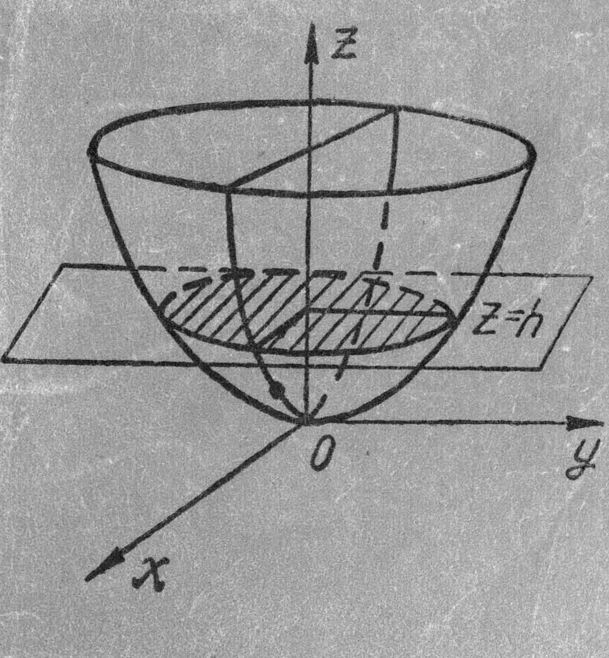
В. Р. КАЙГОРОДОВ

КУРС аналитической геометрии и линейной алгебры



В. Р. КАЙГОРОДОВ

КУРС АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ И ЛИНЕЙНОЙ АЛГЕБРЫ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов университетов, обучающихся по специальности „Физика"

ИЗДАТЕЛЬСТВО
КАЗАНСКОГО УНИВЕРСИТЕТА
1985

Печатается по постановлению
Редакционно-издательского совета
Казанского университета

Рецензенты: проф. Позняк Э. Г. (МГУ)
проф. Базылев В. Т. (МГПИ)

Учебное пособие представляет собой обработанные автором лекции, читаемые на физическом факультете Казанского государственного университета. Отличительной особенностью его является преимущественное изложение материала по аналитической геометрии в векторной форме и использование языка линейных операторов в алгебре, в нем подробно разобраны решения задач, типичных для каждой излагаемой темы.

Пособие полностью соответствует ныне действующей программе курса аналитической геометрии и линейной алгебры для студентов первого курса физических и радиофизических специальностей университетов.

$К \frac{1704020000 - 044}{075(02) - 85}$ КБ-49-005-84

ПРЕДИСЛОВИЕ

Материал, излагаемый ниже, представляет собой годовой курс лекций, прочитанный автором на первом курсе физического факультета Казанского университета в 1981—1983 учебных годах, рассчитан как для аудиторного обучения, так и для самообразования.

Выпуск данного пособия обусловлен тем, что согласно последним учебным планам 2 отдельных ранее курса „Аналитическая геометрия" и „Высшая алгебра" объединены в один годовой курс с сокращением числа учебных часов. Кроме того, методы линейной алгебры находят все большее приложение при рассмотрении прикладных вопросов и студенты 1-го курса физического факультета сразу же привлекаются к работе на ЭВМ при рассмотрении систем линейных алгебраических уравнений. Все это потребовало пересмотра последовательности традиционного изложения материала и начать курс с теории систем линейных уравнений, ввести сразу же такие понятия из алгебры, как матрица, детерминант (определитель) матрицы, ранг матрицы и т. д., что позволило использовать эти понятия при изложении ряда вопросов из курса аналитической геометрии.

Автор сознательно сдвинул в середину курса введение аксиоматического определения линейного векторного пространства, исподволь приближая студентов к этому абстрактному понятию через вопросы, разбираемые в теории систем линейных уравнений и при исследовании линейной зависимости векторов в аналитической геометрии.

Необходимость подготовить материал для прохождения темы „Геометрические и физические приложения определенных интегралов" в курсе математического анализа обусловливает проведение исследования общего уравнения кривых второго порядка и канонических уравнений поверхностей второго порядка в первом семестре, не прибегая к методам линейной алгебры. К более строгому и последовательному изложению этих вопросов автор возвращается в курсе линейной алгебры.

Данный курс читается параллельно с курсом математического анализа, где на первых лекциях затрагивается теория множеств, дается определение числового поля, подробно исследуются свойства полей вещественных и комплексных чисел. Поэтому автор не останавливается на этих понятиях и под числовым полем **K** подразумевается либо поле вещественных чисел **R**, либо поле комплексных чисел **C**. На протяжении курса об этом будет каждый раз указываться отдельно.

Студенты первого курса обычно испытывают затруднения при применении теории к решению задач. На каждую тему в книге приводится решение 1—2 примеров с подробным объяснением хода рассуждений.

Лекционный материал в порядке изложения разбивается на следующие разделы:

I. Системы линейных алгебраических уравнений.

II. Векторная алгебра в трехмерном пространстве.

III. Задачи аналитической геометрии на плоскости и в пространстве.

IV. Линейные пространства и линейные отображения.

V. Билинейные и квадратичные формы.

VI. Аффинные, евклидовы и унитарные пространства.

Читатели, заинтересованные в более детальном и глубоком изучении курса, могут обратиться к литературе, указанной в конце книги.

Лекция 1

ТЕРМИНОЛОГИЯ, ОПРЕДЕЛЕНИЯ И ЗАДАЧИ ТЕОРИИ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ. ДЕТЕРМИНАНТ (ОПРЕДЕЛИТЕЛЬ) КВАДРАТНОЙ МАТРИЦЫ

В первых лекциях излагаются вопросы, связанные с теорией систем линейных алгебраических уравнений. Все результаты справедливы как для поля вещественных чисел **R**, так и для поля комплексных чисел **C**. Ниже для числового поля употребляется общий символ **K**.

Под системой m линейных уравнений с n неизвестными понимается m рассматриваемых в совокупности соотношений вида

$$\left.\begin{array}{l} a_1^1 x^1 + a_2^1 x^2 + \ldots + a_n^1 x^n = b^1 \\ a_1^2 x^1 + a_2^2 x^2 + \ldots + a_n^2 x^n = b^2 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_1^m x^1 + a_2^m x^2 + \ldots + a_n^m x^n = b^m \end{array}\right\} \qquad (1.1)$$

или в краткой форме

$$\sum_{j=1}^{n} a_j^i x^j = b^i \quad (i = \overline{1, m}), \qquad (1.2)$$

где $a_j^i \in \mathbf{K}$ — *коэффициенты системы*, $b^i \in \mathbf{K}$ — *свободные члены*, $x^j \in \mathbf{K}$ — *неизвестные*, подлежащие определению.

Таблица чисел

$$A = \begin{bmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ a_1^m & a_2^m & \ldots & a_n^m \end{bmatrix} \qquad (1.3)$$

называется *матрицей* системы уравнений. Она имеет m строк и n столбцов. Если $m = n$, то матрица называется *квадратной*, если же $m \neq n$ — *прямоугольной*. Числа a_j^i называются *элементами* матрицы.

Матрица, содержащая только одну строку, называется *матрицей-строкой* (с n элементами)

$$A = [a_1, a_2, \ldots, a_n]. \qquad (1.4)$$

Матрица, содержащая один столбец, называется *матрицей-столбцом* (с m элементами)

$$A=\begin{bmatrix} a^1 \\ a^2 \\ \vdots \\ a^m \end{bmatrix}. \tag{1.5}$$

Если в (1.5) все $a^i=0$, то имеем *нуль-столбец*, обозначаемый символом 0. Два столбца

$$A=\begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}, \quad B=\begin{bmatrix} b^1 \\ \vdots \\ b^m \end{bmatrix} \tag{1.6}$$

считаем *равными* (пишем $A=B$), есл и $a^i=b^i$ $(i=\overline{1, m})$. Если хотя бы одно из чисел a^k (k — фиксировано) не совпадает с соответствующим числом b^k, то столбцы считаются *различными* $(A \neq B)$. Под *суммой* столбцов A и B (пишем $A+B$) понимаем столбец

$$\begin{bmatrix} a^1+b^1 \\ a^2+b^2 \\ \vdots \\ a^m+b^m \end{bmatrix}. \tag{1.7}$$

Под *умножением* столбца A на число $\lambda \in \mathbf{K}$ (пишем λA) понимаем столбец вида

$$\begin{bmatrix} \lambda a^1 \\ \lambda a^2 \\ \vdots \\ \lambda a^m \end{bmatrix}. \tag{1.8}$$

Введенные операции позволяют рассматривать *линейные комбинации* матриц-столбцов. Пусть имеем k столбцов

$$A_1=\begin{bmatrix} a_1^1 \\ \vdots \\ a_1^m \end{bmatrix}, \quad A_2=\begin{bmatrix} a_2^1 \\ \vdots \\ a_2^m \end{bmatrix}, \ldots, A_k=\begin{bmatrix} a_k^1 \\ \vdots \\ a_k^m \end{bmatrix}.$$

Линейной комбинацией столбцов $A_1, \dots, A_k$ с коэффициентами $\lambda_1, \dots, \lambda_k$ (пишем $\lambda_1 A_1 + \lambda_2 A_2 + \dots + \lambda_k A_k$) называется столбец вида

$$\begin{bmatrix} \sum_{j=1}^{k} \lambda_j a_j^1 \\ \vdots \\ \sum_{j=1}^{k} \lambda_j a_j^m \end{bmatrix}. \tag{1.9}$$

Столбцы $A_1, \dots, A_k$ называются *линейно-зависимыми*, если существуют такие числа $\lambda_1, \dots, \lambda_k$, среди которых хотя бы одно отлично от нуля, что выполнено

$$\lambda_1 A_1 + \lambda_2 A_2 + \dots + \lambda_k A_k = 0 \ (0 \text{ — нуль-столбец}).$$

Если равенство $\lambda_1 A_1 + \lambda_2 A_2 + \dots + \lambda_k A_k = 0$ выполнено тогда и только тогда, когда $\lambda_1 = 0, \lambda_2 = 0, \dots, \lambda_k = 0$, то система матриц-столбцов $\{A_1, \dots, A_k\}$ называется *линейно-независимой.*

Например, система столбцов

$$A_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 2 \end{bmatrix}$$

является линейно-зависимой, так как $A_1 + 2A_2 - A_3 = 0$, а система столбцов

$$A_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

является линейно-независимой, поскольку из соотношения $\lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3 = 0$ на основании равенства матриц-столбцов следует

$$\left.\begin{aligned} \lambda_1 + \lambda_2 &= 0 \\ 2\lambda_2 + \lambda_3 &= 0 \\ \lambda_3 &= 0 \end{aligned}\right\} \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Очевидно, если между столбцами $A_1, \dots, A_k$ $(k > 1)$ существует линейная зависимость, то *один из этих столбцов является линейной комбинацией других.* В самом деле, пусть выполнено $\lambda_1 A_1 + \lambda_2 A_2 + \dots + \lambda_k A_k = 0$, где, например, $\lambda_1 \neq 0$. Тогда

$$A_1 = \left(-\frac{\lambda_2}{\lambda_1}\right) A_2 + \dots + \left(-\frac{\lambda_k}{\lambda_1}\right) A_k.$$

Обратно, *если какой-либо столбец есть линейная комбинация других, то данная система столбцов линейно-зависима.*

Действительно, пусть $A_1 = \mu_2 A_2 + \dots + \mu_k A_k$. Тогда имеем

$$(-1) A_1 + \mu_2 A_2 + \dots + \mu_k A_k = 0,$$

где первый коэффициент отличен от нуля.

Всё приведенное выше для матриц-столбцов тривиальным образом переносится и для матриц-строк, если ввести в рассмотрение нуль-строку $0 = [0, 0, \dots, 0]$, сумму $A + B = [a_1 + b_1, \dots, a_n + b_n]$, умножение на число $\lambda A = [\lambda a_1, \dots, \lambda a_n]$.

Вернемся к рассмотрению систем вида (1.1). *Решением системы* (1.1) называется матрица-столбец

$$X = \begin{bmatrix} c^1 \\ c^2 \\ \vdots \\ c^n \end{bmatrix}$$

если после подстановки чисел $c^1, \dots, c^n$, вместо неизвестных $x^1, \dots, x^n$ соотношения в (1.1) обращаются в тождества. При этом решения

$$X_1 = \begin{bmatrix} c_1^1 \\ \vdots \\ c_1^n \end{bmatrix}, \quad X_2 = \begin{bmatrix} c_2^1 \\ \vdots \\ c_2^n \end{bmatrix}$$

считаются *различными*, если $X_1 \neq X_2$.

Система называется *определенной*, если она имеет *единственное решение*, и *неопределенной*, если имеются по крайней мере два различных решения. В случае отсутствия решений система называется *несовместной*, при наличии решений — *совместной*.

Например, система

$$\begin{cases} x^1 - x^2 - x^3 = 0 \\ x^1 + x^2 + x^3 = 0 \end{cases}$$

является неопределенной, так как

$$X_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

— два различных решения указанной системы.

Система уравнений

$$\begin{cases} x^1 + x^2 - 2x^3 = 1 \\ -x^1 - x^2 + 2x^3 = 1 \end{cases}$$

является несовместной системой, так как правые части обоих уравнений одинаковы, а левые отличаются знаком.

В теории систем линейных уравнений рассматривают три основные задачи: (1). Выяснить, является ли система совместной или несовместной. (2). Если система совместна, то является ли она определенной или неопределенной. В случае определенной системы необходимо указать способ нахождения ее единственного решения. (3). В случае неопределенной системы описать все множество ее решений.

При решении этих задач важную роль играют два понятия: (1) детерминант (определитель) квадратной матрицы и (2) ранг матрицы.

Остановимся на первом понятии. Мы сформулируем правило, согласно которому каждой квадратной матрице $A = [a_j^i]$ $(i, j = \overline{1, n})$ сопоставим число, называемое *детерминантом* (*определителем* n-го порядка) матрицы. Это число обозначают символом $\det A$, а сам определитель обозначают так:

$$\det A = \begin{vmatrix} a_1^1 & a_2^1 \ldots a_n^1 \\ a_1^2 & a_2^2 \ldots a_n^2 \\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ a_1^n & a_2^n \ldots a_n^n \end{vmatrix}.$$

Термин *вычислить* (*раскрыть*) определитель означает указать данное число, следуя сформулированному ниже определению детерминанта с использованием свойств определителей, изложенных во второй лекции.

Для введения детерминанта используем *альтернатор* $\delta^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k}$ $(i_1, \ldots, i_k, j_1, \ldots, j_k = \overline{1, n})$. Он определяется следующими условиями: $\delta^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k} = \pm 1$, если $j_1, j_2, \ldots, j_k$ есть некоторая перестановка значений индексов $i_1, i_2, \ldots, i_k$, считая, что все эти значения различны; при этом берется $+1$, если перестановка *четная*, и -1, если она *нечетная*. Во всех остальных случаях $\delta^{i_1 \ldots i_k}_{j_1 \ldots j_k} = 0$ (т. е. если среди значений $i_1, i_2, \ldots, i_k$ или среди значений $j_1, j_2, \ldots, j_k$ есть одинаковые, а также если среди значений $i_1, \ldots, i_k$ есть такие, каких нет среди $j_1, \ldots, j_k$ и наоборот).

Так, при $k = 1$ имеем альтернатор $\delta^i_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$, называемый **символом Кронекера** (очень часто используемым в последующих лекциях).

