из другого фокуса*** (рис. 48).

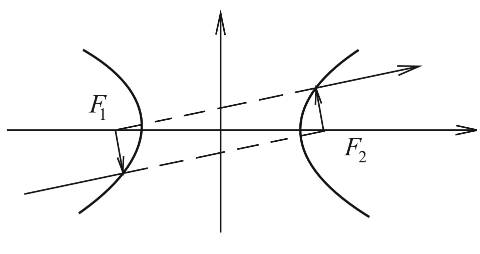


Рис. 48

10.1.9. Парабола

Определение. ***Параболой*** называется геометрическое место точек, равноудаленных от точки, именуемой ***фокусом***, и прямой, называемой ***директрисой***.

Для канонической системы координат за ось Ox принимаем прямую, проходящую через фокус перпендикулярно директрисе, ось Oy проводим через середину расстояния между фокусом и директрисой. Положительным направлением на оси Ox считаем направление от директрисы к фокусу. Расстояние между фокусом и директрисой обозначается p и называется параметром параболы или фокальным параметром.

Выведем каноническое уравнение параболы.

Пусть точка $M(x;\ y)$ – произвольная точка параболы (рис. 49).

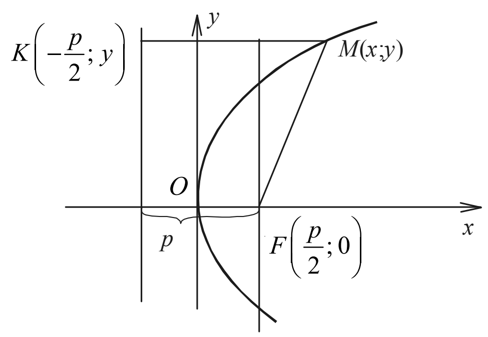


Рис. 49

Тогда по определению $|\overline{MF}| = |\overline{MK}|$, т. е.

$$\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}=\sqrt{\left(x+\frac{p}{2}\right)^2}$$

или, избавляясь от радикала и приведя подобные, получим

$$\left(x-\frac{p}{2}\right)^2+y^2=\left(x+\frac{p}{2}\right)^2;$$

$$x^2-px+\frac{p^2}{4}+y^2=x^2+px+\frac{p^2}{4};$$

$$y^2 = 2px.$$

Это уравнение называется ***каноническим уравнением параболы***. Из определения параболы следует, что для нее ***эксцентриситет равен единице***. Поскольку уравнение параболы не меняется при замене ***y*** на (–***y***), ***график параболы симметричен относительно оси Ох***. В отличие от эллипса и гиперболы ***парабола не обладает центральной симметрией***.

Директриса параболы имеет уравнение

$$x = -\frac{p}{2}.$$

Нетрудно найти формулу для фокального радиуса. Если точка $M(x;\ y)$ лежит на параболе, то

$$r = |\overline{MF}| = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \sqrt{\left(x^2 - px + \frac{p^2}{4} + 2px\right)} =$$

$$= \sqrt{\left(x^2 + px + \frac{p^2}{4}\right)} = \sqrt{\left(x + \frac{p}{2}\right)^2} = x + \frac{p}{2}.$$

В школьной программе парабола появляется как график квадратного трехчлена $y = ax^2 + bx + c$. Чтобы такое уравнение привести к каноническому виду, достаточно выделить полный квадрат, т. е. привести уравнение к виду $y - y_0 = a(x - x_0)^2$, а затем сделать замену переменных $x - x_0 = x'$; $y - y_0 = y'$, после чего получим $y' = ax'^2$ или $x'^2 = \frac{1}{a}y'$, т. е. в этой параболе переменные x и y поменялись ролями и для нее осью симметрии служит ось Oy.

10.1.10. Оптические свойства параболы. Если считать, что внутренняя сторона параболы зеркальна и источник света поместить в фокус, то все лучи, отразившись от параболы, будут направлены по прямым, параллельным фокальной оси (рис. 50).

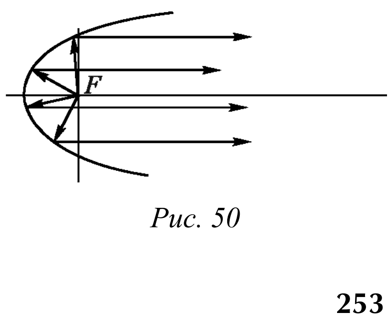


Рис. 50

Наоборот, если лучи света падают на внутреннюю сторону линии параллельно фокальной оси, то, отразившись, все они собираются в фокусе. На этих свойствах основано устройство отражателей, прожекторов и различных антенн.

10.2. УНИВЕРСАЛЬНОЕ ОПРЕДЕЛЕНИЕ ЭЛЛИПСА, ГИПЕРБОЛЫ И ПАРАБОЛЫ

10.2.1. Определение с помощью эксцентриситета. Как мы видим, для каждой из перечисленных кривых существует определение, использующее эксцентриситет. Все определения отличаются друг от друга только величиной эксцентриситета. Это позволяет дать универсальное определение всех трех кривых.

Определение. ***Геометрическое место точек, отношение расстояния от которых до данной точки, называемой фокусом, к расстоянию до данной прямой, называемой директрисой, равно постоянному числу*** ε***, есть кривая второго порядка, а именно*****:**

– эллипс **(*при*** $0<\varepsilon<1$**)**;

– парабола **(*при*** $\varepsilon=1$**)**;

– гипербола **(*при*** $\varepsilon>1$**)**.

Подчеркнем, что окружность не «вписывается» в это определение, так как ***для окружности*** $\varepsilon=0$.

10.2.2. Определение с помощью конических сечений. Однако есть еще одно универсальное определение кривых второго порядка, куда попадает уже и окружность. Речь идет о так называемой теории конических сечений. Доскональное изучение этой теории выходит за рамки нашей задачи, мы лишь покажем «картинки», иллюстрирующие ее.

Рассмотрим коническую поверхность, образованную вращением двух пересекающихся прямых (так называемых «образующих») вокруг одной из биссектрис вертикальных углов, образованных этими прямыми. Сечения этой поверхности различными плоскостями дают все кривые второго порядка. Если исключить случай, когда секущая плоскость проходит через вершину конусов (в этом случае получаются либо одна точка, либо одна прямая, либо пара пересекающихся прямых), то все зависит от того, пересакает плоскость один из конусов или оба. Если плоскость пересекает только один из конусов, получаются либо окружности (плоскость перпендикулярна оси симметрии), либо параболы (плоскость параллельна одной из образующих), либо эллипсы. Если же плоскость пересекает оба конуса, получаются гиперболы (рис. 51).