Определение. *Детерминантом* (*определителем*) матрицы A называется число:

$$\det A = \sum_{i_1, \ldots, i_n = 1}^{n} \delta^{1\,2 \ldots n}_{i_1 i_2 \ldots i_n} a^{i_1}_1 a^{i_2}_2 \ldots a^{i_n}_n .$$

Поскольку существует $n!$ перестановок, то, согласно определению альтернатора, *определитель есть сумма $n!$ произведений элементов матрицы, выбранных по одному из каждого столбца на различных строках и взятых со знаком плюс в случае четной перестановки номеров строк и со знаком минус в случае нечетной перестановки.*

Пример вычисления определителя второго порядка

$$\begin{vmatrix} a^1_1 & a^1_2 \\ a^2_1 & a^2_2 \end{vmatrix} = \sum_{i_1, i_2 = 1}^{2} \delta^{1\,2}_{i_1 i_2} a^{i_1}_1 a^{i_2}_2 =$$

$$= \delta^{1\,2}_{1\,1} a^1_1 a^1_2 + \delta^{1\,2}_{1\,2} a^1_1 a^2_2 + \delta^{1\,2}_{2\,1} a^2_1 a^1_2 + \delta^{1\,2}_{2\,2} a^2_1 a^2_2 .$$

По определению альтернатора $\delta^{12}_{11}=0$, $\delta^{12}_{22}=0$, $\delta^{12}_{12}=1$, $\delta^{12}_{21}=-1$. Поэтому для любого определителя второго порядка имеем

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb. \tag{1.10}$$

Пример вычисления определителя третьего порядка (члены с нулевым альтернатором не пишем)

$$\begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} = \sum_{i_1, i_2, i_3 = 1}^{3} \delta^{123}_{i_1 i_2 i_3} a^{i_1}_1 a^{i_2}_2 a^{i_3}_3 =$$

$$= \delta^{123}_{123} a_1^1 a_2^2 a_3^3 + \delta^{123}_{132} a_1^1 a_2^3 a_3^2 + \delta^{123}_{213} a_1^2 a_2^1 a_3^3 +$$
$$+ \delta^{123}_{231} a_1^2 a_2^3 a_3^1 + \delta^{123}_{312} a_1^3 a_2^1 a_3^2 + \delta^{123}_{321} a_1^3 a_2^2 a_3^1 .$$

Используя определение альтернатора, приходим к правилу Саррюса вычисления определителя 3-го порядка

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} \begin{matrix} a & b \\ d & e \\ g & h \end{matrix} = aek + bfg + cdh - gec - hfa - kdb.$$

Лекция 2

СВОЙСТВА ОПРЕДЕЛИТЕЛЕЙ. АЛГЕБРАИЧЕСКОЕ ДОПОЛНЕНИЕ И МИНОР ЭЛЕМЕНТА ОПРЕДЕЛИТЕЛЯ

Рассмотрим матрицу из m строк и n столбцов

$$A = \begin{bmatrix} a_1^1 & a_2^1 \dots & a_n^1 \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_1^m & a_2^m \dots & a_n^m \end{bmatrix}.$$

Ей можно сопоставить матрицу из n строк и m столбцов, у которой каждая строка является столбцом матрицы A с тем же номером. Эту матрицу

$$B = \begin{bmatrix} a_1^1 & a_1^2 \dots a_1^m \\ a_2^1 & a_2^2 \dots a_2^m \\ \dots & \dots \\ a_n^1 & a_n^2 \dots a_n^m \end{bmatrix}$$

называют *транспонированной* матрицей и обозначают символом A^{tr}, а переход от A к B — *транспонированием*. В случае квадратных матриц транспонирование можно определить как поворот вокруг главной диагонали.

Элементы b_j^i транспонированной матрицы связаны с элементами матрицы A соотношением $b_j^i = a_i^j$ $(i = \overline{1, n};\ j = \overline{1, m})$.

Свойство 1. При транспонировании квадратной матрицы значение определителя не меняется, т. е. $\det A^{tr} = \det A$.

В самом деле,

$$\det A^{tr} = \sum \delta^{1\,2\,\dots\,n}_{i_1 i_2 \dots i_n} b_1^{i_1} b_2^{i_2} \dots b_n^{i_n} =$$

$$= \sum \delta^{1\,2\,\dots\,n}_{i_1 i_2 \dots i_n} a^1_{i_1} a^2_{i_2} \dots a^n_{i_n}.$$

Согласно определению альтернатора $\delta^{1\,2\,\dots\,n}_{i_1 i_2 \dots i_n} = \delta^{i_1 i_2 \dots i_n}_{1\,2\,\dots\,n}$.

Вследствие этого $\det A^{tr} = \sum \delta^{i_1 \dots i_n}_{1 \dots n} a^1_{i_1} \dots a^n_{i_n}$. Мы получили сумму $n!$ тех же произведений элементов (с измененным порядком перемножения) матрицы A и взятых с тем же знаком, что при вычислении суммы .

$$\sum \delta^{1\,2\,\dots\,n}_{i_1 i_2 \dots i_n} a_1^{i_1} a_2^{i_2} \dots a_n^{i_n} = \det A.$$

Это означает, что $\det A^{tr} = \det A$.

Доказанное предложение позволяет формулировать и доказывать свойства определителей с использованием лишь столбцов определителя. Автоматически все переносится и на его строки.

Свойство 2. При перестановке двух столбцов (строк) в матрице ее детерминант меняет знак.

Сначала рассмотрим тот случай, когда переставлены соседние столбцы (например, s-й и $(s+1)$-й). До перестановки имеем:

$$\sum \delta^{1 \ldots s\ s+1 \ldots n}_{i_1 \ldots i_s\ i_{s+1} \ldots i_n} a^{i_1}_1 \ldots a^{i_s}_s a^{i_{s+1}}_{s+1} \ldots a^{i_n}_n .$$

После перестановки:

$$\sum \delta^{1 \ldots s+1\ s \ldots n}_{i_1 \ldots i_s\ i_{s+1} \ldots i_n} a^{i_1}_1 \ldots a^{i_s}_{s+1} a^{i_{s+1}}_s \ldots a^{i_n}_n .$$

Но

$$\delta^{1 \ldots s+1\ s \ldots n}_{i_1 \ldots i_s i_{s+1} \ldots i_n} = -\delta^{1 \ldots s\ s+1 \ldots n}_{i_1 \ldots i_s\ i_{s+1} \ldots i_n} .$$

Свойство 2 для рассматриваемого случая доказано.

Обратимся к более общему случаю, когда меняются местами s-й и l-й столбец, между которыми находится k столбцов. Такую перестановку можно осуществить последовательными перестановками s-го столбца с соседними столбцами, двигаясь к l-му столбцу на его место. Понадобится $k+1$ перестановок. Передвигая после этого аналогичным образом l-й столбец на место s-го, произведем еще k перестановок. Определитель получит множитель $(-1)^{2k+1}$, т. е. изменит свой знак на противоположный.

Следствие свойства 2. Определитель, имеющий 2 одинаковых столбца (строки), равен нулю.

В самом деле, переставляя одинаковые столбцы, мы, с одной стороны, не изменим определителя, а, с другой, в силу свойства 2, он изменяет знак на противоположный, т. е. $\det A = -\det A$. Следовательно, $\det A = 0$.

Свойство 3. Если элементы s-го столбца матрицы A представляют собой линейную комбинацию вида $a^i_s = \lambda_1 a^i_{1s} + \lambda_2 a^i_{2s} + \ldots + \lambda_k a^i_{ks}$, то $\det A = \lambda_1 \det A_1 + \lambda_2 \det A_2 + \ldots + \lambda_k \det A_k$, где матрицы $A_1, \ldots, A_k$ получаются из матрицы A путем замены s-го столбца на элементы $a^i_{1s}, a^i_{2s}, \ldots, a^i_{ks}$.

В самом деле,

$$\det A = \sum \delta^{1 \ldots s \ldots n}_{i_1 \ldots i_s \ldots i_n} a^{i_1}_1 \ldots a^{i_s}_s \ldots a^{i_n}_n =$$

$$= \sum \delta^{1 \ldots s \ldots n}_{i_1 \ldots i_s \ldots i_n} a^{i_1}_1 \ldots (\lambda_1 a^{i_s}_{1s} + \lambda_2 a^{i_s}_{2s} + \ldots + \lambda_k a^{i_s}_{ks}) \ldots a^{i_n}_n =$$

$$= \lambda_1 \sum \delta^{1 \ldots s \ldots n}_{i_1 \ldots i_s \ldots i_n} a^{i_1}_1 \ldots a^{i_s}_{1s} \ldots a^{i_n}_n + \ldots +$$

$$+\lambda_k \sum \delta^{1\ldots s\ldots n}_{i_1\ldots i_s\ldots i_n} a^{i_1}_1 \ldots a^{i_s}_{ks} \ldots a^{i_n}_n .$$

Свойство 3 доказано.

Следствие свойства 3. Общий множитель у всех элементов столбца (строки) можно вынести за знак определителя, т. е. если $a^i_s = \lambda \overset{*}{a}{}^i_s$, то $\det A = \lambda \det \overset{*}{A}$.

Свойство 4. Если какой-либо столбец (строка) есть линейная комбинация других столбцов (строк), то определитель равен нулю.

Действительно, применяя свойство 3, мы обнаружим в определителях матриц $A_1, \ldots, A_k$ одинаковые столбцы, и в силу следствия свойства 2 их детерминанты равны нулю, что и доказывает свойство 4.

Свойство 5. Определитель не изменится, если к элементам столбца (строки) прибавить соответствующие элементы другого столбца (строки), умноженного на какое-либо число.

В самом деле, пусть к s-му столбцу матрицы A прибавляем l-й, умноженный на λ. Во вновь полученной матрице $\overset{*}{A}$, s-й столбец будет состоять из элементов: $\overset{*}{a}{}^i_s = a^i_s + \lambda a^i_l$. По свойству 3 $\det \overset{*}{A} = \det A + \lambda \det A_1$, где у матрицы A_1 s-й и l-й столбцы совпадают. Следовательно, $\det A_1 = 0$ и $\det \overset{*}{A} = \det A$.

Пример. Требуется вычислить определитель

$$\Delta = \begin{vmatrix} 1 & a_1 & a_2 \ldots a_n \\ 1 & x_1 & a_2 \ldots a_n \\ 1 & a_1 & x_2 \ldots a_n \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ 1 & a_1 & a_2 \ldots x_n \end{vmatrix}.$$

Умножая последовательно первый столбец на $a_n, a_{n-1}, \ldots, a_1$ и вычитая результат после каждого умножения из $(n+1)$-го, n-го и т. д. столбцов, получим по свойству 5

$$\Delta = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & x_1 - a_1 & 0 & \ldots & 0 \\ 1 & 0 & x_2 - a_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & \ldots & x_n - a_n \end{vmatrix}.$$

Вычитая первую строку из всех остальных, мы придем к определителю, у которого все элементы, за исключением главной диагонали (она останется неизменной), нули. Из всех $(n+1)!$ произведений не нулевым останется лишь одно произведение элементов по главной диагонали. В итоге $\Delta=(x_1-a_1)(x_2-a_2)\ldots$ $\ldots(x_n-a_n)$.

Важную роль в теории определителей играют понятия *алгебраического дополнения и минора* соответствующих фиксированному элементу a_s^k. Чтобы их ввести в рассмотрение, зафиксируем s-й столбец и представим определитель как сумму n слагаемых, вынося в качестве множителя элементы s-го столбца

$$\begin{aligned}\det A &= \sum \delta_{i_1 \ldots i_s \ldots i_n}^{1 \ldots s \ldots n} a_1^{i_1} \ldots a_s^{i_s} \ldots a_n^{i_n} = \\ &= a_s^1 \sum \delta_{i_1 \ldots 1 \ldots i_n}^{1 \ldots s \ldots n} a_1^{i_1} \ldots a_{s-1}^{i_{s-1}} a_{s+1}^{i_{s+1}} \ldots a_n^{i_n} + \\ &+ a_s^2 \sum \delta_{i_1 \ldots 2 \ldots i_n}^{1 \ldots s \ldots n} a_1^{i_1} \ldots a_{s-1}^{i_{s-1}} a_{s+1}^{i_{s+1}} \ldots a_n^{i_n} + \ldots + \\ &+ a_s^n \sum \delta_{i_1 \ldots n \ldots i_n}^{1 \ldots s \ldots n} a_1^{i_1} \ldots a_{s-1}^{i_{s-1}} a_{s+1}^{i_{s+1}} \ldots a_n^{i_n}.\end{aligned}$$

Обозначим соответственно суммы, стоящие после множителей $a_s^1, \ldots, a_s^n$ через $A_s^1, A_s^2, \ldots, A_s^n$, и назовем их *алгебраическими дополнениями* элементов $a_s^1, a_s^2, \ldots$ $\ldots, a_s^n$.

Итак,

$$\det A = a_s^1 A_s^1 + a_s^2 A_s^2 + \ldots + a_s^n A_s^n = \sum_{k=1}^{n} a_s^k A_s^k. \quad (2.1)$$

Формула (2.1) носит название *разложения определителя по элементам s-го столбца.*

Приведенное выше определение алгебраического дополнения A_s^k не имеет конструктивного характера, удобного для вычисления. Чтобы преодолеть эту трудность, дадим определение *минора* M_s^k элемента a_s^k и установим связь между алгебраическим дополнением A_s^k и минором M_s^k.

Определение. Если в определителе мысленно вычеркнуть s-й столбец и k-ую строку, то оставшийся определитель $(n-1)$-го порядка называется ***минором*** элемента a_s^k и обозначается символом M_s^k.

Теорема. *Имеет место следующая связь между алгебраическими дополнениями и соответствующими минорами:*

$$A_s^k = (-1)^{k+s} M_s^k. \qquad (2.2)$$

Докажем сначала теорему для частного случая, когда $k = s = 1$. С этой целью из выражения

$$\det A = \sum \delta_{i_1 i_2 \dots i_n}^{1\,2\,\dots\,n} a_1^{i_1} a_2^{i_2} \dots a_n^{i_n}$$

выберем все слагаемые, содержащие множителем a_1^1. Вынося a_1^1 за скобки, получим, что в скобках останется следующее выражение: $\sum \delta_{1\, i_2 \dots i_n}^{1\,2\,\dots\,n} a_2^{i_2} \dots a_n^{i_n}$. Но данная сумма есть (согласно определению детерминанта) определитель $(n-1)$-го порядка, полученный из исходного вычеркиванием 1-го столбца и 1-й строки, т. е. $A_1^1 = M_1^1$ (или $A_1^1 = (-1)^{1+1} M_1^1$).

Докажем сейчас (2.2) для любых k и s. Рассмотрим элемент $b = a_s^k$ и, переставляя последовательно s-й столбец и k-ую строку, переведем его в левый верхний угол определителя. Вновь полученный определитель обозначим через $\overset{*}{\Delta}$. Согласно свойству 2 получаем, что $\overset{*}{\Delta} = (-1)^{s-1+k-1}\Delta = (-1)^{k+s}\Delta$. В силу определения минора имеем, что $\overset{*}{M}{}_1^1 = M_s^k$. По вышедоказанному $\overset{*}{A}{}_1^1 = \overset{*}{M}{}_1^1 = M_s^k$. В силу равенства $\overset{*}{\Delta} = (-1)^{k+s}\Delta$ и согласно правилу выделения алгебраического дополнения A_s^k как суммы слагаемых с общим множителем a^{sk} получим, что $\overset{*}{A}{}_1^1 = (-1)^{k+s} A_s^k$, и, следовательно, имеем цепочку равенств $(-1)^{k+s} A_s^k = \overset{*}{A}{}_1^1 = \overset{*}{M}{}_1^1 = M_s^k$, откуда тотчас следует выполнение (2.2).

Пример. Вычислить алгебраическое дополнение $A(\alpha)$ элемента α в определителе

$$\Delta = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 2 & \alpha & 0 & \beta \\ -1 & 0 & -1 & 1 \end{vmatrix}. \qquad (2.3)$$

Согласно формуле (2.2) имеем

$$A(\alpha) = (-1)^{3+2}\begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{vmatrix} = (-1)^5\cdot 1\cdot 2\cdot 1 = -2.$$

Пример. Разложить определитель (2.3) по элементам третьего столбца и вычислить его

$$\Delta = 2\cdot(-1)^{2+3}\begin{vmatrix} 1 & 1 & 0 \\ 2 & \alpha & \beta \\ -1 & 0 & 1 \end{vmatrix} +$$

$$+(-1)\cdot(-1)^{4+3}\begin{vmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 2 & \alpha & \beta \end{vmatrix} = 4 - 2\alpha + \beta.$$

Наконец, приведем два следующих замечания.

Замечание 1. В силу свойства 1 имеет место аналогичная формуле (2.1) формула *разложения определителя по элементам строки*

$$\det A = \sum_{k=1}^{n} a_k^l A_k^l \quad (l \text{ — фиксировано!}). \tag{2.4}$$

Замечание 2. Пусть определитель имеет два одинаковых столбца (например, j-й и s-й, т. е. $a_j^k = a_s^k$ $(k = \overline{1, n})$). Он, как указано выше, равен нулю. Разложим его по элементам j-го столбца. Имеем

$$a_s^1 A_j^1 + a_s^2 A_j^2 + \ldots + a_s^n A_j^n = 0 \quad (j \neq s). \tag{2.5}$$

Этот результат можно сформулировать так. *Сумма произведений элементов какого-либо столбца (строки) на алгебраические дополнения соответствующих элементов другого столбца (строки) всегда равна нулю.*

Лекция 3

ПРАВИЛО КРАМЕРА. РАНГ МАТРИЦЫ. ТЕОРЕМА О БАЗИСНОМ МИНОРЕ

Перейдем к исследованию систем линейных уравнений. Прежде всего рассмотрим систему n уравнений с n неизвестными, когда определитель матрицы системы отличен от нуля. Наша цель состоит в доказательстве, что такая система уравнений всегда совместна и имеет единственное решение.

Пусть задана система уравнений $\sum_{j=1}^{n} a_j^i x^j = b^i$ $(i = \overline{1, n})$, где $A = [a_j^i]$ — матрица системы и $\det A \neq 0$.