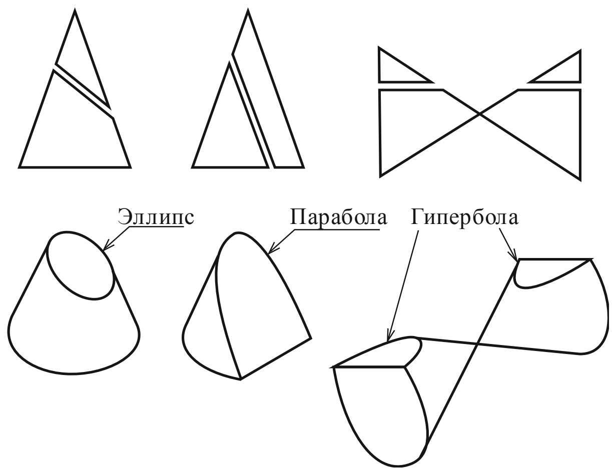


Рис. 51

10.3. ПРИВЕДЕНИЕ ОБЩЕГО УРАВНЕНИЯ ЛИНИЙ ВТОРОГО ПОРЯДКА К КАНОНИЧЕСКОМУ ВИДУ

10.3.1. Предварительные замечания. Вспомним общее уравнение линии второго порядка (10.1.1) и проведем некоторые предварительные рассуждения.

1. Все канонические уравнения линий второго порядка не содержат смешанного произведения переменных.

2. Первые три слагаемых общего уравнения являются квадратичной формой.

3. При приведении квадратичной формы к каноническому виду все смешанные произведения исчезают.

Отсюда следует вывод: для приведения общего уравнения линии второго порядка к каноническому виду нужно прежде всего привести к каноническому виду квадратичную форму. Эта задача решена нами в п. 9.5.2 для произвольной квадратичной формы. Здесь же мы детально опишем этот алгоритм для линии второго порядка.

10.3.2. Пример. Для простоты возьмем конкретный пример. Привести к каноническому виду следующее уравнение кривой второго порядка:

$$7x^2+12xy+2y^2+82x+20y+17=0 \ .$$

1. Составим симметрическую матрицу квадратичной формы $Q(x)=7x^2+12xy+2y^2$:

$$A=\begin{pmatrix} 7 & 6 \\ 6 & 2 \end{pmatrix}.$$

2. Найдем собственные числа матрицы A, для чего решим характеристическое уравнение $|A-\lambda E|=0$. Воспользуемся формулой п. 8.3.3.

$$|A-\lambda E|=(-\lambda)^2+(-\lambda)(7+2)+(-22)=\lambda^2-9\lambda-22=0.$$

$$\lambda_1=11, \quad \lambda_2=-2 .$$

3. Для каждого λ найдем собственный вектор и нормируем его. Пусть $\lambda_1=11$. Составляем расширенную матрицу однородной системы линейных уравнений с матрицей $A-\lambda_1 E$:

$$\left(\begin{array}{cc|c} 7-11 & 6 & 0 \\ 6 & 2-11 & 0 \end{array}\right)=\left(\begin{array}{cc|c} -4 & 6 & 0 \\ 6 & -9 & 0 \end{array}\right)\sim(2 \quad -3 \mid 0);$$

$$2x_1-3x_2=0.$$

Пусть $x_2=2$, тогда $x_1=3$. Собственным вектором, соответствующим $\lambda_1=11$, будет $\vec{c}_1=\begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Нормируем этот вектор

$$\vec{e}_1=\frac{1}{|\vec{c}_1|}\vec{c}_1=\frac{1}{\sqrt{13}}\begin{pmatrix} 3 \\ 2 \end{pmatrix}=\begin{pmatrix} \dfrac{3}{\sqrt{13}} \\ \dfrac{2}{\sqrt{13}} \end{pmatrix}.$$

Теперь пусть $\lambda_2 = -2$.

$$\left(\begin{array}{cc|c} 7+2 & 6 & 0 \\ 6 & 2+2 & 0 \end{array}\right) = \left(\begin{array}{cc|c} 9 & 6 & 0 \\ 6 & 4 & 0 \end{array}\right) \sim (3 \quad 2 \mid 0),$$

$$3x_1 + 2x_2 = 0.$$

Пусть $x_2 = 3$, тогда $x_1 = -2$. Собственный вектор, соответствующий $\lambda_2 = -2$, будет $\vec{c}_2 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$. Нормируем его:

$$\vec{e}_2 = \frac{1}{|\overline{c}_2|}\vec{c}_2 = \frac{1}{\sqrt{13}}\begin{pmatrix} -2 \\ 3 \end{pmatrix} = \begin{pmatrix} -\dfrac{\sqrt{2}}{13} \\ \dfrac{3}{\sqrt{13}} \end{pmatrix}.$$

Итак, в базисе $\overline{e}_1$, $\overline{e}_2$ собственных векторов матрица A квадратичной формы имеет диагональный вид. Пусть матрицей перехода от базиса $\overline{i}$, $\overline{j}$ к базису $\overline{e}_1$, $\overline{e}_2$ служит матрица H. Тогда столбцами матрицы H^{-1} служат координаты векторов $\overline{e}_1$, $\overline{e}_2$ в базисе $\overline{i}$, $\overline{j}$:

$$H^{-1} = \begin{pmatrix} \dfrac{3}{\sqrt{13}} & -\dfrac{2}{\sqrt{13}} \\ \dfrac{2}{\sqrt{13}} & \dfrac{3}{\sqrt{13}} \end{pmatrix}.$$

Сделав преобразование координат с матрицей H^{-1}, приведем квадратичную форму к каноническому виду, т. е. к сумме квадратов:

$$\vec{x} = H^{-1}\vec{x}',$$

или

$$\begin{pmatrix} x \\ y \end{pmatrix} = H^{-1}\begin{pmatrix} x' \\ y' \end{pmatrix},$$

что означает

$$x = \frac{3}{\sqrt{13}}x' - \frac{2}{\sqrt{13}}y'; \qquad y = \frac{2}{\sqrt{13}}x' + \frac{3}{\sqrt{13}}y'.$$