Покажем, что система совместна. С этой целью рассмотрим набор чисел $c^s = \frac{\Delta_s}{\det A}$ $(s = \overline{1, n})$, где Δ_s есть определитель вида

$$\Delta_s = \begin{vmatrix} a_1^1 \dots b^1 \dots a_n^1 \\ a_1^2 \dots b^2 \dots a_n^2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_1^n \dots b^n \dots a_n^n \end{vmatrix}, \tag{3.1}$$

т. е. в матрице A s-й столбец заменен столбцом из свободных членов системы. Подставим числа c^s вместо неизвестных x^s в систему уравнений, заметив предварительно, что разложение детерминанта (3.1) по элементам s-го столбца дает следующее выражение: $\Delta_s = \sum_{k=1}^{n} b^k A_s^k$, где A_s^k — алгебраические дополнения элементов s-го столбца матрицы A. Имеем

$$\sum_{j=1}^{n} a_j^i c^j = \sum_{j=1}^{n} a_j^i \frac{\Delta_j}{\det A} = \frac{1}{\det A} \sum_{j=1}^{n} a_j^i \sum_{k=1}^{n} b^k A_j^k =$$

$$= \frac{1}{\det A} \sum_{k=1}^{n} b^k \sum_{j=1}^{n} a_j^i A_j^k. \tag{3.2}$$

Вспомнив формулу (2.1) и замечание 2 в конце второй лекции, мы получим, что

$$\sum_{j=1}^{n} a_j^i A_j^k = \begin{cases} \det A, & \text{если } i = k, \\ \ \ 0\ \ , & \text{если } i \neq k. \end{cases} \tag{3.3}$$

Поэтому при суммировании по k в правой части (3.2) останется лишь одно слагаемое, равное $b^i \cdot \det A$, т. е.

$$\sum_{j=1}^{n} a_j^i c^j = \frac{1}{\det A} \cdot b^i \cdot \det A = b^i$$

имеем тождественное выполнение уравнений системы.

Итак,

$$X = \begin{bmatrix} \dfrac{\Delta_1}{\det A} \\ \vdots \\ \dfrac{\Delta_n}{\det A} \end{bmatrix} \qquad (3.4)$$

есть решение системы (3.1).

Докажем, что оно единственное. Предположим, что имеется решение $x^s = \bar{c}^s$, отличное от (3.4). Имеем n тождеств $\sum\limits_{j=1}^{n} a_j^i \bar{c}^j \equiv b^i$ $(i = \overline{1, n})$. Умножим каждое тождество на алгебраические дополнения A_s^i (s — фиксировано!) и просуммируем их по индексу i от 1 до n. Имеем следующие тождества:

$$\sum_{i=1}^{n} A_s^i \sum_{j=1}^{n} a_j^i \bar{c}^j = \sum_{i=1}^{n} b^i A_s^i \quad (s = \overline{1, n})$$

или

$$\sum_{j=1}^{n} \bar{c}^j \sum_{i=1}^{n} a_j^i A_s^i = \Delta_s \quad (s = \overline{1, n}). \qquad (3.5)$$

В силу (2.1) и (2.5) вытекает, что $\bar{c}^s \det A = \Delta_s$ $(s = \overline{1, n})$.

Единственность решения системы (3.1) доказана. Оно всегда имеет вид (3.4) и носит название *формул Крамера*.

Исследование более общих систем линейных уравнений, чем системы вида (3.1), приводит нас к введению важнейшего понятия *ранга матрицы*.

С этой целью в матрице $A = [a_j^i]$ $(i = \overline{1, m};\ j = \overline{1, n})$ зафиксируем p произвольно выбранных строк и p произвольно выбранных столбцов матрицы. Из элементов, стоящих на пересечении выбранных строк и столбцов, построим квадратную матрицу p-го порядка A_p.

Определение. Детерминант матрицы A_p называется *минором p-го порядка* матрицы A.

Например, для матрицы $\begin{bmatrix} 1 & -1 & -1 \\ 2 & 0 & -1 \end{bmatrix}$ имеем следующие миноры второго порядка:

$$\begin{vmatrix} 1 & -1 \\ 2 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix}, \quad \begin{vmatrix} -1 & -1 \\ 0 & -1 \end{vmatrix},$$

а миноры первого порядка совпадают с элементами матрицы.

Определение. Число r называется *рангом* матрицы A (пишем ранг $A = r$), если у нее имеется хотя бы один минор r-го порядка, отличный от нуля, а все миноры $(r+1)$-го порядка (следовательно, в силу формулы разложения (2.1) $(r+2)$-го и т. д.) равны нулю. При этом всякий отличный от нуля минор r-го порядка называется *базисным*, а соответствующие столбцы и строки матрицы A называются *базисными* столбцами и строками.

Пример. Найти ранги матриц

$$A = \begin{bmatrix} 0 & 3 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Очевидно, что ранг $A = 1$. Ранг $B = 1$, так все миноры второго порядка равны нулю. Ранг $C = 2$, ибо существует минор $\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix}$ $\left(\text{также } \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix}\right)$, который отличен от нуля, а все миноры третьего порядка содержат нулевую строку и поэтому обращаются в нуль. В матрице C каждый из указанных миноров 2-го порядка может быть выбран за базисный минор.

Теорема (о базисном миноре). *Любой столбец (строка) матрицы A есть линейная комбинация ее базисных столбцов (строк).*

Прежде всего отметим, что за счет перенумерации строк и столбцов можно добиться того, что базисный минор находится в левом верхнем углу матрицы.

Рассмотрим минор $(r+1)$-го порядка, полученный окаймлением k-й строкой и j-м столбцом ($k = \overline{1, m}$; $j = \overline{1, n}$)

$$\Delta_{r+1} = \begin{vmatrix} a_1^1 \dots a_r^1 & a_j^1 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ a_1^r \dots a_r^r & a_j^r \\ a_1^k \dots a_r^k & a_j^k \end{vmatrix}.$$

Этот определитель всегда равен нулю, так как если $j \leqslant r$, то Δ_{r+1} содержит 2 одинаковых столбца; если $j > r$, то Δ_{r+1} есть минор $(r+1)$-го порядка и он равен нулю, поскольку ранг $A = r$. К такому же выводу приходим, когда дело имеем с окаймляющей строкой.

Разложим Δ_{r+1} по элементам последней строки. Получим

$$a_1^k C_1 + \dots + a_r^k C_r + a_j^k C_j = 0, \tag{3.6}$$

где $C_1, \dots, C_r, C_j$ означают алгебраические дополнения элементов $a_1^k, \dots, a_r^k, a_j^k$ в определителе Δ_{r+1}.

Поскольку C_j совпадает со значением базисного минора и не равно нулю, то из (3.6) выводим

$$a_j^k = \left(-\frac{C_1}{C_j}\right) a_1^k + \left(-\frac{C_2}{C_j}\right) a_2^k + \dots + \left(-\frac{C_r}{C_j}\right) a_r^k .$$

Придавая k значения от 1 до m, получим

$$\begin{bmatrix} a_j^1 \\ \vdots \\ a_j^m \end{bmatrix} = \left(-\frac{C_1}{C_j}\right) \begin{bmatrix} a_1^1 \\ \vdots \\ a_1^m \end{bmatrix} + \dots + \left(-\frac{C_r}{C_j}\right) \begin{bmatrix} a_r^1 \\ \vdots \\ a_r^m \end{bmatrix}.$$

Тем самым для столбцов теорема доказана. Если бы мы разложили Δ_{r+1} по элементам последнего столбца и затем рассмотрели получившиеся соотношения для $j = \overline{1, n}$, то получили бы, что каждая строка матрицы A есть линейная комбинация ее базисных строк.

Перейдем к выводу следствий из только что доказанной теоремы.

Следствие 1. Пусть дана квадратная матрица A и известно, что $\det A = 0$. Это означает, что ранг $A < n$, и, следовательно, по крайней мере один из столбцов (строк) является (на основании теоремы о базисном миноре) линейной комбинацией остальных столбцов (строк), что говорит о линейной зависимости столбцов определителя. Учитывая свойство 4 для определителей и только что полученный факт, можно сделать вывод: *детерминант матрицы A равен нулю тогда и только тогда, когда его столбцы (строки) линейно-зависимы.*

Следствие 2. Ранг матрицы равен максимальному числу линейно-независимых столбцов (строк) матрицы.

В самом деле, пусть ранг $A = r$. Базисные столбцы (строки) являются линейно-независимыми столбцами (строками), так как в противном случае по следствию 1 был бы равен нулю базисный минор. Число линейно-независимых столбцов (строк) не может и превысить числа r, потому что в противном случае нашелся бы столбец (строка), не являющийся линейной комбинацией базисных (см. лекцию 1 о линейной зависимости матриц-столбцов), что противоречит теореме о базисном миноре. Таким образом, *максимальное число линейно-независимых столбцов матрицы всегда совпадает с максимальным числом линейно-независимых строк матрицы и равно ее рангу.*

На основании этого следствия можно дать второе определение ранга матрицы.

Определение. Рангом матрицы называется максимальное число линейно-независимых столбцов (строк) этой матрицы.

Следствие 3. Ранг матрицы не изменится, если (1) переставить местами столбцы (строки); (2) умножить столбец (строку) на число, отличное от нуля; (3) прибавить к столбцу (строке) другой столбец (строку), умноженный на некоторый общий множитель.

Данное утверждение очевидным образом выполняется для преобразований (1) и (2), так как после них максимальное число линейно-независимых столбцов (строк) не изменится. Докажем справедливость утверждения для преобразования (3). Если основной s-й столбец, к которому прибавляются другие со множителем, не входит в число базисных, то во вновь преобразованной матрице согласно теореме о базисном миноре он будет линейной комбинацией базисных и ранг матрицы не изменится. Если же основной s-й столбец, к которому прибавляются со множителями другие столбцы, входит в число базисных, то опять-таки максимальное число линейно-независимых столбцов остается неизменным. Действительно, если бы

s-й столбец в преобразованной матрице стал бы являться линейной комбинацией других базисных столбцов, то мы сумели бы s-й столбец, взятый до преобразования, выразить как линейную комбинацию других базисных столбцов матрицы A и прибавленного столбца.

Если прибавленный столбец базисный, то тогда между базисными столбцами матрицы A существует линейная зависимость. Противоречие. Если же прибавленный столбец не входит в число базисных, то он представляет собой линейную комбинацию базисных, и, следовательно, опять-таки между базисными столбцами матрицы A существует линейная зависимость. Снова приходим к противоречию.

Преобразования (1)—(3), не меняющие ранга матрицы, позволяют дать следующий конструктивный метод нахождения ранга матрицы. Пусть в матрице $A=[a_j^i]$ $(i=\overline{1, m};\ j=1, n)$ имеется элемент $\alpha_1 \neq 0$. Переставляя строки и столбцы, поместим его в левом верхнем углу. Затем, вычитая из каждой строки матрицы первую строку, умноженную на соответствующий множитель, обратим в нуль все оставшиеся элементы первого столбца. Аналогичную операцию проводим и со столбцами, обращая в нуль все элементы первой строки. Имеем

$$A=\begin{bmatrix} a_1^1 & a_2^1 \ldots a_n^1 \\ a_1^2 & a_2^2 \ldots a_n^2 \\ \cdots & \cdots \\ a_1^m & a_2^m \ldots a_n^m \end{bmatrix} \backsim \begin{bmatrix} \alpha_1 & 0 & \ldots 0 \\ 0 & b_2^2 & \ldots b_n^2 \\ \cdots & \cdots & \cdots \\ 0 & b_2^m & \ldots b_n^m \end{bmatrix}.$$

Знак $\backsim$ означает *эквивалентность* матриц в смысле неизменности ранга у обеих матриц. Аналогично поступаем с матрицей $B=[b_j^i]$ $(i=\overline{2, m};\ j=\overline{2, n})$. В итоге мы придем к матрице вида

$$\left[\begin{array}{ccccc:c} \alpha_1 & & & & & \\ & \alpha_2 & 0 & & & \\ & & \ddots & & & 0 \\ & 0 & & & \alpha_r & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & 0 & & & 0 \end{array}\right].$$

Число не равных нулю и расположенных по главной диагонали блока элементов совпадает с рангом матрицы A.

Пример. Найти ранг матрицы

$$A=\begin{bmatrix} 2 & 3 & 4 & 1 \\ -1 & 2 & -1 & 6 \\ 1 & 5 & 3 & 5 \end{bmatrix} \sim \begin{bmatrix} -1 & 2 & -1 & 6 \\ 2 & 3 & 4 & 1 \\ 1 & 5 & 3 & 5 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 7 & 2 & 11 \\ 0 & 7 & 2 & 11 \end{bmatrix} \sim \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ранг матрицы A равен двум.

Лекция 4

ТЕОРЕМА КРОНЕКЕРА-КАПЕЛЛИ. УСЛОВИЯ НЕТРИВИАЛЬНОЙ СОВМЕСТНОСТИ ОДНОРОДНОЙ СИСТЕМЫ. ОБЩЕЕ РЕШЕНИЕ СИСТЕМЫ. ФУНДАМЕНТАЛЬНАЯ СИСТЕМА РЕШЕНИЙ

Пусть задана система уравнений

$$\sum_{j=1}^{n} a_j^i \cdot x^j = b^i \quad (i = \overline{1, m}).$$

Если хотя бы одно из чисел b^i отлично от нуля, то система называется *неоднородной*. Если все свободные члены b^i равны нулю, то система называется *однородной*.

Перейдем к установлению условий совместности неоднородной системы уравнений. Ее можно записать в виде

$$x^1 \begin{bmatrix} a_1^1 \\ a_1^2 \\ \vdots \\ a_1^m \end{bmatrix} + x^2 \begin{bmatrix} a_2^1 \\ a_2^2 \\ \vdots \\ a_2^m \end{bmatrix} + \ldots + x^n \begin{bmatrix} a_n^1 \\ a_n^2 \\ \vdots \\ a_n^m \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{bmatrix}. \qquad (4.1)$$

Кроме матрицы системы $A = [a_j^i]$ будем рассматривать *расширенную* матрицу системы

$$B=\begin{bmatrix} a_1^1 \dots a_n^1 & b^1 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot \\ a_1^m \dots a_n^m & b^m \end{bmatrix}.$$

Теорема (Кронекера-Капелли). *Для того, чтобы неоднородная система линейных уравнений была совместна, необходимо и достаточно, чтобы ранг матрицы системы и ранг расширенной матрицы совпадали.*

Необходимость. Дано, что система уравнений совместна. Следовательно, существует набор чисел $c^1, \dots, c^n$ таких, что выполнено

$$c^1\begin{bmatrix} a_1^1 \\ \vdots \\ a_1^m \end{bmatrix} + c^2\begin{bmatrix} a_2^1 \\ \vdots \\ a_2^m \end{bmatrix} + \dots + c^n\begin{bmatrix} a_n^1 \\ \vdots \\ a_n^m \end{bmatrix} = \begin{bmatrix} b^1 \\ \vdots \\ b^m \end{bmatrix}.$$

Это означает, что последний столбец расширенной матрицы B есть линейная комбинация столбцов матрицы A. Ранги матриц совпадают.

Достаточность. Дано: $r =$ ранг $A =$ ранг B. В силу строения матрицы B это означает, что существует один и тот же базисный минор как для A, так и для B. В таком случае последний столбец матрицы B (по теореме о базисном миноре) есть линейная комбинация базисных столбцов, и мы можем его представить в виде (для простоты считаем первые r столбцов базисными)

$$\begin{bmatrix} b^1 \\ \vdots \\ b^m \end{bmatrix} = \lambda_1\begin{bmatrix} a_1^1 \\ \vdots \\ a_1^m \end{bmatrix} + \lambda_2\begin{bmatrix} a_2^1 \\ \vdots \\ a_2^m \end{bmatrix} + \dots + \lambda_r\begin{bmatrix} a_r^1 \\ \vdots \\ a_r^m \end{bmatrix} +$$

$$+ 0\cdot\begin{bmatrix} a_{r+1}^1 \\ \vdots \\ a_{r+1}^m \end{bmatrix} + \dots + 0\cdot\begin{bmatrix} a_n^1 \\ \vdots \\ a_n^m \end{bmatrix},$$

откуда следует, что набор чисел $\lambda_1, \dots, \lambda_r, 0, \dots, 0$, подставленный в систему вместо $x^1, \dots, x^n$, обратит уравнения системы в тождества. Система совместна.

Пример. Совместна ли система уравнений?

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 1, \\ x_1 - x_2 + 2x_3 = 0, \\ 2x_1 + x_2 - x_3 = 1. \end{cases}$$

Решение:

$$B=\left[\begin{array}{rrr:r}1 & 2 & -3 & 1\\ 1 & -1 & 2 & 0\\ 2 & 1 & -1 & 1\end{array}\right]\sim\left[\begin{array}{rrr:r}1 & 2 & -3 & 1\\ 0 & -3 & 5 & -1\\ 0 & -3 & 5 & -1\end{array}\right]\sim$$

$$\sim\left[\begin{array}{rrr:r}1 & 2 & -3 & 1\\ 0 & -3 & 5 & -1\\ 0 & 0 & 0 & 0\end{array}\right].$$

Ранг $A=$ ранг $B=2$. Система совместна.

Обратимся к рассмотрению однородных систем уравнений $\sum\limits_{j=1}^{n} a_j^i x^j=0$ $(i=\overline{1,m})$. Такая система всегда совместна, так как набор $0, 0, \ldots, 0$ всегда удовлетворяет системе. Нулевое решение называется *тривиальным* и нас не интересует.

Поставим вопрос об условиях *нетривиальной совместности* однородной системы уравнений. Ответ дает следующее утверждение.

Теорема. *Для того, чтобы однородная система линейных уравнений была нетривиально совместна, необходимо и достаточно, чтобы ранг матрицы системы был меньше числа неизвестных.*

Необходимость. Дано: однородная система нетривиально совместна. Следовательно, существует набор чисел $c^1, \ldots, c^n$, среди которых хотя бы одно отлично от нуля, что выполнено

$$c^1\begin{bmatrix}a_1^1\\ \vdots\\ a_1^m\end{bmatrix}+c^2\begin{bmatrix}a_2^1\\ \vdots\\ a_2^m\end{bmatrix}+\ldots+c^n\begin{bmatrix}a_n^1\\ \vdots\\ a_n^m\end{bmatrix}=\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}.$$

Последнее соотношение определяет линейную зависимость столбцов матрицы системы. Поэтому ранг $A<n$.

Достаточность. Дано, что ранг $A<n$. В таком случае (см. лекцию 3, следствие 1) столбцы матрицы линейно-зависимы. Это означает, что существуют такие числа $\lambda_1, \ldots, \lambda_n$, среди которых по крайней мере одно отлично от нуля, что

$$\lambda_1 \begin{bmatrix} a_1^1 \\ \vdots \\ a_1^m \end{bmatrix} + \lambda_2 \begin{bmatrix} a_2^1 \\ \vdots \\ a_2^m \end{bmatrix} + \ldots + \lambda_n \begin{bmatrix} a_n^1 \\ \vdots \\ a_n^m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Следовательно, набор чисел $\lambda_1, \ldots, \lambda_n$, подставленный вместо $x^1, \ldots, x^n$ в систему, обратит уравнения в тождества. Система нетривиально совместна.

Как описать все множество решений системы линейных уравнений? Чтобы ответить на этот вопрос, дадим определение общего решения для системы линейных уравнений.

Определение. Общим решением называется набор функций $\Phi_1(a; b; c^1, \ldots, c^q), \ldots, \Phi_n(a; b; c^1, \ldots, c^q)$, зависящих от коэффициентов системы, свободных членов и q числовых параметров, удовлетворяющий двум условиям: (1) подстановка $\Phi_1, \ldots, \Phi_n$ в систему вместо $x^1, \ldots, x^n$ обращает уравнения в тождества; (2) любое решение системы может быть получено из этого набора путем фиксирования определенных значений параметров $c^1, \ldots, c^q$.

Как конструктивно найти общее решение системы? Предлагается следующий метод.

Пусть r — ранг матрицы системы. За счет перенумерации переменных и уравнений можно добиться того, что базисный минор матрицы эквивалентной системы уравнений будет находиться в левом верхнем углу матрицы. Все дальнейшие рассуждения проводятся для этого случая.

Рассмотрим первые r уравнений сисемы (4.1) и перепишем их в следующем виде:

$$\begin{cases} a_1^1 x^1 + \ldots + a_r^1 x^r = b^1 - a_{r+1}^1 c^{r+1} - \ldots - a_n^1 c^n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_1^r x^1 + \ldots + a_r^r x^r = b^r - a_{r+1}^r c^{r+1} - \ldots - a_n^r c^n. \end{cases} \tag{4.2}$$

Здесь переменные $x^{r+1}, \ldots, x^n$ являются параметрами и обозначены через $c^{r+1}, \ldots, c^n$ ($c^q \in \mathbf{K}$, $q = \overline{r+1, n}$).

Определитель Δ данной системы уравнений совпадает с базисным минором и, применяя формулы Крамера, получим

$$x^p = \frac{\Delta_p}{\Delta} \quad (p = \overline{1, r}), \tag{4.3}$$

где

$$\Delta_p = \begin{vmatrix} \overbrace{a_1^1 \dots b^1 - \sum_{q=r+1}^{n} a_q^1 c^q}^{p} \dots a_r^1 \\ a_1^2 \dots b^2 - \sum_{q=r+1}^{n} a_q^2 c^q \dots a_r^2 \\ \dots\dots\dots\dots\dots \\ a_1^r \dots b^r - \sum_{q=r+1}^{n} a_q^r c^q \dots a_r^r \end{vmatrix}. \tag{4.4}$$

Из приведенных рассуждений следует, что набор чисел $\frac{\Delta_1}{\Delta}, \dots, \frac{\Delta_r}{\Delta}, c^{r+1}, \dots, c^n$, подставленных вместо неизвестных, обращает в тождество первые r уравнений системы линейных уравнений (4.1). Покажем, что этот набор обращает в тождества и оставшиеся $(m-r)$ уравнений рассматриваемой неоднородной системы. С этой целью вспомним, что $(r+1)$-я, ..., m-я строка расширенной матрицы есть линейная комбинация первых r базисных строк. Для системы линейных уравнений это означает, что каждое уравнение, начиная с $(r+1)$-го и кончая m-ым, есть линейная комбинация первых r уравнений. Вследствие тождественного удовлетворения рассматриваемым набором первых r уравнений тождественно удовлетворяются и оставшиеся $(m-r)$ уравнений системы.

Итак, матрица-столбец

$$X = \begin{bmatrix} \frac{\Delta_1}{\Delta} \\ \vdots \\ \frac{\Delta_r}{\Delta} \\ c^{r+1} \\ \vdots \\ c^n \end{bmatrix} \tag{4.5}$$

есть решение исследуемой системы уравнений.

Докажем, что X задает общее решение системы. Для этого нам необходимо показать, что **любое** решение системы содержится в X при определенных значениях параметров $c^{r+1}, \dots, c^n$.

Пусть набор $\bar{c}^1, \bar{c}^2, \dots, \bar{c}^n$ — некоторое решение системы. Оно является также решением системы (4.2). Зафиксируем в правой части (4.2) значения параметров c^q, положив $c^{r+1} = \bar{c}^{r+1}, \dots, c^n = \bar{c}^n$. Как известно, по формулам Крамера решение определяется *однозначным образом*. Поэтому с необходимостью числа $\bar{c}^p$ $(p = \overline{1, r})$ будут совпадать с числами $\frac{\bar{\Delta}_p}{\Delta}$, где за $\bar{\Delta}_p$ обозначены определители (4.4), у которых вместо параметров c^q выступают фиксированные значения $\bar{c}^q$ $(q = \overline{r+1, n})$.

Пример. Решить систему уравнений

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 1 \\ x_1 - x_2 + 2x_3 = 0 \\ 2x_1 + x_2 - x_3 = 1. \end{cases}$$

Разбирая предыдущий пример, мы показали, что система совместна. Ранг основной и расширенной матриц равен двум. Минор 2-го порядка в левом верхнем углу

$$\Delta = \begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = -3$$

является базисным.

Поступаем согласно разработанному алгоритму. Имеем

$$\begin{cases} x_1 + 2x_2 = 1 + 3c_3 \\ x_1 - x_2 = -2c_3, \end{cases}$$

$$x_1 = \frac{\begin{vmatrix} 1+3c_3 & 2 \\ -2c_3 & -1 \end{vmatrix}}{-3} = \frac{1}{3}(1 - c_3),$$

$$x_2 = \frac{\begin{vmatrix} 1 & 1+3c_3 \\ 1 & -2c_3 \end{vmatrix}}{-3} = \frac{1}{3}(5c_3 + 1).$$

Получили общее решение системы.

При нахождении общего решения однородной системы линейных уравнений важную роль играют понятия *нормальной фундаментальной системы решений и фундаментальной системы решений.*

Пусть задана однородная система $\sum_{j=1}^{n} a_j^i x^j = 0$ $(i = \overline{1, m})$. Пусть ранг матрицы $A = [a_j^i]$ равен r. Согласно предложенному выше методу решение системы находится по формулам (4.5), где в определителях Δ_p необходимо положить $b^p = 0$ $(p = \overline{1, r})$.

Зададимся $(n-r)$ наборами значений параметров $c^{r+1}, \ldots, c^n$, фиксируя их следующим образом:

$$\begin{bmatrix} c_1^{r+1} \\ c_1^{r+2} \\ \vdots \\ c_1^n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \begin{bmatrix} c_2^{r+1} \\ c_2^{r+2} \\ \vdots \\ c_2^n \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} c_{n-r}^{r+1} \\ c_{n-r}^{r+2} \\ \vdots \\ c_{n-r}^n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \quad (4.6)$$

Этим наборам, согласно (4.5), будет соответствовать $(n-r)$ решений

$$X_1 = \begin{bmatrix} x_1^1 \\ \vdots \\ x_1^r \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_2^1 \\ \vdots \\ x_2^r \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, X_{n-r} = \begin{bmatrix} x_{n-r}^1 \\ \vdots \\ x_{n-r}^r \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad (4.7)$$

где $x_1^p, \ldots, x_{n-r}^p$ подсчитываются по формулам (4.3), с учетом фиксированных значений c^q $(q = \overline{r+1, n})$. Они имеют вид

$$x_1^p = \frac{1}{\Delta} \begin{vmatrix} a_1^1 \ldots \underbrace{(-a_{r+1}^1)}_{} \ldots a_r^1 \\ a_1^2 \ldots (-a_{r+1}^2) \ldots a_r^2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_1^r \ldots \underbrace{(-a_{r+1}^r)}_{p} \ldots a_r^r \end{vmatrix}, \ldots,$$

$$x_{n-r}^{p} = \frac{1}{\Delta}\begin{vmatrix} a_1^1 \dots (-a_n^1) \dots a_r^1 \\ a_1^2 \dots (-a_n^2) \dots a_r^2 \\ \dots\dots\dots\dots\dots \\ a_1^r \dots \underbrace{(-a_n^r)}_{p} \dots a_r^r \end{vmatrix}. \qquad (4.8)$$

Система полученных решений образует линейно-независимую систему матриц-столбцов, так как в матрице, содержащей n строк и $n-r$ столбцов

$$\begin{bmatrix} x_1^1 & x_2^1 \dots & x_{n-r}^1 \\ \vdots & \vdots & \vdots \\ x_1^r & x_2^r \dots & x_{n-r}^r \\ 1 & 0 \dots & 0 \\ 0 & 1 \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \dots & 1 \end{bmatrix},$$

имеется минор $(n-r)$-го порядка

$$\begin{vmatrix} 1 & 0 \dots & 0 \\ 0 & 1 \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \dots & 1 \end{vmatrix},$$

отличный от нуля.

Определение. Система решений (4.7) однородной системы линейных уравнений, соответствующая $(n-r)$ наборам (4.6) параметров c^q $(q = \overline{r+1, n})$, называется *нормальной фундаментальной системой решений.*

Теорема. *Общее решение однородной системы уравнений есть линейная комбинация с произвольными постоянными нормальной фундаментальной системы решений.*

В самом деле, рассматривая общее решение (4.5), где в определителях Δ_p постоянные b^i $(i = \overline{1, m})$ равны нулю, используя свойство 3 определителей, получим

$$
\begin{cases}
x^p = \dfrac{1}{\Delta}\begin{vmatrix} a_1^1 \dots \underbrace{(-a_{r+1}^1 c^{r+1})}_{} \dots a_r^1 \\ \vdots \qquad \vdots \qquad \vdots \\ a_1^r \dots \underbrace{(-a_{r+1}^r c^{r+1})}_{p} \dots a_r^r \end{vmatrix} + \dots + \\
\qquad\qquad + \dfrac{1}{\Delta}\begin{vmatrix} a_1^1 \dots (-a_n^1 c^n) \dots a_r^1 \\ \vdots \qquad \vdots \qquad \vdots \\ a_1^r \dots \underbrace{(-a_n^r c^n)}_{p} \dots a_r^r \end{vmatrix} \\
x^q = c^q \quad (q = \overline{r+1, n};\ p = \overline{1, r})
\end{cases}
\tag{4.9}
$$

или

$$
\begin{cases}
x^p = c^{r+1}\left(\dfrac{\Delta_p^1}{\Delta}\right) + c^{r+2}\left(\dfrac{\Delta_p^2}{\Delta}\right) + \dots + c^n\left(\dfrac{\Delta_p^{n-r}}{\Delta}\right) \\
x^q = c^q \quad (q = \overline{r+1, n};\ p = \overline{1, r}),
\end{cases}
\tag{4.10}
$$

где

$$
\Delta_p^1 = \begin{vmatrix} a_1^1 \dots (-a_{r+1}^1) \dots a_r^1 \\ \vdots \qquad \vdots \qquad \vdots \\ a_1^r \dots \underbrace{(-a_{r+1}^r)}_{p} \dots a_r^r \end{vmatrix}, \dots,
$$

$$
\Delta_p^{n-r} = \begin{vmatrix} a_1^1 \dots (-a_n^1) \dots a_r^1 \\ \vdots \qquad \vdots \qquad \vdots \\ a_1^r \dots (-a_n^r) \dots a_r^r \end{vmatrix}.
$$

С учетом (4.7), (4.8) общее решение (4.10) может быть записано в виде

$$
X = c^{r+1} X_1 + c^{r+2} X_2 + \dots + c^n X_{n-r}, \tag{4.11}
$$

что и доказывает теорему.

Зададимся сейчас $(n-r)$ наборами параметров $c^{r+1}, \dots, c^n$, фиксируя их произвольным образом

$$
\begin{bmatrix} \overset{*}{c}{}_1^{r+1} \\ \vdots \\ \overset{*}{c}{}_1^n \end{bmatrix}, \begin{bmatrix} \overset{*}{c}{}_2^{r+1} \\ \vdots \\ \overset{*}{c}{}_2^n \end{bmatrix}, \dots, \begin{bmatrix} \overset{*}{c}{}_{n-r}^{r+1} \\ \vdots \\ \overset{*}{c}{}_{n-r}^n \end{bmatrix} \tag{4.12}
$$

с одним лишь условием, чтобы детерминант матрицы

$$N = \begin{bmatrix} \overset{*}{c}{}_1^{r+1} & \overset{*}{c}{}_2^{r+1} & \dots & \overset{*}{c}{}_{n-r}^{r+1} \\ \vdots & \vdots & & \vdots \\ \overset{*}{c}{}_1^{n} & \overset{*}{c}{}_2^{n} & \dots & \overset{*}{c}{}_{n-r}^{n} \end{bmatrix}$$

был отличен от нуля.

Наборам (4.12) отвечает $n-r$ линейно-независимых (из-за $\det N \neq 0$) решений

$$\overset{*}{X}_1 = \begin{bmatrix} \overset{*}{x}{}_1^1 \\ \vdots \\ \overset{*}{x}{}_1^r \\ \overset{*}{c}{}_1^{r+1} \\ \vdots \\ \overset{*}{c}{}_1^n \end{bmatrix}, \quad \overset{*}{X}_2 = \begin{bmatrix} \overset{*}{x}{}_2^1 \\ \vdots \\ \overset{*}{x}{}_2^r \\ \overset{*}{c}{}_2^{r+1} \\ \vdots \\ \overset{*}{c}{}_2^n \end{bmatrix}, \dots, \overset{*}{X}_{n-r} = \begin{bmatrix} \overset{*}{x}{}_{n-r}^1 \\ \vdots \\ \overset{*}{x}{}_{n-r}^r \\ \overset{*}{c}{}_{n-r}^{r+1} \\ \vdots \\ \overset{*}{c}{}_{n-r}^n \end{bmatrix}. \qquad (4.13)$$

Система решений (4.13), отвечающая наборам (4.12) параметров $c^{r+1}, \dots, c^n$ при условии $\det N \neq 0$, называется фундаментальной системой решений.

Как и выше, можно доказать, что *общее решение однородной системы линейных уравнений есть линейная комбинация с произвольными постоянными фундаментальной системы решений.*

Доказательство проводить не будем, заметив только, что каждое из $\overset{*}{X}_s$ $(s = \overline{1, n-r})$ по доказанной выше теореме есть конкретная линейная комбинация нормальной фундаментальной системы решений.

Из-за линейной независимости систем $\{X_s\}$ и $\{\overset{*}{X}_s\}$ каждое из X_s также представляет собой фиксированную линейную комбинацию матриц-столбцов $\overset{*}{X}_s$ (ее конкретный вид может быть получен по формулам Крамера). Подстановка этих комбинаций в (4.11) приведет нас к желаемому результату.