Подставим эти выражения координат x, y через x', y' в исходное уравнение линии второго порядка. Тогда получается

$$7\left(\frac{3}{\sqrt{13}}x'-\frac{2}{\sqrt{13}}y'\right)^2+12\left(\frac{3}{\sqrt{13}}x'-\frac{2}{\sqrt{13}}y'\right)\left(\frac{2}{\sqrt{13}}x'+\frac{3}{\sqrt{13}}y'\right)+$$

$$+2\left(\frac{2}{\sqrt{13}}x'+\frac{3}{\sqrt{13}}y'\right)^2+82\left(\frac{3}{\sqrt{13}}x'-\frac{2}{\sqrt{13}}y'\right)+$$

$$+20\left(\frac{2}{\sqrt{13}}x'+\frac{3}{\sqrt{13}}y'\right)+17=$$

$$=x'^2\left(\frac{63}{13}+\frac{72}{13}+\frac{8}{13}\right)+x'y'\left(-\frac{84}{13}-\frac{48}{13}+\frac{108}{13}+\frac{24}{13}\right)+$$

$$+y'^2\left(\frac{28}{13}-\frac{72}{13}+\frac{18}{13}\right)+x'\left(\frac{246}{\sqrt{13}}+\frac{40}{\sqrt{13}}\right)+y'\left(-\frac{164}{\sqrt{13}}+\frac{60}{\sqrt{13}}\right)+17=$$

$$=11x'^2-2y'^2+22\sqrt{13}\,x'-8\sqrt{13}\,y'+17=0\,.$$

Для приведения уравнения к каноническому виду выделим полные квадраты по переменным x' и y':

$$11\left(x'^2+2\sqrt{13}\,x'+13-13\right)-2\left(y'^2+4\sqrt{13}\,y'+52-52\right)+17=$$

$$=11\left(x'+\sqrt{13}\right)^2-143-2\left(y'+2\sqrt{13}\right)^2+104+17=$$

$$=11\left(x'+\sqrt{13}\right)^2-2\left(y'+2\sqrt{13}\right)^2-22=0\,,$$

или

$$\frac{\left(x'+\sqrt{13}\right)^2}{2}-\frac{\left(y'+2\sqrt{13}\right)^2}{11}=1\,.$$

Последнее уравнение в системе координат $Ox'y'$ является уравнением гиперболы с действительной полуосью $a=\sqrt{2}$, мнимой полуосью $b=\sqrt{11}$ и центром в точке $\left(-\sqrt{13};\ -2\sqrt{13}\right)$.

Ось Ox' направлена по вектору $\overline{e}_1 = \left(\frac{3}{\sqrt{13}}; \frac{2}{\sqrt{13}} \right)$,

ось Oy' – по вектору $\overline{e}_2 = \left(-\frac{2}{\sqrt{13}}; \quad \frac{3}{\sqrt{13}} \right)$.

Обозначим $x'' = x' + \sqrt{13}, \quad y'' = y' + 2\sqrt{13}$. В системе координат $O''x''y''$, полученной из системы $Ox'y'$ переносом начала координат, уравнение гиперболы примет вид

$$\frac{x''^2}{2} - \frac{y''^2}{11} = 1.$$

Выполним чертеж (рис. 52).

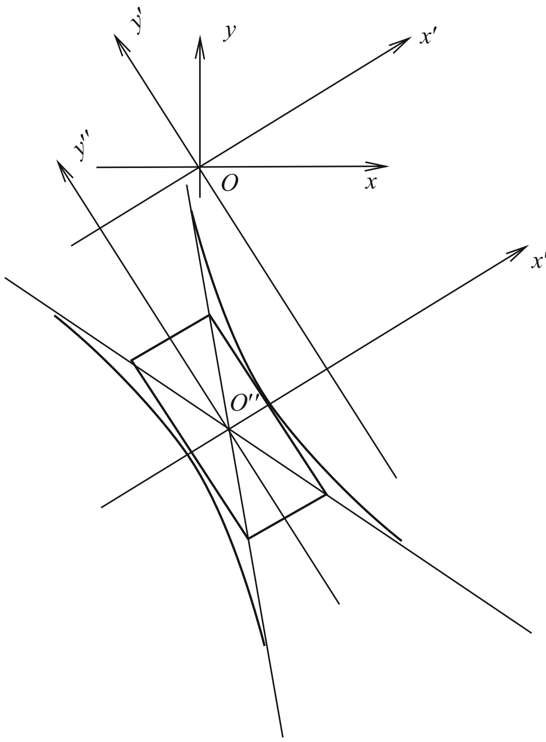


Рис. 52

Тем самым решена задача приведения общего уравнения линии второго порядка к каноническому виду.

10.4. ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА

10.4.1. Эллипсоиды

Определение. ***Эллипсоидом*** называется поверхность, уравнение которой в некоторой прямоугольной системе координат имеет вид

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \qquad a\ge b\ge c>0.$$

Эта система координат называется канонической для данного эллипсоида, и данное уравнение тоже называется ***каноническим уравнением эллипсоида***.

При $a=b=c$ оно задает сферу с центром в начале координат и радиусом, равным a. Поэтому сфера – частный случай эллипсоида. Но на сферу можно посмотреть как на поверхность вращения. Например, при вращении полуокружности

$$x=\sqrt{a^2-z^2}$$

вокруг оси Oz получится сфера, задаваемая уравнением

$$x^2+y^2+z^2=a^2.$$

Аналогично, если правую часть эллипса $\frac{x^2}{a^2}+\frac{z^2}{c^2}=1$ вращать вокруг оси Oz, получится ***эллипсоид вращения.***

Как же будет выглядеть уравнение такого эллипсоида? Чтобы понять это, проведем рассуждения в общем случае. Пусть на плоскости Oxz задана кривая с уравнением

$$f(x, z)=0.$$

Перейдем к цилиндрическим координатам:

$$x=\rho\cos\varphi; \qquad y=\rho\sin\varphi; \qquad z=z.$$

Напомним, что в цилиндрической системе координат φ, ρ, z каждая точка характеризуется полярными координатами φ и ρ своей проекции на плоскость Oxy и своей аппликатой z, той же, что и в декартовой системе координат. В цилиндрической системе координат уравнение поверхности вращения вокруг оси Oz не должно зависеть от угла φ, поэтому оно будет иметь вид

$$f(\rho,\ z) = 0 \quad \text{или} \quad f\left(\sqrt{x^2 + y^2},\ z\right) = 0,$$

т. е. чтобы получить уравнение поверхности вращения вокруг оси Oz линии $f(x,\ z) = 0$, нужно в уравнение линии вместо x подставить $\sqrt{x^2 + y^2}$. Тогда уравнение эллипсоида вращения получается следующее:

$$\frac{\left(\sqrt{x^2 + y^2}\right)^2}{a^2} + \frac{z^2}{c^2} = 1 \quad \text{или} \quad \frac{x^2 + y^2}{a^2} + \frac{z^2}{c^2} = 1.$$

Аналогично, если эллипс

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

вращать вокруг оси Ox, получим эллипсоид вращения с уравнением

$$\frac{x^2}{a^2} + \frac{y^2 + z^2}{b^2} = 1.$$

Теперь понятно, что эллипсоид, задаваемый каноническим уравнением, будет эллипсоидом вращения, когда из его трех параметров a, b, c хотя бы два совпадают.