В заключение вернемся к неоднородным системам линейных уравнений $\sum\limits_{j=1}^{n} a_j^i x^j = b^i$ $(i = \overline{1, m})$. Система $\sum\limits_{j=1}^{n} a_j^i x^j = 0$ $(i = \overline{1, m})$ с той же матрицей $A = [a_j^i]$

называется однородной системой, *соответствующей* рассматриваемой неоднородной системе.

Теорема. *Общее решение неоднородной системы есть сумма частного решения неоднородной системы и общего решения соответствующей однородной системы.*

В самом деле, пусть

$$X_0 = \begin{bmatrix} x_0^1 \\ x_0^2 \\ \vdots \\ x_0^n \end{bmatrix}$$

— частное решение неоднородной системы

$$\left(\sum_{j=1}^{n} a_j^i x_0^j \equiv b^i \quad (i = \overline{1, m})\right).$$

Вместо неизвестных x^j в неоднородной системе введем новые неизвестные y^j, положив $x^j = x_0^j + y^j$. Имеем

$$\sum_{j=1}^{n} a_j^i (x_0^j + y^j) = \sum_{j=1}^{n} a_j^i x_0^j + \sum_{j=1}^{n} a_j^i y^j = b^i \quad (i = \overline{1, m}).$$

Поскольку $\sum_{j=1}^{n} a_j^i x_0^j \equiv b^i$, то получаем, что

$$\sum_{j=1}^{n} a_j^i y^j = 0 \quad (i = \overline{1, m}).$$

Неизвестные y^j являются неизвестными для соответствующей однородной системы. Пусть Y — общее решение этой системы уравнений, найденное по формуле (4.11). Тогда для неоднородной системы общее решение имеет вид

$$X = X_0 + \sum_{s=1}^{n-r} c^{r+s} X_s, \tag{4.14}$$

где $\{X_s\}$ образуют нормальную фундаментальную систему решений (фундаментальную систему решений).

На этом мы заканчиваем рассмотрение теории систем линейных алгебраических уравнений и переходим к вопросам векторной алгебры в трехмерном евклидовом пространстве.

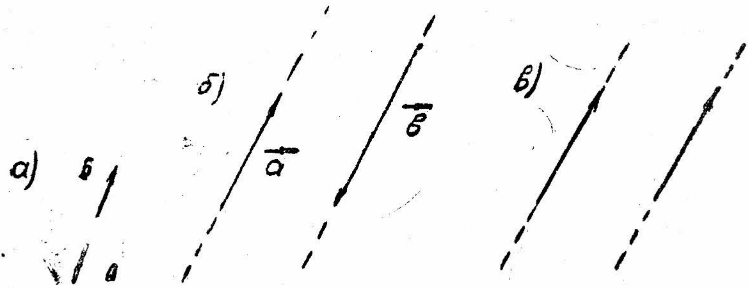


Рис. 1. а) — вектор $\overrightarrow{AB}$; б) — коллинеарные векторы; в) — равные векторы.

Определение 2. Всякую упорядоченную пару точек называют *направленным отрезком* или геометрическим *вектором* и обозначают символом $\overrightarrow{AB}$.

На вектор можно смотреть как на оператор сдвига точки A в точку B. Его обозначают обычно малой латинской буквой со стрелкой (чертой) сверху: $\vec{a}, \vec{b}, \ldots, \vec{m}, \ldots$ Поэтому вектор $\overrightarrow{AB}$ можно записать в виде $\vec{a}(A) = B$. Но такую громоздкую запись обычно не используют, оставляя лишь первую букву со стрелкой (чертой) наверху.

Определение 3. Расстояние между точками A и B называют *длиной* вектора $\overrightarrow{AB}$ и для обозначения длины пользуются символом модуля (абсолютной величины): $|\overrightarrow{AB}|$, $|\vec{a}|, \ldots$

Определение 4. Векторы называются *коллинеарными*, если они лежат на одной прямой либо на параллельных прямых.

Определение 5. Вектор называется *нулевым* (пишется $\vec{0}$ либо просто 0), если его начало и конец совпадают. Очевидно, что длина нулевого вектора равна нулю, а направление не определено (можно счи-

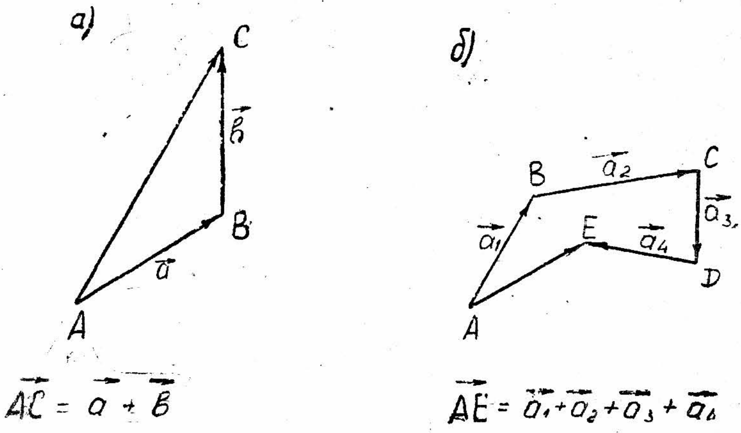


Рис. 2.

тать, что его направление какое угодно). Вследствие этого *нулевой вектор коллинеарен с любым вектором пространства.*

Определение 6. Два вектора считаются *равными*, если их длины равны и они имеют одинаковые направления (сонаправлены). Пишем: $\vec{a} = \vec{b}$ (рис. 1).

Из последнего определения следует, что каковы бы ни были точка P и вектор $\vec{a}$, всегда найдется такая точка Q, что $\overrightarrow{PQ} = \vec{a}$. Это означает, что точка приложения вектора $\vec{a}$ может быть выбрана *произвольно* (оператор сдвига $\vec{a}$ одновременно действует на все точки пространства). Такие векторы называются *свободными*. В дальнейшем мы будем рассматривать только свободные векторы, хотя в статике твердого тела используют *скользящие* векторы. Для них определение равных векторов состоит в том, что два вектора считаются равными, если их длины равны, они сонаправлены и лежат на одной прямой.

Для свободных векторов введем две линейные операции (пришедшие из физики) и укажем их свойства.

Определение. *Суммой* векторов $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ называется вектор $\vec{a}$ (пишем $\vec{a} = \vec{a}_1 + \vec{a}_2 + \ldots + \vec{a}_k$), Отметим, что в этом курсе мы не останавливаемся на численных методах (с использованием ЭВМ) решения систем уравнений, так как эти вопросы для студентов физических факультетов вынесены в отдельный учебный курс.

Лекция 5

ВЕКТОРЫ. СЛОЖЕНИЕ ВЕКТОРОВ. УМНОЖЕНИЕ ВЕКТОРА НА ЧИСЛО. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ ВЕКТОРОВ В ТРЕХМЕРНОМ ПРОСТРАНСТВЕ

С этой лекции мы переходим к изложению вопросов аналитической геометрии на плоскости и в пространстве, заметив, что ареной действия является трехмерное евклидово пространство. Строгое аксиоматическое определение n-мерного евклидова пространства будет дано позднее (в курсе линейной алгебры), а сейчас мы будем опираться лишь на те геометрические понятия евклидова пространства, что рассмотрены в школьном курсе элементарной геометрии.

Хотя содержание аналитической геометрии концентрируется вокруг понятия координат, целесообразно начать ее изложение с понятия вектора и аксиоматически ввести операции с векторами, а затем уже переходить к координатной записи. Отметим, что физика и механика дают нам многочисленные примеры использования векторных величин, как, например, скорость, ускорение тела, сила и т. д. В геометрии абстрагируются от конкретного физического содержания векторных величин и говорят просто о векторе. Отметим, что в лекциях 5—14 в качестве числового поля выступает поле вещественных чисел.

Определение 1. Если для двух точек A, B указано, какая из них является начальной и какая конечной, то говорят об *упорядоченной паре точек* (A, B) (A — начальная точка, B — конечная точка).

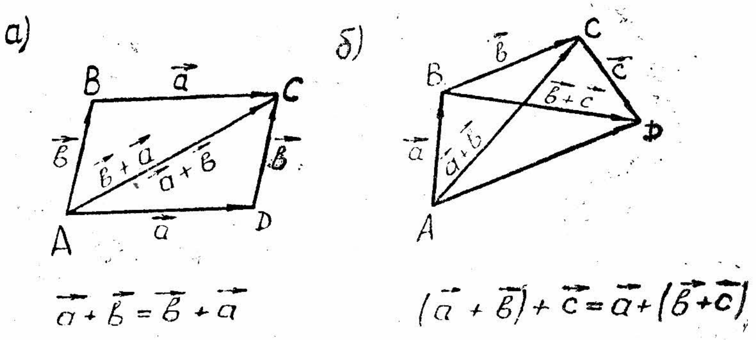


Рис. 3.

который замыкает ломаную линию, построенную из данных векторов так, что начало каждого последующего вектора совмещается с концом предыдущего. Замыкающий вектор $\vec{a}$ направлен из начала первого вектора к концу последнего (см. рис. 2).

Операция сложения векторов обладает следующими свойствами.

Свойство 1 (коммутативность). Для любых векторов $\vec{a}+\vec{b}=\vec{b}+\vec{a}$. В самом деле, для коллинеарных векторов свойство 1 очевидно. В случае не коллинеарных векторов (рис. 3а) $\overrightarrow{AD}=\vec{a}=\overrightarrow{BC}$, $\overrightarrow{DC}=\vec{b}=\overrightarrow{AB}$. Поэтому $\vec{a}+\vec{b}=\overrightarrow{AC}=\vec{b}+\vec{a}$.

Свойство 2 (ассоциативность). Для любых векторов $(\vec{a}+\vec{b})+\vec{c}=\vec{a}+(\vec{b}+\vec{c})$. Действительно (см. рис. 3б), $\overrightarrow{AD}=\overrightarrow{AC}+\overrightarrow{CD}$ и $\overrightarrow{AD}=\overrightarrow{AB}+\overrightarrow{BD}$. Поэтому $(\vec{a}+\vec{b})+\vec{c}=\vec{a}+(\vec{b}+\vec{c})$.

Свойство 3. Нуль-вектор является нулем сложения: $\vec{a}+\vec{0}=\vec{a}$.

Прежде чем переходить к формулировке четвертого свойства, дадим определение противоположного вектора.

Определение. Вектор $\vec{a}'$ (обозначается также символом $-\vec{a}$) называется *противоположным* по отношению к вектору $\vec{a}$, если длины векторов $\vec{a}$ и $\vec{a}'$ равны, а направления противоположны.

Поскольку для любых точек выполнено $\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA}$, то можно сформулировать

Свойство 4. Для любого вектора $\vec{a}$ существует противоположный вектор $\vec{a}'$, такой, что $\vec{a} + \vec{a}' = \vec{0}$.

Перейдем к следующей операции — умножению вектора на вещественное число.

Определение. *Произведением* вектора $\vec{a}$ на число λ называется вектор (символ $\lambda\vec{a}$) такой, что (1) $|\lambda\vec{a}| = |\lambda|\cdot|\vec{a}|$; (2) векторы $\vec{a}$ и $\lambda\vec{a}$ сонаправлены, если $\lambda > 0$, и противоположно направлены, если $\lambda < 0$, и $\lambda\vec{a} = \vec{0}$, если $\lambda = 0$.

Операция произведения вектора на число обладает следующими четырьмя свойствами, доказательства которых просты и предоставляются читателю.

Свойство 1. $1\cdot\vec{a} = \vec{a}$ для любого вектора $\vec{a} \neq \vec{0}$.

Свойство 2. $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$ для любых $\lambda, \mu \in \mathbf{R}$ и любого вектора $\vec{a}$.

Свойство 3. $(\lambda\mu)\vec{a} = \lambda(\mu\vec{a})$ для любых $\lambda, \mu \in \mathbf{R}$ и любого $\vec{a}$.

Свойство 4. $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$ для любого $\lambda \in \mathbf{R}$ и любых векторов $\vec{a}$ и $\vec{b}$.

Из определения операции умножения вектора на число следует, что противоположный вектор $\vec{a}'$ может быть получен путем умножения вектора $\vec{a}$ на -1: $\vec{a}' = (-1)\cdot\vec{a}$. Отсюда и обозначение $-\vec{a}$, принятое для противоположного вектора.

О п р е д е л е н и е. *Разностью* векторов $\vec{a}$ и $\vec{b}$ называем сумму вектора $\vec{a}$ и противоположного вектора $\vec{b}'$. Обозначаем разность символом $\vec{a}-\vec{b}$. Итак, $\vec{a}-\vec{b} \overset{\text{def}}{=} \vec{a}+(-1)\cdot\vec{b}$ (см. рис. 4).

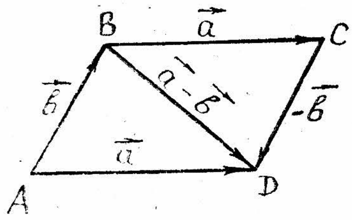


Рис. 4.

Пользуясь операцией умножения вектора на число, выведем формулу для нахождения орта вектора.

О п р е д е л е н и е. *Ортом* вектора $\vec{a}$ называется вектор единичной длины, имеющий то же направление, что и вектор $\vec{a}$. Орт вектора обозначают символом $\vec{a}^0$ (либо $\vec{e}$).

По определению $\vec{a}^0=\lambda\vec{a}$, где $\lambda>0$. Но $|\vec{a}^0|=1$. Поэтому $|\lambda\vec{a}|=\lambda\cdot|\vec{a}|=1$. Следовательно, множитель $\lambda=\dfrac{1}{|\vec{a}|}$. Окончательно можно записать, что

$$\vec{a}^0=\frac{1}{|\vec{a}|}\vec{a}=\frac{\vec{a}}{|\vec{a}|}. \tag{5.1}$$

Обе линейные операции позволяют нам рассматривать *линейные комбинации* векторов $\vec{a}_1, \dots, \vec{a}_s$ с коэффициентами $\lambda_1, \dots, \lambda_s$

$$\vec{l}=\lambda_1\vec{a}_1+\lambda_2\vec{a}_2+\dots+\lambda_s\vec{a}_s. \tag{5.2}$$

О п р е д е л е н и е. Система векторов $\vec{a}_1, \dots, \vec{a}_s$ называется *линейно-зависимой*, если существуют такие числа $\lambda_1, \dots, \lambda_s$, среди которых хотя бы одно не равно нулю, что выполнено

$$\lambda_1\vec{a}_1+\lambda_2\vec{a}_2+\dots+\lambda_s\vec{a}_s=\vec{0}. \tag{5.3}$$

Если же (5.3) для системы $\vec{a}_1, \ldots, \vec{a}_s$ выполнено тогда и только тогда, когда все $\lambda_i = 0$, то система векторов называется *линейно-независимой*.

Приводимые ниже теоремы полностью разъясняют геометрический смысл линейной зависимости векторов.

Теорема 1. *Для того, чтобы система, состоящая из одного вектора, была линейно-зависимой, необходимо и достаточно, чтобы он был нулевым вектором.*

В самом деле, пусть система $\{\vec{a}\}$, состоящая из одного вектора, линейно-зависима. Следовательно, $\lambda\vec{a} = \vec{0}$, где $\lambda \neq 0$. Вектор $\vec{a}$ — нулевой. Пусть, обратно, вектор $\vec{a}$ есть нуль-вектор. Очевидно, что при любом $\lambda \neq 0$ имеем $\lambda\vec{a} = \vec{0}$.

Вывод из этой теоремы можно сформулировать и так: *каждый не нулевой вектор есть линейно-независимый вектор.*

Теорема 2. *Для того, чтобы система, состоящая из двух векторов, была линейно-зависимой, необходимо и достаточно, чтобы векторы были коллинеарны.*

Действительно, если система $\{\vec{a}, \vec{b}\}$ из двух векторов линейно-зависима, то $\lambda\vec{a} + \mu\vec{b} = \vec{0}$, где $\lambda \neq 0$ или $\mu \neq 0$. Если $\lambda \neq 0$, то $\vec{a} = \left(-\frac{\mu}{\lambda}\right)\vec{b}$ и векторы коллинеарны. Если $\mu \neq 0$, то $\vec{b} = \left(-\frac{\lambda}{\mu}\right)\vec{a}$. Имеем коллинеарные векторы.

Пусть, обратно, векторы $\vec{a}$ и $\vec{b}$ коллинеарны между собой. Если хотя бы один из них нулевой (например, $\vec{a} = \vec{0}$), то очевидным образом выполняется $\lambda\vec{a} + 0\cdot\vec{b} = \vec{0}$, где $\lambda \neq 0$. Система линейно-зависима. Если же ни один из векторов не нулевой, то согласно определениям операции умножения вектора на число и орта вектора $\vec{a}$ получим $\vec{b} = \varepsilon|\vec{b}|\frac{\vec{a}}{|\vec{a}|}$, где $\varepsilon = +1$, когда направления векторов $\vec{a}$ и $\vec{b}$ совпадают, $\varepsilon = -1$, когда направления векторов $\vec{a}$ и $\vec{b}$ противоположны. Обозначив $\varepsilon \cdot \frac{|\vec{b}|}{|\vec{a}|} = \lambda$, получим $\lambda\vec{a} - \vec{b} = \vec{0}$, где $\lambda \neq 0$. Линейная зависимость имеет место.