Вернемся к эллипсоиду, задаваемому своим каноническим уравнением, и исследуем его форму так называемым методом параллельных сечений. Геометрически этот метод означает, что исследуются формы кривых, которые получаются при сечении исследуемой поверхности плоскостями, параллельными координатным плоскостям.

Рассмотрим сечение плоскостью $x = h$, параллельной плоскости Oyz. В сечении получаем

$$\frac{h^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

или
$$\frac{y^2}{b^2}+\frac{z^2}{c^2}=1-\frac{h^2}{a^2},$$

и, разделив все уравнение на правую часть $\left(1-\frac{h^2}{a^2}\right)$, получим

$$\frac{y^2}{b^2\left(1-\frac{h^2}{a^2}\right)}+\frac{z^2}{c^2\left(1-\frac{h^2}{a^2}\right)}=1.$$

Исследуем это уравнение. Если $|h|>a$, то точек пересечения нет, так как в этом случае $1-\frac{h^2}{a^2}<0$, а $\frac{y^2}{b^2}+\frac{z^2}{c^2}>0$. Если $|h|=a$, то $y=0$, $z=0$ и имеется только точка пересечения $(a;\ 0;\ 0)$ или $(-a;\ 0;\ 0)$, лежащая на оси Ox.

Если же $|h|<a$, то в сечении образуется эллипс с полуосями

$$b\sqrt{1-\frac{h^2}{a^2}} \quad \text{и} \quad c\sqrt{1-\frac{h^2}{a^2}},$$

которые при уменьшении $|h|$ от a до 0 увеличиваются от 0 до b или от 0 до c соответственно.

Аналогичные результаты получаем при сечении эллипсоида плоскостями $y=h$ и $z=h$.

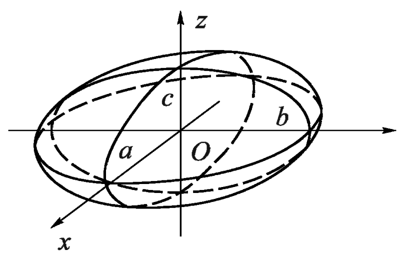


Рис. 53

Исследование показывает, что эллипсоид целиком расположен внутри параллелепипеда с центром в точке $O(0;\ 0;\ 0)$ с гранями, параллельными координатным плоскостям, и со сторонами, имеющими длины $2a$, $2b$ и $2c$ (рис. 53).

Кроме того, поскольку уравнение эллипсоида не меняется от замены x, y, z на $(-x)$, $(-y)$, $(-z)$, плоскости координат служат его плоскостями симметрии, а начало координат – его центром симметрии.

10.4.2. Двуполостные гиперболоиды

Определение. ***Двуполостным гиперболоидом*** называется поверхность, имеющая в некоторой прямоугольной системе координат уравнение вида

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1,$$

где $a \geq b > 0$ и $c > 0$.

Эта система координат называется канонической, и уравнение называется ***каноническим***.

Если на плоскости Oxz рассмотреть гиперболу

$$-\frac{x^2}{a^2}+\frac{z^2}{c^2}=1$$

с действительной осью Oz, то поверхность, образуемая вращением этой гиперболы вокруг оси Oz, имеет уравнение

$$-\frac{x^2+y^2}{a^2}+\frac{z^2}{c^2}=1,$$

или

$$\frac{x^2+y^2}{a^2}-\frac{z^2}{c^2}=-1 \text{ (рис. 54).}$$

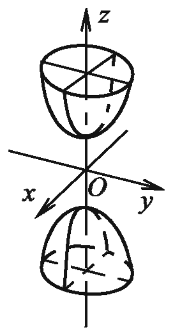


Рис. 54

Последнее уравнение получается из канонического уравнения при $a = b$. Форма двуполостного гиперболоида вращения вполне понятна и объясняет эпитет «двуполостный».

Исследуем гиперболоид методом параллельных сечений. Начнем с плоскостей, параллельных плоскости Oxy, т. е. имеющих уравнение $z = h$. Получим уравнение

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{h^2}{c^2}=-1,$$

или

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{h^2}{c^2}-1.$$

Если $|h| < c$, то точек пересечения нет, так как в этом случае

$$\frac{x^2}{a^2}+\frac{y^2}{b^2} \geq 0, \quad \text{а} \quad \frac{h^2}{c^2}-1<0.$$

Если $|h| = c$, то

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=0,$$

откуда $x = 0$, $y = 0$, и имеется только две точки пересечения: (0; 0; c) или (0; 0; $-c$), лежащие на оси Oz. Если же $|h| > c$, то уравнение линии пересечения имеет вид

$$\frac{x^2}{a^2\left(\frac{h^2}{c^2}-1\right)}+\frac{y^2}{b^2\left(\frac{h^2}{c^2}-1\right)}=1,$$

откуда следует, что линией пересечения является эллипс с полуосями

$$a\sqrt{\frac{h^2}{c^2}-1} \quad \text{и} \quad b\sqrt{\frac{h^2}{c^2}-1}.$$

Понятно, что при неограниченном увеличении высоты сечения $|h|$, начиная со значения большего c до ∞, полуоси неограниченно растут.

Теперь исследуем сечения двуполостного гиперболоида плоскостями, параллельными плоскости Oxz, т. е. имеющими уравнение $y = h$. Получим

$$\frac{x^2}{a^2}+\frac{h^2}{b^2}-\frac{z^2}{c^2}=-1, \quad \text{или} \quad \frac{x^2}{a^2}-\frac{z^2}{c^2}=-1-\frac{h^2}{b^2},$$

или

$$-\frac{x^2}{a^2\left(1+\frac{h^2}{b^2}\right)}+\frac{z^2}{c^2\left(1+\frac{h^2}{b^2}\right)}=1.$$

Это уравнение представляет собой уравнение гиперболы с действительной осью Oz и полуосями:

$a\sqrt{1+\dfrac{h^2}{b^2}}$ – мнимая полуось;

$c\sqrt{1+\dfrac{h^2}{b^2}}$ – действительная полуось.

При $h=0$ получается гипербола с наименьшими полуосями, равными a и c.

При неограниченном увеличении $|\,h\,|$ начиная с 0 до ∞ полуоси тоже неограниченно увеличиваются начиная со значений a и c соответственно до ∞.

Аналогично при сечении двуполостного гиперболоида плоскостями, параллельными плоскости Oyz, получаем гиперболы с полуосями:

$b\sqrt{1+\dfrac{h^2}{a^2}}$ – мнимая полуось;

$c\sqrt{1+\dfrac{h^2}{a^2}}$ – действительная полуось,

с действительной осью, параллельной оси Oz, мнимой осью, параллельной оси Oy.

Из приведенного выше следует, что двуполостный гиперболоид состоит из двух частей, располагающихся в полупространствах $z \ge c$ и $z \le -c$. Так же, как и в случае эллипсоида, двуполостный гиперболоид симметричен относительно всех координатных плоскостей и относительно начала координат.

10.4.3. Однополостные гиперболоиды

Определение. ***Однополостным гиперболоидом*** называется поверхность, имеющая в некоторой прямоугольной системе координат уравнение вида

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1.$$

Эта система координат называется канонической для данного гиперболоида, а уравнение называется ***каноническим уравнением*** однополостного гиперболоида.

Как и в п. 10.4.2, попробуем вращать соответствующую гиперболу. А именно рассмотрим поверхность, образуемую вращением гиперболы

$$\frac{x^2}{a^2}-\frac{z^2}{c^2}=1$$

вокруг оси Oz. Тогда получим уравнение

$$\frac{\left(\sqrt{x^2+y^2}\right)^2}{a^2}-\frac{z^2}{c^2}=1, \quad \text{или} \quad \frac{x^2+y^2}{a^2}-\frac{z^2}{c^2}=1.$$

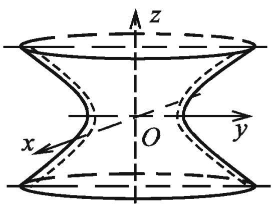


Рис. 55

Последнее уравнение получается из канонического уравнения при $a = b$. Значит, однополостный гиперболоид вращения имеет вид (рис. 55)

Исследуем однополостный гиперболоид методом параллельных сечений. Вначале рассмотрим сечения гиперболоида плоскостями, параллельными плоскости Oxy, т. е. имеющими уравнения $z = h$. Получим

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{h^2}{c^2}=1, \quad \text{или} \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=1+\frac{h^2}{c^2}.$$

Сразу видим, что при любых значениях величины h линия пересечения существует и представляет собой эллипс с полуосями

$$a\sqrt{1+\frac{h^2}{c^2}} \quad \text{и} \quad b\sqrt{1+\frac{h^2}{c^2}}.$$

При $h = 0$ полуоси равны соответственно a и b, а при неограниченном возрастании $|h|$ начиная с 0 полуоси тоже неограниченно возрастают от значений a или b. Теперь рассмотрим линии пересечения однополостного гиперболоида с плоскостями, параллельными плоскости Oxz, т. е. имеющими уравнение $y = h$. Получаем

$$\frac{x^2}{a^2}+\frac{h^2}{b^2}-\frac{z^2}{c^2}=1, \quad \text{или} \quad \frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{h^2}{b^2}.$$

При $0 \le |h| < b$ получаются гиперболы с действительной осью Ox и такими полуосями:

$a\sqrt{1-\dfrac{h^2}{b^2}}$ – действительная полуось;

$c\sqrt{1-\dfrac{h^2}{b^2}}$ – мнимая полуось.

При возрастании $|h|$ в полуинтервале $[0, b)$ эти полуоси уменьшаются соответственно от значений a и c. При $|h| = b$ получаем уравнение

$$\frac{x^2}{a^2}-\frac{z^2}{c^2}=0, \quad \text{или} \quad \left(\frac{x}{a}-\frac{z}{c}\right)\left(\frac{x}{a}+\frac{z}{c}\right)=0.$$

Это означает, что линия пересечения гиперболоида с плоскостями $y = b$ или $y = -b$ представляет собой две прямые, которые описываются уравнениями

$$\begin{cases} \dfrac{x}{a}-\dfrac{z}{c}=0; \\ y=b. \end{cases} \qquad \begin{cases} \dfrac{x}{a}+\dfrac{z}{c}=0; \\ y=b. \end{cases}$$

Если $|h| > b$, то линия пересечения имеет уравнение

$$\frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{h^2}{b^2} \quad \text{или} \quad -\frac{x^2}{a^2}+\frac{z^2}{c^2}=\frac{h^2}{b^2}-1.$$

Это гипербола с действительной осью Oz и такими полуосями:

$a\sqrt{\dfrac{h^2}{b^2}-1}$ – мнимая полуось;

$c\sqrt{\dfrac{h^2}{b^2}-1}$ – действительная полуось.

При неограниченном возрастании величины $|h|$ эти полуоси неограниченно возрастают.

Обратим внимание читателя на то, что при $0 \le |h| < b$ получаются гиперболы с действительной осью, параллельной оси Ox, и мнимой осью, параллельной оси Oz, а при $|h| > b$ образуются гиперболы «наоборот»: с действительной осью, параллельной оси Oz, и мнимой осью, параллельной оси Ox.

Исследование сечений однополостного гиперболоида плоскостями, параллельными плоскости Oyz, дает аналогичные результаты с заменой x на y.

Так же, как и в предыдущих пунктах, заметим, что однополостный гиперболоид симметричен относительно координатных осей и центра координат.

10.4.4. Гиперболические параболоиды

Определение. ***Гиперболическим параболоидом*** называется поверхность, имеющая в некоторой прямоугольной системе координат уравнение вида

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z,$$

где $p > 0$, $q > 0$. Эта система координат называется ***канонической*** для данного параболоида, а уравнение – его ***каноническим уравнением.***

К сожалению, в этом случае представить себе форму поверхности с помощью поверхностей вращения не получается, так как эта поверхность ни при каких значениях параметров не является поверхностью вращения. Остается только метод параллельных сечений.

Начнем с плоскостей, параллельных плоскости Oxy, т. е. имеющих уравнения $z = h$. Получим

$$\frac{x^2}{p} - \frac{y^2}{q} = 2h.$$

При $h > 0$ получаем гиперболы с действительной осью, параллельной оси Ox, и такими полуосями:

$\sqrt{2ph}$ – действительная полуось;

$\sqrt{2qh}$ – мнимая полуось.

При неограниченном возрастании h полуоси тоже неограниченно возрастают. При $h = 0$ получим

$$\frac{x^2}{p} - \frac{y^2}{q} = 0, \quad \text{или} \quad \left(\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}}\right)\left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}}\right) = 0,$$

т. е. уравнение описывает две прямые:

$$\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = 0 \text{ и } \frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} = 0,$$

лежащие на плоскости Oxy и проходящие через начало координат.