Вывод из доказанной теоремы может быть сформулирован следующим образом. *Любая пара не коллинеарных векторов образует систему линейно-независимых векторов.*

Определение. Векторы, расположенные в одной плоскости или параллельные одной и той же плоскости, называются *компланарными*.

Теорема 3. *Для того, чтобы система, состоящая из трех векторов, была линейно-зависимой, необходимо и достаточно, чтобы тройка векторов была компланарной.*

Необходимость. Дано, что $\{\vec{a}, \vec{b}, \vec{c}\}$ — линейно-зависимая тройка. Это означает, что выполнено $\lambda\vec{a} + \mu\vec{b} + \nu\vec{c} = \vec{0}$, где хотя бы один из коэффициентов отличен от нуля. Например, $\lambda \neq 0$. В таком случае $\vec{a} = \left(-\frac{\mu}{\lambda}\right)\vec{b} + \left(-\frac{\nu}{\lambda}\right)\vec{c}$ — вектор $\vec{a}$ есть диагональ параллелограмма, построенного на векторах $\left(-\frac{\mu}{\lambda}\right)\vec{b}$ и $\left(-\frac{\nu}{\lambda}\right)\vec{c}$. Это означает (в силу коллинеарности векторов $\vec{b}$ и $\left(-\frac{\mu}{\lambda}\right)\vec{b}$, $\vec{c}$ и $\left(-\frac{\nu}{\lambda}\right)\vec{c}$), что тройка векторов $\{\vec{a}, \vec{b}, \vec{c}\}$ лежит в плоскости указанного параллелограмма и, следовательно, компланарна. В случае, когда хотя бы один из тройки векторов нулевой, доказательство становится очевидным.

Достаточность. Пусть тройка $\{\vec{a}, \vec{b}, \vec{c}\}$ компланарна. Если хотя бы один из векторов нулевой (например, $\vec{a} = \vec{0}$), то имеет место равенство $\lambda \cdot \vec{a} + 0 \cdot \vec{b} + 0 \cdot \vec{c} = \vec{0}$, где $\lambda \neq 0$, и в данном случае тройка является линейно-зависимой. Пусть ни один из век-

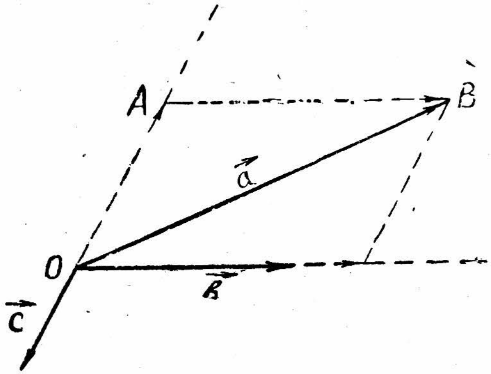


Рис. 5.

торов не является нулевым. Приведем их к общему началу 0 (см. рис. 5) и спроектируем конец вектора $\vec{a}$ (проектирующие прямые строим параллельно векторам $\vec{b}$ и $\vec{c}$) на прямые, на которых лежат векторы $\vec{b}$ и $\vec{c}$: $\vec{a}=\overrightarrow{OB}=\overrightarrow{OA}+\overrightarrow{AB}$. Но $\overrightarrow{OA}=\lambda\vec{c}$ ($\overrightarrow{OA}$ коллинеарен $\vec{c}$), $\overrightarrow{AB}=\mu\vec{b}$ ($\overrightarrow{AB}$ коллинеарен $\vec{b}$). Таким образом, имеет место равенство $\vec{a}-\lambda\vec{c}-\mu\vec{b}=\vec{0}$. Тройка векторов линейно-зависима. Данные рассуждения справедливы, когда $\vec{b}$ и $\vec{c}$ — не коллинеарные векторы (что и отражено на рис. 5). Если же $\vec{b}$ и $\vec{c}$ коллинеарны, то в силу теоремы 2 $\vec{b}-\lambda\vec{c}+0\vec{a}=\vec{0}$ — имеем линейно-зависимую систему векторов.

Вывод из теоремы 3 таков: *всякая не компланарная тройка векторов есть линейно-независимая система векторов.*

Теорема 4. *Всякая четверка векторов (в трехмерном пространстве) линейно-зависима.*

Если все четыре вектора компланарны, то в силу теоремы 3 они линейно-зависимы. Пусть векторы $\vec{a}, \vec{b}, \vec{c}$ не компланарны, а $\vec{x}$ — произвольный четвертый вектор. Приведем четверку к общему началу (рис. 6). Из конца вектора $\overrightarrow{OB}=\vec{x}$ проведем прямую,

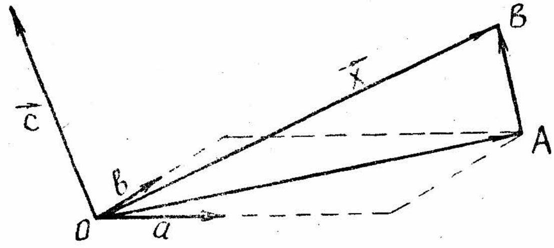


Рис. 6.

параллельную вектору $\vec{c}$, до встречи в точке A с плоскостью векторов $\vec{a}$ и $\vec{b}$. Из точки A проведем прямые, параллельные $\vec{a}$ и $\vec{b}$, до пересечения с прямыми, на которых лежат $\vec{a}$ и $\vec{b}$. Имеем, что $\vec{x}=\overrightarrow{OB}=\overrightarrow{OA}+\overrightarrow{AB}$. Но $\overrightarrow{OA}=\lambda\vec{a}+\mu\vec{b}$, $\overrightarrow{AB}=\nu\vec{c}$ и, следовательно,

$$\vec{x}=\lambda\vec{a}+\mu\vec{b}+\nu\vec{c}, \tag{5.4}$$

что и доказывает линейную зависимость четверки векторов.

Лекция 6

БАЗИС И АФФИННЫЕ КООРДИНАТЫ. ПРОЕКЦИЯ ВЕКТОРА НА ОСЬ. ПРЯМОУГОЛЬНАЯ СИСТЕМА КООРДИНАТ. ПОЛЯРНАЯ СИСТЕМА КООРДИНАТ НА ПЛОСКОСТИ

Используем результаты теорем 3 и 4, доказанных в конце предыдущей лекции, для введения базиса и аффинных координат в пространстве и на плоскости.

Определение 1. Тройка (пара) $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ ($\{\vec{e}_1, \vec{e}_2\}$) линейно-независимых векторов называется базисом в пространстве (на плоскости), если *любой* вектор $\vec{x}$ может быть представлен в виде линейной комбинации: $\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2+x_3\vec{e}_3$ ($\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2$). Числа (x_1, x_2, x_3)

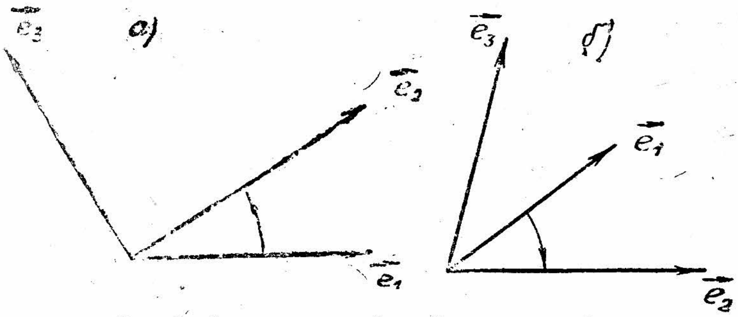


Рис. 7. а) — правая тройка; б) — левая тройка.

(на плоскости (x_1, x_2)) называются *координатами* $\vec{x}$ в указанном базисе.

Из этого определения на основании теоремы 3 вытекает, что *всякая пара не коллинеарных векторов образует базис на плоскости*, а на основании теоремы 4 — *всякая тройка не компланарных векторов образует базис в пространстве.*

Определение 2. Упорядоченная тройка не компланарных векторов $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ называется *правой* (*левой*) тройкой, если, будучи приведенной к общему началу, кратчайший поворот вокруг $\vec{e}_3$ от $\vec{e}_1$ к $\vec{e}_2$ совершается против часовой стрелки (по часовой стрелке) (см. рис. 7). Соответственно базис в пространстве называется *правым* (*левым*).

Определение 3. Упорядоченная пара не коллинеарных векторов $\{\vec{e}_1, \vec{e}_2\}$ называется *правой* (*левой*), если, будучи приведенной к общему началу, кратчайший поворот на плоскости от $\vec{e}_1$ к $\vec{e}_2$ происходит против часовой стрелки (по часовой стрелке). Соответственно базис на плоскости называется *правым* (*левым*).

Теорема 1. *Координаты вектора в данном базисе определены однозначным образом.*

Действительно, если предположить, что это не так $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$ и $\vec{x} = y_1\vec{e}_1 + y_2\vec{e}_2 + y_3\vec{e}_3$, то при вычитании векторов получим

$$(x_1 - y_1)\vec{e}_1 + (x_2 - y_2)\vec{e}_2 + (x_3 - y_3)\vec{e}_3 = \vec{0}. \qquad (6.1)$$

Но для тройки линейно-независимых векторов равенство (6.1) имеет место лишь тогда, когда все коэффициенты равны нулю. Таким образом, $x_1 = y_1$, $x_2 = y_2$, $x_3 = y_3$. Теорема доказана.

Теорема 2. *При сложении векторов их координаты складываются, а при умножении вектора на число его координаты умножаются на это число.*

В самом деле, пусть $\vec{a} = \sum_{i=1}^{3} a_i \vec{e}_i$, $\vec{b} = \sum_{i=1}^{3} b_i \vec{e}_i$. Тогда $\vec{a} + \vec{b} = \sum_{i=1}^{3} a_i \vec{e}_i + \sum_{i=1}^{3} b_i \vec{e}_i = \sum_{i=1}^{3} (a_i + b_i) \vec{e}_i$. Согласно определению $(a_1 + b_1,\ a_2 + b_2,\ a_3 + b_3)$ есть координаты вектора $\vec{a} + \vec{b}$. Умножим вектор $\vec{a}$ на число λ. Имеем

$$\lambda \vec{a} = \lambda \sum_{i=1}^{3} a_i \vec{e}_i = \sum_{i=1}^{3} (\lambda a_i) \vec{e}_i .$$

Числа $(\lambda a_1, \lambda a_2, \lambda a_3)$ есть координаты вектора $\lambda \vec{a}$.

Обе теоремы имеют место и для векторов плоскости.

Определение 4. *Аффинной системой координат* в пространстве $\{0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ (на плоскости $\{0, \vec{e}_1, \vec{e}_2\}$) называется совокупность некоторой точки O и базиса $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ ($\{\vec{e}_1, \vec{e}_2\}$), приведенного к общему началу в этой точке.

Определение 5. *Аффинными координатами* точки M в пространстве (на плоскости) в данной аффинной системе координат $\{0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ ($\{0, \vec{e}_1, \vec{e}_2\}$) называются координаты вектора $\overrightarrow{OM}$ относительно базиса $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ ($\{\vec{e}_1, \vec{e}_2\}$) (рис. 8).

Координаты $\overrightarrow{OM}$, как показано нами, определены однозначным образом в данном базисе. Тем самым показано, что каждой точке M однозначным способом сопоставлено три числа—ее аффинные координаты.

Отметим, что в аналитической геометрии употребляются только правые системы координат, т. е. тройка

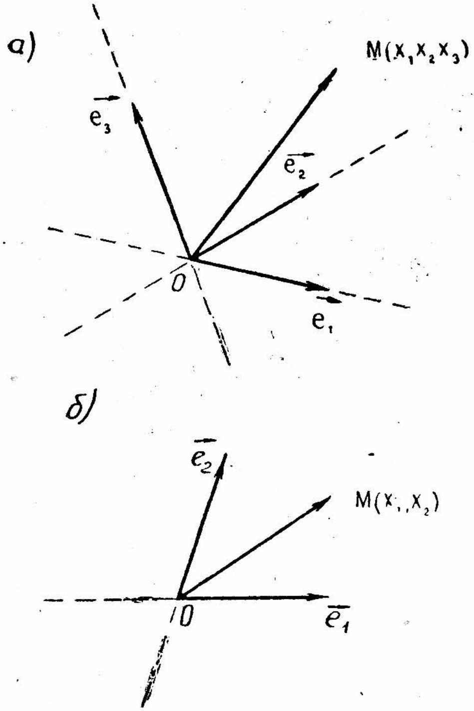


Рис. 8. а) — аффинная система координат в пространстве;
б) — аффинная система координат на плоскости.

(пара) $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ ($\{\vec{e}_1, \vec{e}_2\}$) в системе координат $\{0, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ ($\{0, \vec{e}_1, \vec{e}_2\}$) есть правая тройка векторов.

Прежде чем переходить к наиболее используемым в приложениях прямоугольным системам координат, являющимся частным случаем аффинных, остановимся на свойствах проекции вектора на ось.

Пусть задана прямая и с помощью орта $\vec{i}$ на ней указано направление. Такая прямая называется *осью*.

Определение 6. Проекцией вектора $\overrightarrow{AB}$ на ось Ox называется длина отрезка $A'B'$ этой оси, заключенного между основаниями перпендикуляров, опущенных из A и B на ось, взятая со знаком плюс, если направление $\overrightarrow{A'B'}$ совпадает с направлением $\vec{i}$, и взя-

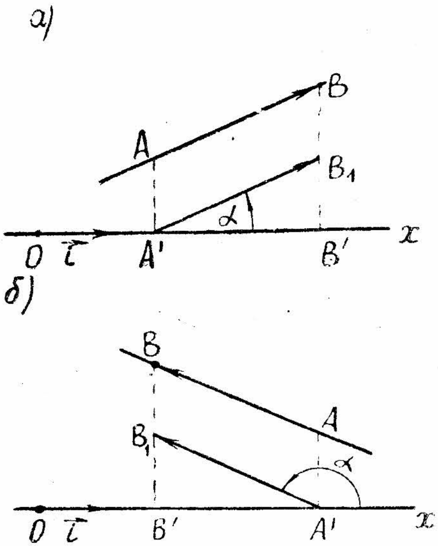


Рис. 9.

тая со знаком минус, если направления векторов $\overrightarrow{A'B'}$ и $\vec{i}$ противоположны. Символически записываем следующим образом:

$$\text{Пр}_{Ox}\,\overrightarrow{AB} = \begin{cases} |\overrightarrow{A'B'}|, & \text{если сонаправлены векторы } \vec{i} \text{ и } \overrightarrow{A'B'}, \\ -|\overrightarrow{A'B'}|, & \text{если направления } \vec{i} \text{ и } \overrightarrow{A'B'} \text{ противоположны.} \end{cases}$$

Углом вектора $\overrightarrow{AB}$ $(= \overrightarrow{A'B_1})$ с осью Ox называется угол α (рис. 9), на который нужно кратчайшим способом повернуть ось Ox около точки A', чтобы совместить с вектором $\overrightarrow{A'B_1}$ $(0 \leqslant \alpha \leqslant \pi)$.

Теорема 3. *Проекция вектора на ось равна длине вектора, умноженного на косинус угла* α.

Пусть α — острый угол. Из треугольника $A'B_1B'$ (рис. 9а) имеем:

$$\text{Пр}_{Ox}\,\overrightarrow{AB} = |\overrightarrow{A'B'}| = |\overrightarrow{A'B_1}|\cos\alpha = |\overrightarrow{AB}|\cos\alpha.$$

Пусть α — тупой угол (рис. 9б). Из треугольника $B'B_1A'$ находим:

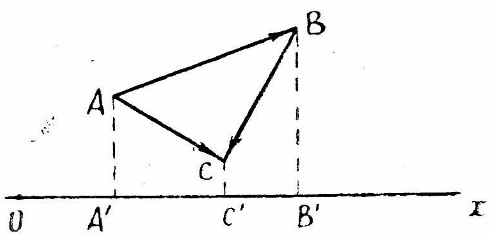


Рис. 10.

$$\text{Пр}_{Ox}\,\overrightarrow{AB} = -\,|\,\overrightarrow{A'B'}\,| = -\,|\,A'B_1\,|\cos(\pi-\alpha) =$$

$$= |\,\overrightarrow{A'B_1}\,|\cos\alpha = |\,\overrightarrow{AB}\,|\cos\alpha.$$

Если длина вектора $\overrightarrow{AB}$ равна единице (имеем орт-вектор), то $\text{Пр}_{Ox}\,\overrightarrow{AB} = \cos\alpha$. Косинус угла α называется *направляющим* косинусом единичного вектора с осью Ox.

Теорема 4. *При умножении вектора на число λ его проекция умножается на то же число.*

Пусть вектор $\vec{a}$ составляет с осью Ox угол α и $\lambda > 0$. В таком случае вектор $\lambda\vec{a}$ с осью Ox составляет тот же угол, а вектор $(-\lambda\vec{a})$ — угол $(\pi-\alpha)$. Согласно теореме 3 имеем

$$\text{Пр}_{Ox}\,(\lambda\vec{a}) = |\,\lambda\vec{a}\,|\cos\alpha = \lambda\,|\,\vec{a}\,|\cos\alpha = \lambda\,\text{Пр}_{Ox}\,\vec{a}$$

и

$$\text{Пр}_{Ox}\,(-\lambda\vec{a}) = |-\lambda\vec{a}\,|\cos(\pi-\alpha) = -\lambda\,|\,\vec{a}\,|\cos\alpha =$$

$$= -\lambda\,\text{Пр}_{Ox}\,a.$$

Теорема 5. *Проекция суммы двух векторов на ось равна сумме проекций этих векторов на ту же ось.*

Рассмотрим сумму двух векторов $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$. Для доказательства теоремы спроектируем точки A, B, C на ось Ox и получим три точки A', B', C' на оси. В общем случае возможны следующие 6 случаев расположения в направлении штрихованных точек на оси: (1) A', B', C'; (2) A', C', B'; (3) B', A', C'; (4) B', C', A'; (5) C', A', B'; (6) C', B', A'.

Доказательство проведем для случая (2) (рис. 10). Имеем

$$\begin{cases} \text{Пр}_{Ox}(\overrightarrow{AB}+\overrightarrow{BC}) = \text{Пр}_{Ox}\overrightarrow{AC} = |\overrightarrow{A'C'}|, \\ \text{Пр}_{Ox}\overrightarrow{AB} + \text{Пр}_{Ox}\overrightarrow{BC} = |\overrightarrow{A'C'}| - |\overrightarrow{B'C'}| = |\overrightarrow{A'C'}|. \end{cases}$$

Из сравнения соотношений вытекает, что

$$\text{Пр}_{Ox}(\overrightarrow{AB}+\overrightarrow{BC}) = \text{Пр}_{Ox}\overrightarrow{AB} + \text{Пр}_{Ox}\overrightarrow{BC}.$$

Аналогично доказываются остальные случаи. Теорема 5 по индукции распространяется на любое число слагаемых векторов.

Определение. Система координат $\{0, \vec{i}, \vec{j}, \vec{k}\}$ в пространстве (на плоскости $\{0, \vec{i}, \vec{j}\}$) называется *декартовой прямоугольной системой координат*, если базисные векторы $\{\vec{i}, \vec{j}, \vec{k}\}$ (пара $\{\vec{i}, \vec{j}\}$) попарно ортогональны и имеют длину, равную единице.

Базисные векторы $\vec{i}, \vec{j}, \vec{k}$ соответственно называются *ортами координатных осей* Ox, Oy, Oz (на плоскости Ox, Oy).

Определение. Декартовыми прямоугольными координатами точки M в пространстве (на плоскости) относительно выбранной прямоугольной системы координат называются координаты вектора $\overrightarrow{OM}$ в ортонормированном базисе $\{\vec{i}, \vec{j}, \vec{k}\}$ (на плоскости $\{\vec{i}, \vec{j}\}$), т. е.

$$\overrightarrow{OM} = x\vec{i} + y\vec{j} + z\vec{k} \Rightarrow M(x; y; z).$$

Из определения проекции вектора на ось следует геометрический смысл прямоугольных координат точки M

$$x = \text{Пр}_{Ox}\overrightarrow{OM}, \quad y = \text{Пр}_{Oy}\overrightarrow{OM}, \quad z = \text{Пр}_{Oz}\overrightarrow{OM}.$$

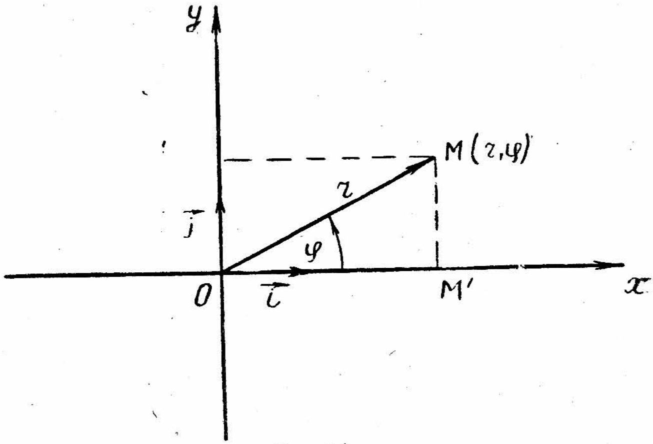


Рис. 11.

Для плоскости имеем буквальное повторение сказанного.

Некоторые задачи аналитической геометрии на плоскости целесообразно решать используя не прямоугольную систему координат на плоскости, а *полярную систему координат*, которая строится следующим образом. Выбираем точку O на плоскости (*полюс*) и ось с ортом $\vec{i}$ (*полярная ось*) (рис. 11). Любой точке плоскости можно однозначно сопоставить (за исключением полюса) пару чисел (*полярные координаты*) $r = |\overrightarrow{OM}|$ и φ — угол между $\overrightarrow{OM}$ и полярной осью. При этом считаем $\varphi > 0$, если поворот от $\vec{i}$ к $\overrightarrow{OM}$ совершается против часовой стрелки. Полюс задается одним числом $r = 0$. Когда $0 \leqslant r < \infty$ и $0 \leqslant \varphi < 2\pi$ (или $-\pi < \varphi \leqslant \pi$), мы пробегаем все множество точек плоскости. Из определения проекции вектора на ось и с учетом геометрического смысла прямоугольных координат точки следует связь полярных и прямоугольных координат

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi, \end{cases} \qquad \begin{cases} r = \sqrt{x^2 + y^2} \\ \operatorname{tg}\varphi = \dfrac{y}{x}. \end{cases} \tag{6.4}$$

Замечание. В задачах, связанных с перемещением материальной точки на плоскости, часто требуют, чтобы переменная φ менялась в пределах $-\infty < \varphi < +\infty$.

Лекция 7

СКАЛЯРНОЕ ПРОИЗВЕДЕНИЕ ВЕКТОРОВ И ЕГО ПРИЛОЖЕНИЯ. ВЕКТОРНОЕ ПРОИЗВЕДЕНИЕ ВЕКТОРОВ И ЕГО ПРИЛОЖЕНИЯ

Определение 1. Углом α между векторами $\vec{a} \neq \vec{0}$ и $\vec{b} \neq \vec{0}$ называется *наименьший* угол между этими векторами, приведенными к общему началу. Его обозначение $\alpha = \widehat{(\vec{a}, \vec{b})}$. Очевидно, что $0 \leqslant \alpha \leqslant \pi$.

Определение 2. Скалярным произведением двух векторов $\vec{a}$ и $\vec{b}$ (пишем $(\vec{a} \cdot \vec{b})$) называется число, равное произведению длин этих векторов на косинус угла между ними

$$(\vec{a}\,\vec{b}) = |\vec{a}|\,|\vec{b}| \cos \widehat{(\vec{a}, \vec{b})}. \tag{7.1}$$

Обращаясь к формуле, по которой вычисляется проекция вектора на ось, можно (7.1) записать в виде

$$\text{(а).}\quad (\vec{a}\,\vec{b}) = |\vec{a}|\,\text{Пр}_{\vec{a}}\,\vec{b};\quad \text{(б).}\quad (\vec{a}\,\vec{b}) = |\vec{b}|\,\text{Пр}_{\vec{b}}\,\vec{a}. \tag{7.2}$$

Таким образом, *скалярное произведение двух векторов равно длине одного из них, умноженной на проекцию второго вектора на ось, направление которой определяется первым вектором.*

Свойство 1. Для любых $\vec{a}$ и $\vec{b}$ имеем $(\vec{a}\,\vec{b}) = (\vec{b}\,\vec{a})$.

В самом деле, из равенства $\cos \widehat{(\vec{a}, \vec{b})} = \cos \widehat{(\vec{b}, \vec{a})}$ следует коммутативность скалярного произведения.

Свойство 2. Для любого $\vec{a} \neq \vec{0} : (\vec{a}\,\vec{a}) = \vec{a}^2 = |\vec{a}|^2 > 0$.

Скалярный квадрат вектора равен нулю тогда и только тогда, когда $\vec{a}$ есть нуль-вектор.

Свойство 3. Для любого $\lambda \in \mathbf{R}$ и любых векторов $\vec{a}$ и $\vec{b}$ $((\lambda\vec{a})\,\vec{b}) = \lambda(\vec{a}\,\vec{b})$.

Действительно, $((\lambda\vec{a})\,\vec{b}) = |\vec{b}|\,\text{Пр}_{\vec{b}}(\lambda\vec{a}) = |\vec{b}|\,\lambda\,\text{Пр}_{\vec{b}}\,\vec{a} = \lambda(\vec{a}\,\vec{b})$.

Свойство 4. Для любых векторов $\vec{a}, \vec{b}, \vec{c}$ выполнено $(\vec{a}(\vec{b}+\vec{c})) = (\vec{a}\,\vec{b}) + (\vec{a}\,\vec{c})$.

На основании (7.2 а) имеем

$$(\vec{a}(\vec{b}+\vec{c})) = |\vec{a}|\,\text{Пр}_{\vec{a}}(\vec{b}+\vec{c}) = |\vec{a}|\,(\text{Пр}_{\vec{a}}\,\vec{b} + \text{Пр}_{\vec{a}}\,\vec{c}) =$$
$$= (\vec{a}\,\vec{b}) + (\vec{a}\,\vec{c}).$$

Теорема. *Два вектора ортогональны друг другу тогда и только тогда, когда их скалярное произведение равно нулю.*

Пусть векторы $\vec{a}$ и $\vec{b}$ не нулевые и ортогональны между собой: $(\widehat{\vec{a}, \vec{b}}) = \frac{\pi}{2}$. Так как $\cos(\widehat{\vec{a}, \vec{b}}) = \cos\frac{\pi}{2} = 0$, то $(\vec{a}\,\vec{b}) = 0$. Если же один из векторов нулевой, то его длина равна нулю и $(\vec{a}\,\vec{b}) = 0$.

Обратно, если $(\vec{a}\,\vec{b}) = |\vec{a}|\,|\vec{b}|\cos(\widehat{\vec{a}, \vec{b}}) = 0$, то хотя бы один из множителей равен нулю. В случае равенства нулю одного из первых двух множителей один из векторов нулевой, который всегда можно считать ортогональным к любому вектору. Если $\cos(\widehat{\vec{a}, \vec{b}}) = 0$, то $(\widehat{\vec{a}, \vec{b}}) = \frac{\pi}{2}$ — векторы ортогональны между собой.

Все перечисленные выше определения и свойства установлены безотносительно к какой-либо системе координат — они *инвариантны* относительно выбора системы координат.

Введем сейчас прямоугольную систему координат и установим, как вычисляется скалярное произведение векторов, заданных своими координатами в базисе $\{\vec{i}, \vec{j}, \vec{k}\}$. Все приводимые ниже формулы выведены для пространства (для плоскости во всех формулах третью координату следует положить равной нулю).

Учитывая, что базис $\{\vec{i}, \vec{j}, \vec{k}\}$ есть ортонормированный базис, получим

$$(\vec{i}\,\vec{i}) = 1,\ (\vec{i}\,\vec{j}) = 0,\ (\vec{i}\,\vec{k}) = 0,\ (\vec{j}\,\vec{j}) = 1,\ (\vec{j}\,\vec{k}) = 0,$$
$$(\vec{k}\,\vec{k}) = 1. \tag{7.3}$$

Если $\vec{a} = x_1\vec{i} + y_1\vec{j} + z_1\vec{k}$, $\vec{b} = x_2\vec{i} + y_2\vec{j} + z_2\vec{k}$, то, воспользовавшись свойствами 3 и 4, имеем

$$(\vec{a}\,\vec{b}) = x_1x_2(\vec{i}\,\vec{i}) + x_1y_2(\vec{i}\,\vec{j}) + x_1z_2(\vec{i}\,\vec{k}) + y_1x_2(\vec{j}\,\vec{i}) +$$
$$+ y_1y_2(\vec{j}\,\vec{j}) + y_1z_2(\vec{j}\,\vec{k}) + z_1x_2(\vec{k}\,\vec{i}) + z_1y_2(\vec{k}\,\vec{j}) + z_1z_2(\vec{k}\,\vec{k}).$$

С учетом (7.3) и свойства 1 окончательно получим

$$(\vec{a}\,\vec{b}) = x_1x_2 + y_1y_2 + z_1z_2. \tag{7.4}$$

Итак, *скалярное произведение векторов в ортонормированном базисе равно сумме произведений одноименных координат этих векторов.*

На основании (7.4) могут быть тотчас получены вычислительные формулы для длины вектора, орта вектора, проекции вектора, косинуса угла между векторами и другие геометрические приложения.

Длина вектора $\vec{a}(x_1, y_1, z_1)$

$$|\vec{a}| = \sqrt{(\vec{a}\,\vec{a})} = \sqrt{x_1^2 + y_1^2 + z_1^2}. \tag{7.5}$$

Орт вектора $\vec{a}(x_1, y_1, z_1)$ $(\vec{a} \neq \vec{0})$

$$\vec{a}^0 = \frac{\vec{a}}{|\vec{a}|} = \frac{x_1\vec{i} + y_1\vec{j} + z_1\vec{k}}{\sqrt{x_1^2 + y_1^2 + z_1^2}}. \tag{7.6}$$

Поэтому для направляющих косинусов вектора $\vec{a}$ имеем

$$\cos\alpha = \frac{x_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}},\quad \cos\beta = \frac{y_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}},$$
$$\cos\gamma = \frac{z_1}{\sqrt{x_1^2 + y_1^2 + z_1^2}}. \tag{7.7}$$

Проекция вектора $\vec{b}(x_2, y_2, z_2)$ на ось с направлением $\vec{a}(x_1, y_1, z_1)$ $(\vec{a} \neq \vec{0})$

$$\text{Пр}_{\vec{a}}\vec{b} = \frac{(\vec{a}\,\vec{b})}{|\vec{a}|} = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2}}. \tag{7.8}$$

Косинус угла между векторами $(\vec{a} \neq \vec{0},\ \vec{b} \neq \vec{0})$

$$\cos(\widehat{\vec{a}, \vec{b}}) = \frac{(\vec{a}\,\vec{b})}{|\vec{a}|\,|\vec{b}|} = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2}\sqrt{x_2^2 + y_2^2 + z_2^2}}. \tag{7.9}$$

На основании доказанной выше теоремы получаем, что необходимым и достаточным условием ортогональности двух векторов является следующее условие:

$$x_1x_2 + y_1y_2 + z_1z_2 = 0. \tag{7.10}$$

Пример 1. Найти орт вектора $\vec{a}$ (1, —1, 5).

Согласно (7.6) имеем

$$\vec{a}^0 = \frac{\vec{i} - \vec{j} + 5\vec{k}}{\sqrt{1^2 + (-1)^2 + 5^2}} = \frac{1}{\sqrt{27}}\vec{i} - \frac{1}{\sqrt{27}}\vec{j} + \frac{5}{\sqrt{27}}\vec{k}.$$

Пример 2. Дан вектор $\vec{a}$ $(6, -18, z)$. Найти z, если $|\vec{a}| = 21$.

В силу (7.5) имеем

$$21 = \sqrt{6^2 + (-18)^2 + z^2},\quad z = \pm\sqrt{441 - 360} = \pm 9.$$

Понятие скалярного произведения векторов пришло из физики, и мы остановимся на одном из физических

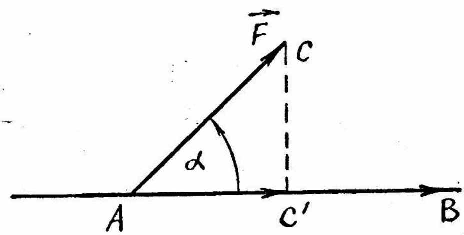


Рис. 12.

приложений скалярного произведения для подсчета *работы силы.*

Пусть требуется вычислить работу W силы $\vec{F}$ по перемещению материальной точки из точки A в точку B по прямолинейному пути (рис. 12).

Если бы материальная точка двигалась по направлению действия силы $\vec{F}$ (угол $\alpha = 0$), то, по определению, работа силы равна произведению величины силы на длину перемещения: $W = |\vec{F}|\,|\overrightarrow{AB}| = (\vec{F}\cdot\overrightarrow{AB})$. Если же точка движется под углом α к направлению силы, то работает только составляющая $\overrightarrow{AC'}$, направленная по линии перемещения $\overrightarrow{AB}$. Перпендикулярная составляющая силы уравновешивается сопротивлением. Поэтому

$$W = (\text{Пр}_{\overrightarrow{AB}}\,\vec{F})\,|\overrightarrow{AB}| = (\vec{F}\cdot\overrightarrow{AB}). \qquad (7.11)$$

Другие физические приложения скалярного произведения векторов мы находим в курсе „Общей физики“, читаемом параллельно с данным курсом.