При $h < 0$ опять получаем гиперболы, но теперь действительная и мнимая ось меняются местами, а именно действительная ось проходит параллельно оси Oy, а мнимая – оси Ox. Соответственно полуоси гипербол равны:

$\sqrt{-2qh}$ – действительная полуось;

$\sqrt{-2ph}$ – мнимая полуось.

И здесь при неограниченном возрастании значения $|\,h\,|$ полуоси тоже неограниченно растут.

Заметим, что гиперболы при $h > 0$ и при $h < 0$ ведут себя точно так же, как гиперболы в случае однополостного гиперболоида, т. е. мнимые оси гипербол при $h > 0$ параллельны действительным осям гипербол при $h < 0$, а действительные оси гипербол при $h > 0$ параллельны мнимым осям гипербол при $h < 0$.

Теперь рассмотрим сечения гиперболического параболоида плоскостями, параллельными плоскости Oxz, т. е. имеющими уравнение $y = h$. Получим

$$\frac{x^2}{p} - \frac{h^2}{q} = 2z, \text{ или } x^2 = p\left(2z + \frac{h^2}{q}\right), \text{ или } x^2 = 2p\left(z + \frac{h^2}{2q}\right).$$

Последнее уравнение представляет собой уравнение параболы с параметром p, осью симметрии, параллельной оси Oz, и вершиной в точке $\left(0;\ h;\ -\frac{h^2}{2q}\right)$. При этом ветви параболы направлены вверх.

При сечении гиперболического параболоида плоскостями, параллельными плоскости Oyz, имеющими уравнение $x = h$, получим $\frac{h^2}{p}-\frac{y^2}{q}=2z$, откуда $\frac{y^2}{q}=-2\left(z-\frac{h^2}{2p}\right)$, и окончательно $y^2=-2q\left(z-\frac{h^2}{2p}\right)$.

Это уравнение параболы с параметром q, осью симметрии, параллельной оси Oz, и вершиной в точке $\left(x;\ 0;\ \frac{h^2}{2p}\right)$. При этом ветви параболы направлены вниз.

Чтобы лучше представить себе форму гиперболического параболоида, сделаем некоторые замечания. При $x = 0$ в сечении получаем параболу $y^2=-2qz$. При $y = h$ точки этой параболы имеют координаты $\left(0;\ h;\ -\frac{h^2}{2q}\right)$, а это в точности координаты вершин парабол, получающихся при сечении гиперболического параболоида плоскостями $y = h$.

Таким образом, можно представить себе гиперболический параболоид как поверхность, которая получается при «скольжении» параболы $x^2=2pz$, так что ее вершина движется по параболе $y^2=-2qz$. Вблизи начала координат эта поверхность имеет форму седла.

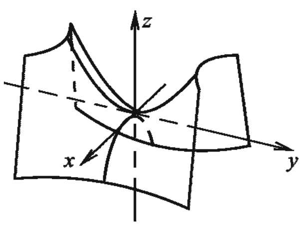


Рис. 56

Поскольку уравнение гиперболического параболоида не меняется при замене x и y на $(-x)$ и $(-y)$, эта поверхность симметрична относительно координатных плоскостей Oxz и Oyz (но не Oxy!) (рис. 56).

10.4.5. Эллиптические параболоиды

Определение. ***Эллиптическим параболоидом*** называется поверхность, имеющая в некоторой прямоугольной системе координат уравнение вида

$$\frac{x^2}{p}+\frac{y^2}{q}=2z,$$

где $p > 0$, $q > 0$.

Эта система координат называется ***канонической*** для данного параболоида, а уравнение называется ***каноническим уравнением*** этого эллиптического параболоида.

Здесь опять можно получить поверхность вращения. А именно, если параболу $x^2 = 2pz$ вращать вокруг оси Oz, то получим поверхность вращения с уравнением

$$x^2 + y^2 = 2pz, \quad \text{или} \quad \frac{x^2}{p} + \frac{y^2}{p} = 2z.$$

Последнее уравнение выводится из канонического уравнения при условии $p = q$. Описанную поверхность вращения легко себе представить (рис. 57).

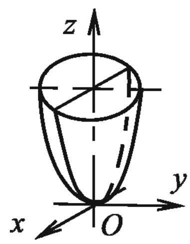


Рис. 57

Исследуем эллиптический параболоид методом параллельных сечений. Рассмотрим сечение этой поверхности плоскостью, параллельной плоскости Oxy, имеющей уравнение $z = h$. Получим

$$\frac{x^2}{p} + \frac{y^2}{q} = 2h.$$

Сразу заметим, что при $h < 0$ точек сечения нет, так как в этом случае равенство величин $\frac{x^2}{p} + \frac{y^2}{q} \geq 0$ и $2h < 0$ недостижимо. Отсюда следует, что весь эллиптический параболоид расположен в верхнем полупространстве, т. е. при $z \geq 0$.

При значении $h = 0$ это дает одну точку 0 (0; 0; 0).

При положительном значении $h > 0$ в сечении образуется эллипс:

$$\frac{x^2}{2ph} + \frac{y^2}{2qh} = 1$$

с полуосями $\sqrt{2ph}$ и $\sqrt{2qh}$. При неограниченном возрастании h полуоси эллипса тоже неограниченно возрастают.

При сечении этой поверхности плоскостями, параллельными плоскости Oxz, имеющими уравнение $y = h$, получим

$$\frac{x^2}{p}+\frac{h^2}{q}=2z\,, \text{ или } x^2=2p\left(z-\frac{h^2}{2q}\right).$$

Это есть уравнение параболы с параметром p, с осью симметрии, параллельной оси Oz, и координатами вершины $\left(0;\ h;\ \dfrac{h^2}{2q}\right)$. При этом ветви направлены вверх.

Аналогично при сечении эллиптического параболоида плоскостями, параллельными плоскости Oyz, имеющими уравнение $x = h$, получим параболы

$$y^2=2q\left(z-\frac{h^2}{2p}\right)$$

с осями симметрии, параллельными оси Oz, с параметром q и координатами вершин $\left(h;\ 0;\ \dfrac{h^2}{2p}\right)$. Ветви направлены вверх.

По аналогии с гиперболическим параболоидом можно представлять себе эллиптический параболоид как поверхность, которая получается при «скольжении» вершины параболы, описывающейся уравнением $y^2=2qz$, по параболе с уравнением $x^2=2pz$.

Эллиптический параболоид симметричен относительно координатных плоскостей Oxz и Oyz.

10.4.6. Конусы второго порядка

Определение. ***Конусом второго порядка*** называется поверхность, имеющая в некоторой прямоугольной системе координат уравнение вида

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0,$$

где $a\ge b>0,\ c>0$.

Эта система координат называется ***канонической*** относительно данного конуса, а уравнение – ***каноническим уравнением*** данного конуса.

Легко видеть, что плоскость $z = h$ при $h \neq 0$ пересекает конус по эллипсу с полуосями $\frac{a}{c}|h|$ и $\frac{b}{c}|h|$, которые при возрастании h тоже возрастают.

Плоскость $z = 0$ пересекает конус в точке 0 (0, 0, 0).

Плоскость $y = h$ при $h \neq 0$ пересекает конус по гиперболе с полуосями

$$\frac{c}{b}|h| \text{ и } \frac{a}{b}|h|,$$

которые при неограниченном возрастании h тоже неограниченно возрастают.

Плоскость $x = h$ при $h \neq 0$ пересекает конус по гиперболе с полуосями

$$\frac{c}{b}|h \quad| \text{ и } \frac{b}{a}|h \quad|,$$

которые при возрастании h тоже возрастают.

Плоскости $y = 0$ и $x = 0$ пересекают конус по паре пересекающихся прямых (рис. 58).

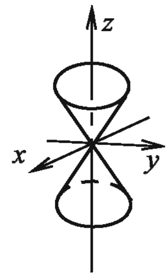


Рис. 58

10.4.7. Цилиндры второго порядка

Определение. Поверхность называется:

а) ***эллиптическим цилиндром***;

б) ***параболическим цилиндром***;

в) ***гиперболическим цилиндром***,

если в некоторой прямоугольной системе координат, называемой ***канонической*** для данной поверхности, она имеет следующее ***каноническое уравнение***:

а) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a \geq b > 0;$

б) $y^2 = 2px, \quad h > 0;$

в) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad a > 0, \ b > 0.$

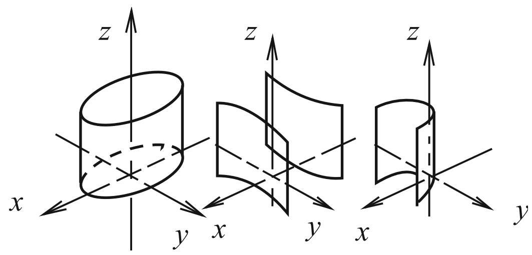


Рис. 59

Так как в эти уравнения не входит координата z, то вместе с некоторой точкой (x_0, y_0, z_0) им удовлетворяет и любая точка вида (x_0, y_0, z). Это означает, что вместе с точкой M_0 каждый из цилиндров содержит всю прямую, проходящую через эту точку параллельно оси Oz, т. е. цилиндр получается из «скольжения» кривой, лежащей в плоскости Oxy, вдоль оси Oz (рис. 59).

10.5. КЛАССИФИКАЦИЯ КРИВЫХ ВТОРОГО ПОРЯДКА

10.5.1. Предварительные замечания. Вспомним, как мы приводили общее уравнение линии второго порядка к каноническому виду в п. 10.3.2. Все началось с отыскания собственных чисел λ_1 и λ_2 матрицы квадратичной формы $ax^2 + 2bxy + cy^2$. *Сразу оговоримся, что случай* $\lambda_1 = \lambda_2 = 0$ *мы не будем рассматривать*, так как он означает, что $a = b = c = 0$, т. е. что уравнение п. 10.1.1 не является уравнением линии второго порядка. Значит, если произведение $\lambda_1 \lambda_2$ равно нулю, то только одно из собственных чисел равно нулю, а другое – не равно нулю. При преобразовании общего уравнения линии второго порядка мы можем делать поворот системы координат, параллельный перенос, смену координатных осей Ox на Oy и Oy на Ox и смену ориентации на осях.

После таких оговорок можно рассмотреть схему возникновения разных случаев при приведении общего уравнения кривых второго порядка к каноническому виду.

10.5.2. Типы уравнений кривых второго порядка. Возникает девять разных случаев. Рассмотрим их по порядку. Сначала заметим, что кривые второго порядка после приведения их уравнений к каноническому виду в зависимости от обращения в нуль некоторых коэффициентов принято делить на типы. Так, кривые, чьи уравнения приводятся к виду

$$a_{11}x^2 + a_{22}y^2 + a_{33} = 0\text{, где } a_{11} \neq 0,\ \ a_{22} \neq 0,$$

называются кривыми I типа, кривые, чьи уравнения приводятся к виду

$$a_{22}y^2 + 2a_{13}x = 0\text{, где } a_{22} \neq 0,\quad a_{13} \neq 0,$$

называются кривыми II типа, и, наконец, кривые, уравнения которых приводятся к виду

$$a_{22}y^2 + a_{33} = 0\text{, где } a_{22} \neq 0,$$

называются кривыми III типа. Можно показать, что всякое общее уравнение линии второго порядка приводится к уравнению только одного из трех типов, и не может быть приведено к уравнению двух или трех типов одновременно.

$\lambda_1\lambda_2 \neq 0$	I	$\lambda_1\lambda_2 < 0$	$a_{11}x^2 - a_{22}y^2 + a_{33} = 0$	$a_{33} \neq 0$	1
				$a_{33} = 0$	2
		$\lambda_1\lambda_2 > 0$	$a_{11}x^2 + a_{22}y^2 + a_{33} = 0$	$a_{33} < 0$	3
				$a_{33} = 0$	4
				$a_{33} > 0$	5
$\lambda_1\lambda_2 = 0$	II	$a_{22}y^2 + 2a_{13}x = 0$			6
	III	$a_{22}y^2 + a_{33} = 0$	$a_{33} \neq 0$	$a_{22}a_{33} < 0$	7
				$a_{22}a_{33} > 0$	8
			$a_{33} = 0$		9

Переходим к рассмотрению случаев 1–9.

1. Этот случай принято называть ***гиперболическим невырожденным***. Умножив, если нужно, уравнения на (–1) и переименовав, если нужно, переменные, считаем, что $a_{11} > 0$, $a_{22} > 0$, $a_{33} < 0$. Тогда положим

$$a^2 = -\frac{a_{33}}{a_{11}}, \quad b^2 = -\frac{a_{33}}{a_{22}}$$

и получим

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad a > 0,\ b > 0,$$

т. е. каноническое уравнение гиперболы.

2. Этот случай называется ***гиперболическим вырожденным***. Можем считать, что $a_{11} \le a_{22}$. Положим

$$a^2 = \frac{a_{11} + a_{22}}{a_{11}}, \quad b^2 = \frac{a_{11} + a_{22}}{a_{22}}$$

и получим

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0,$$

где $a \ge b > 0$ и $\frac{1}{a^2} + \frac{1}{b^2} = 1$. Этому уравнению удовлетворяют точки двух прямых

$$\frac{x}{a} + \frac{y}{b} = 0 \ \text{ и } \ \frac{x}{a} - \frac{y}{b} = 0,$$

пересекающихся в точке (0; 0), т. е. в этом случае линия второго порядка состоит из двух пересекающихся прямых.

3. Этот случай называется ***эллиптическим действительным невырожденным***. Считаем, что

$$a_{11} > 0, \quad a_{22} > 0, \quad a_{11} \le a_{22}.$$

Тогда положим

$$a^2=-\frac{a_{33}}{a_{11}},\qquad b^2=-\frac{a_{33}}{a_{22}}$$

и получим

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\ ,\qquad a\ge b>0\,,$$

т. е. каноническое уравнение эллипса.

4. Случай ***эллиптический вырожденный***. Положим

$$a^2=\frac{a_{11}+a_{22}}{a_{11}},\qquad b^2=\frac{a_{11}+a_{22}}{a_{22}}$$

и получим

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=0\,,$$

где $a\ge b>0$ и $\frac{1}{a^2}+\frac{1}{b^2}=1$. Этому уравнению удовлетворяет только точка (0; 0).

5. Случай ***эллиптический мнимый невырожденный***. Положим

$$a^2=\frac{a_{33}}{a_{11}},\qquad b^2=\frac{a_{33}}{a_{22}}$$

и получим

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=-1\,.$$

В действительной плоскости это уравнение определяет пустое множество.

6. Случай ***параболический невырожденный***. Будем считать, что $a_{22}>0,\quad a_{13}<0$. Положив $p=-\frac{a_{13}}{a_{22}}$, получим уравнение $y^2=2px,\ \ p>0,$ т. е. каноническое уравнение параболы.

Случаи 7, 8 и 9 называются ***параболическими вырожденными***.

7. Положим $b^2=-\frac{a_{33}}{a_{22}}$ и получим $y^2-b^2=0,\ b>0$. Этому уравнению удовлетворяют точки двух параллельных прямых $y+b=0,\ y-b=0$.

8. Положим $b^2 = \frac{a_{33}}{a_{22}}$ и получим $y^2 + b^2 = 0, \quad b > 0$. В действительной плоскости это уравнение определяет пустое множество.

9. Сократим на коэффициент a_{22} и получим $y^2 = 0$. Это уравнение определяет ту же прямую (ось Ox), что и уравнение $y = 0$. Чтобы отметить различие между ними, будем говорить, что уравнение $y^2 = 0$ определяет пару совпадающих прямых.

10.5.3. Таблица типов кривых второго порядка

Тип линии	Вид линии	Невырожденные линии	Вырожденные линии
I	Гиперболический	Гипербола $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	Пара пересекающихся прямых $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$
	Эллиптический	а) Эллипс действительный $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	Одна точка (0, 0) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$
		б) Эллипс мнимый $\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$ (пустое множество)	
II	Параболический	Парабола $y^2 = 2px, \ p > 0$	а) Пара параллельных различных прямых б) Пара совпадающих прямых $y^2 = 0$ в) Пустое множество $y^2 + b^2 = 0, \ b > 0$

Нетрудно видеть, что эта таблица полностью соответствует приведенной выше первой таблице.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Беклемишев Д. В.* Курс аналитической геометрии и линейной алгебры / Д. В. Беклемишев. – Москва : Физматлит, 2009. – 312 с.

2. *Бронштейн И. Н.* Справочник по математике для инженеров и учащихся втузов / И. Н. Бронштейн, К. А. Семендяев. – Москва : Лань, 2010. – 608 с.

3. *Воеводин В. В.* Линейная алгебра / В. В. Воеводин. – Москва : Лань, 2009. – 400 с.

4. *Ивлева А. М.* Линейная алгебра. Аналитическая геометрия / А. М. Ивлева, П. И. Прилуцкая, И. Д. Черных. – Новосибирск : Изд-во НГТУ, 2019. – 182 с.

5. *Ильин В. А.* Линейная алгебра / В. А. Ильин, Э. Г. Позняк. – Москва : Физматлит, 2010. – 280 с.

6. *Клетеник Д. В.* Сборник задач по аналитической геометрии / Д. В. Клетеник. – Изд. 17-е, стер. – Санкт-Петербург : Лань, 2010. – 223 с.

7. *Кострикин А. И.* Введение в алгебру. Ч. 2. Линейная алгебра / А. И. Кострикин. – Москва : Физматлит, 2000. – 368 с.

8. *Курош А. Г.* Курс высшей алгебры / А. Г. Курош. – Санкт-Петербург : Лань, 2008. – 431 с.

9. *Мальцев А. И.* Основы линейной алгебры / А. И. Мальцев. – Санкт-Петербург : Лань, 2009. – 480 с.

10. *Постников М. М.* Аналитическая геометрия / М. М. Постников. – Москва : Лань, 2009. – 416 с.

11. *Постников М. М.* Линейная алгебра / М. М. Постников. – Москва : Лань, 2009. – 400 с.

12. *Проскуряков И. В.* Сборник задач по линейной алгебре / И. В. Проскуряков. – Москва : Бином, 2010. – 383 с.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

УЧЕБНОЕ ИЗДАНИЕ

Ивлева Ася Михайловна
Пинус Александр Георгиевич
Чехонадских Александр Васильевич

ОСНОВЫ АЛГЕБРЫ
И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Учебник

2-е издание

Редактор *И.Л. Кескевич*
Выпускающий редактор *И.П. Брованова*
Художественный редактор *А.В. Ладыжская*
Компьютерная верстка *С.И. Ткачева*

Подписано в печать 01.06.2021
Формат 70 × 100 1/16. Бумага офсетная
Уч.-изд. л. 23,22. Печ. л. 18,0
Тираж 3000 экз. (1-й з-д – 1–40 экз.)
Изд. № 75. Заказ № 609

Налоговая льгота – Общероссийский классификатор продукции
Издание соответствует коду 95 3000 ОК 005-93 (ОКП)

Издательство Новосибирского государственного
технического университета
630073, г. Новосибирск, пр. К. Маркса, 20
Тел. (383) 346-31-87
E-mail: office@publish.nstu.ru

Отпечатано в типографии
Новосибирского государственного технического университета
630073, г. Новосибирск, пр. К. Маркса, 20