Переходим к определению векторного произведения двух векторов.

Определение. *Векторным произведением* векторов $\vec{a}$ и $\vec{b}$ называется *вектор* $\vec{x}$, который: (1) перпендикулярен к плоскости векторов $\vec{a}$ и $\vec{b}$; (2) $|\vec{x}| = |\vec{a}|\,|\vec{b}|\sin(\widehat{\vec{a}, \vec{b}})$; (3) направлен так, что тройка $\{\vec{a}, \vec{b}, \vec{x}\}$ — правая (рис. 13).

Векторное произведение обозначается символом $\vec{x} = [\vec{a}\,\vec{b}]$ (либо $\vec{x} = \vec{a}\times\vec{b}$).

Приведенные условия (1)—(3) однозначно определяют векторное произведение, если сомножители — ненулевые векторы. Если хотя бы один из множителей нуль-вектор, векторное произведение, по определению, есть нулевой вектор.

Отметим также, что из условия (2) вытекает: *модуль векторного произведения численно равен площади параллелограмма, построенного на векторах $\vec{a}$ и $\vec{b}$.*

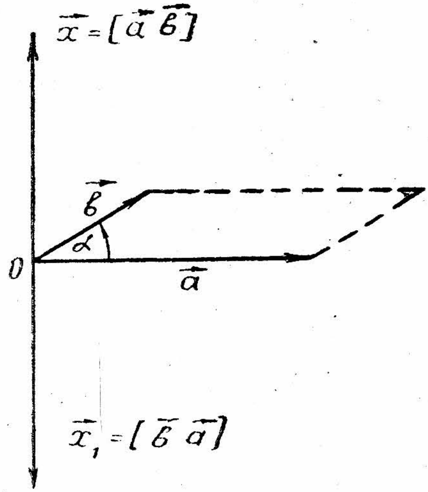


Рис. 13.

Рассмотрим свойства векторного произведения.

Свойство 1. Для любых $\vec{a}$ и $\vec{b}$: $[\vec{a}\,\vec{b}] = -[\vec{b}\,\vec{a}]$. Действительно, для $\vec{x}_1 = [\vec{b}\,\vec{a}]$ выполнены условия (1), (2). Но, чтобы тройка $\{\vec{b}, \vec{a}, \vec{x}_1\}$ была правой, вектор $\vec{x}_1$ должен быть направлен в сторону, противоположную вектору $\vec{x} = [\vec{a}\,\vec{b}]$.

Свойство 2. Для любого $\lambda \in \mathbf{R}$ и любых $\vec{a}$ и $\vec{b}$ имеем

$$[(\lambda\vec{a})\,\vec{b}] = \lambda[\vec{a}\,\vec{b}] = [\vec{a}\,(\lambda\vec{b})]. \tag{7.12}$$

При $\lambda = 0$ справедливость равенства очевидна. Пусть $\lambda > 0$. Тогда $\lambda\vec{a}$ имеет то же направление, что и вектор $\vec{a}$. Длины векторов $[(\lambda\vec{a})\,\vec{b}]$ и $\lambda[\vec{a}\,\vec{b}]$ совпадают, так как $|[(\lambda\vec{a})\,\vec{b}]| = |\lambda\vec{a}|\,|\vec{b}|\sin(\widehat{\vec{a}, \vec{b}}) = \lambda|\vec{a}|\,|\vec{b}|\sin(\widehat{\vec{a}, \vec{b}}) = |\lambda[\vec{a}\,\vec{b}]|$. Направления их также совпадают (ориентация тройки не меняется). Аналогичные рассуждения имеют место и при $\lambda < 0$.

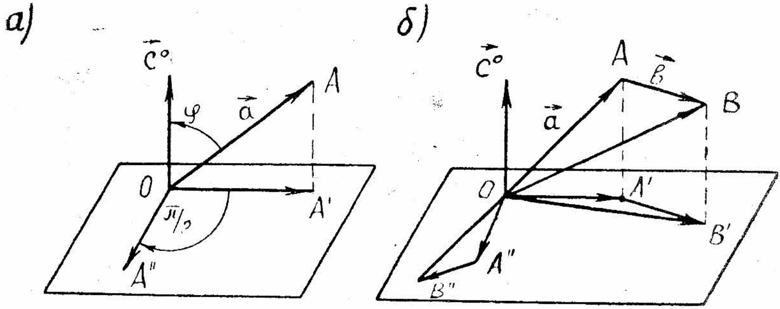


Рис. 14.

Свойство 3. Для любых векторов $\vec{a}, \vec{b}, \vec{c}$ имеем

$$[(\vec{a}+\vec{b})\,\vec{c}] = [\vec{a}\,\vec{c}] + [\vec{b}\,\vec{c}]. \tag{7.13}$$

Предварительно докажем, что имеет место равенство

$$[(\vec{a}+\vec{b})\,\vec{c}^0] = [\vec{a}\,\vec{c}^0] + [\vec{b}\,\vec{c}^0], \tag{7.14}$$

где $\vec{c}^0$ — орт вектора $\vec{c}$. Умножив затем (7.14) на $\lambda = |\vec{c}|$, получим выполнение (7.13).

Для доказательства (7.14) заметим, что вектор $[\vec{a}\,\vec{c}^0]$ можно построить следующим образом. На плоскость, перпендикулярную к $\vec{c}^0$, спроектируем направленный отрезок $\overrightarrow{OA} = \vec{a}$. Затем повернем по часовой стрелке вектор $\overrightarrow{OA'}$ на угол $\frac{\pi}{2}$ и получим вектор $\overrightarrow{OA''}$ (рис. 14а). Имеем: $\overrightarrow{OA''} = [\vec{a}\,\vec{c}^0]$, так как (1) $\overrightarrow{OA''}$ перпендикулярен $\vec{a}$ и $\vec{c}^0$, (2) $|\overrightarrow{OA''}| = |\overrightarrow{OA'}| = |\vec{a}|\cos\left(\frac{\pi}{2} - \varphi\right) = |\vec{a}|\,|\vec{c}^0|\sin\varphi$, (3) тройка $\{\vec{a}, \vec{c}^0, \overrightarrow{OA''}\}$ — правая. Спроектируем далее на плоскость, перпендикулярную к $\vec{c}^0$, векторы $\vec{a}, \vec{b}$, $(\vec{a}+\vec{b})$ (рис. 14б) и получим векторы $\overrightarrow{OA'}, \overrightarrow{A'B'}, \overrightarrow{OB'}$. После поворота на $\alpha = \frac{\pi}{2}$ этих векторов по часовой стрелке можно записать, что $\overrightarrow{OA''} + \overrightarrow{A''B''} = \overrightarrow{OB''}$. Обращаясь к высказанному замечанию, заключаем, что справедливо равенство (7.14), а после умножения его на $\lambda = |\vec{c}|$ убеждаемся в выполнении равенства (7.13).

Теорема. *Необходимым и достаточным условием коллинеарности двух векторов является равенство нулю их векторного произведения.*

Пусть векторы $\vec{a}$ и $\vec{b}$ коллинеарны. Следовательно, либо $(\widehat{\vec{a}, \vec{b}}) = 0$, либо $(\widehat{\vec{a}, \vec{b}}) = \pi$. В обоих случаях $\sin(\widehat{\vec{a}, \vec{b}}) = 0$. Это означает, что $|[\vec{a}\,\vec{b}]| = 0$. Векторное произведение есть нуль-вектор.

Пусть, обратно, $[\vec{a}\,\vec{b}] = \vec{0}$. Тогда $|[\vec{a}\,\vec{b}]| = |\vec{a}||\vec{b}| \times \sin(\widehat{\vec{a}, \vec{b}}) = 0$. Если хотя бы один из первых сомножителей равен нулю, то данный вектор является нулевым и коллинеарность установлена. Если $\sin(\widehat{\vec{a}, \vec{b}}) = 0$, то (1) $(\widehat{\vec{a}, \vec{b}}) = 0$, (2) $(\widehat{\vec{a}, \vec{b}}) = \pi$. Векторы коллинеарны.

Изложенные свойства векторного произведения *инвариантны* относительно выбора системы координат. Пусть задана прямоугольная система координат $\{0, \vec{i}, \vec{j}, \vec{k}\}$. Легко проверить, что выполнены следующие условия:

$$\left.\begin{aligned} &[\vec{i}\,\vec{i}] = \vec{0}, \quad [\vec{j}\,\vec{j}] = \vec{0}, \quad [\vec{k}\,\vec{k}] = \vec{0} \\ &[\vec{i}\,\vec{j}] = \vec{k}, \quad [\vec{j}\,\vec{k}] = \vec{i}, \quad [\vec{k}\,\vec{i}] = \vec{j} \end{aligned}\right\}. \tag{7.15}$$

Пользуясь свойствами 1, 2, 3 и таблицей (7.15), для векторов $\vec{a} = x_1\vec{i} + y_1\vec{j} + z_1\vec{k}$ и $\vec{b} = x_2\vec{i} + y_2\vec{j} + z_2\vec{k}$ получим

$$\begin{aligned}[\vec{a}\,\vec{b}] &= x_1x_2[\vec{i}\,\vec{i}] + x_1y_2[\vec{i}\,\vec{j}] + x_1z_2[\vec{i}\,\vec{k}] + y_1x_2[\vec{j}\,\vec{i}] + \\ &+ y_1y_2[\vec{j}\,\vec{j}] + y_1z_2[\vec{j}\,\vec{k}] + z_1x_2[\vec{k}\,\vec{i}] + z_1y_2[\vec{k}\,\vec{j}] + \\ &+ z_1z_2[\vec{k}\,\vec{k}] = \vec{i}\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} - \vec{j}\begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix} + \vec{k}\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}. \end{aligned} \tag{7.16}$$

Обращаясь к свойству разложения определителя по элементам строки, окончательно получим формулу вычисления векторного произведения векторов, заданных своими координатами в ортонормированном базисе

$$[\vec{a}\,\vec{b}] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}. \qquad (7.17)$$

Заметим, что хотя в первой строке определителя стоят векторы (а не числа!), запись (7.17) законная, так как операции умножения вектора на число и суммы векторов подчиняются тем же правилам, что и числовые элементы.

Поскольку условием коллинеарности двух векторов является равенство нулю их векторного произведения, то, приравнивая определитель в правой части (7.16) нулю и учитывая линейную независимость векторов $\vec{i}, \vec{j}, \vec{k}$, получим, что ранг матрицы $\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{bmatrix}$ равен 1 и, следовательно, *условие*

$$\text{ранг} \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{bmatrix} = 1 \qquad (7.18)$$

является необходимым и достаточным условием коллинеарности векторов $\vec{a}$ *и* $\vec{b}$.

Одним из геометрических приложений векторного произведения является вычисление с его помощью площади треугольника с вершинами в точках

$$A_1(x_1; y_1; z_1),\ A_2(x_2; y_2; z_2),\ A_3(x_3; y_3; z_3).$$

Имеем: $\vec{a} = \overrightarrow{A_1A_2} = \vec{i}(x_2 - x_1) + \vec{j}(y_2 - y_1) + \vec{k}(z_2 - z_1)$, $\vec{b} = \overrightarrow{A_1A_3} = \vec{i}(x_3 - x_1) + \vec{j}(y_3 - y_1) + \vec{k}(z_3 - z_1)$. В силу того, что площадь треугольника равна половине площади параллелограмма, построенного на векторах $\overrightarrow{A_1A_2}$, $\overrightarrow{A_1A_3}$, получим

$$S_\Delta = \frac{1}{2}\left|[\vec{a}\,\vec{b}]\right| = \frac{1}{2}\left\| \begin{matrix} \vec{i} & \vec{j} & \vec{k} \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{matrix} \right\|. \qquad (7.19)$$

Если речь идет о площади треугольника на плоскости, заданного своими вершинами $A_1(x_1; y_1)$, $A_2(x_2; y_2)$, $A_3(x_3; y_3)$, то в формуле (7.19) необходимо положить $z_1 = z_2 = z_3 = 0$ и разложить определитель по элементам последнего столбца

$$S_\Delta = \frac{1}{2}\left|\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_2 - x_1 & y_2 - y_1 & 0 \\ x_3 - x_1 & y_3 - y_1 & 0 \end{vmatrix}\right| = \frac{1}{2}\left|\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}\right|. \tag{7.20}$$

Одним из физических приложений является подсчет *момента силы* с помощью векторного произведения.

Пусть твердое тело закреплено в точке A и в точке B приложена сила $\vec{F}$. Вращающий момент, возникающий в этом случае, вычисляется по следующей формуле (так показывает опыт):

$$\vec{M} = \overrightarrow{AB} \times \vec{F}. \tag{7.21}$$

Лекция 8

СМЕШАННОЕ ПРОИЗВЕДЕНИЕ ВЕКТОРОВ. ДВОЙНОЕ ВЕКТОРНОЕ ПРОИЗВЕДЕНИЕ. ПРЕОБРАЗОВАНИЕ ПРЯМОУГОЛЬНОЙ СИСТЕМЫ КООРДИНАТ НА ПЛОСКОСТИ И В ПРОСТРАНСТВЕ

Зная операции скалярного и векторного умножения двух векторов, что можно сказать о комбинированных произведениях трех векторов? Имеются следующие возможности для комбинированного произведения: (1) $(\vec{a}\,\vec{b})\,\vec{c}$; (2) $([\vec{a}\,\vec{b}]\,\vec{c})$; (3) $[\vec{a}\,[\vec{b}\,\vec{c}]]$. В первом случае ответ простой—получаем вектор, коллинеарный вектору $\vec{c}$. Случаи (2) и (3) требуют более подробного рассмотрения.

Определение. *Смешанным произведением* трех векторов $\vec{a}$, $\vec{b}$, $\vec{c}$ называется число, получаемое от умножения вектора $[\vec{a}\,\vec{b}]$ скалярно на $\vec{c}$. Оно обозначается символом $(\vec{a}\,\vec{b}\,\vec{c})$, т. е. $(\vec{a}\,\vec{b}\,\vec{c}) \overset{\text{def}}{=} ([\vec{a}\,\vec{b}]\,\vec{c})$.

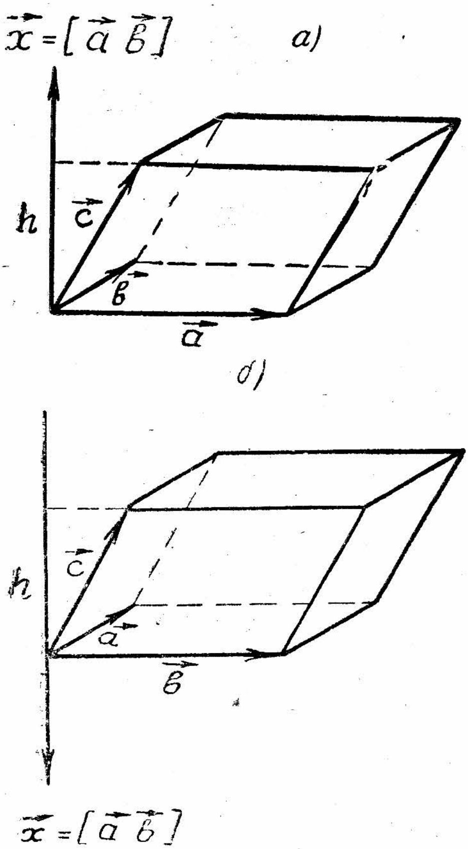


Рис. 15.

Выясним геометрический смысл смешанного произведения, считая, что $\{\vec{a}, \vec{b}, \vec{c}\}$ — не компланарная тройка векторов. Учитывая, что вектор $\vec{x} = [\vec{a}\,\vec{b}]$ имеет длину, равную численно площади параллелограмма, построенного на $\vec{a}$ и $\vec{b}$, и перпендикулярен к плоскости параллелограмма, из равенства

$$([\vec{a}\,\vec{b}]\,\vec{c}) = (\vec{x}\,\vec{c}) = |\vec{x}|\,\text{Пр}_{\vec{x}}\,\vec{c} = \pm\, h \cdot S \qquad (8.1)$$

выводим, что в *случае правой тройки смешанное произведение равно объему V параллелепипеда, построенного на векторах $\vec{a}, \vec{b}, \vec{c}$, а в случае левой тройки — объему параллелепипеда, взятому со знаком минус* (рис. 15).

Теорема. *Тройка векторов $\{\vec{a}, \vec{b}, \vec{c}\}$ компланарна тогда и только тогда, когда их смешанное произведение равно нулю.*

Необходимость. Пусть тройка $\{\vec{a}, \vec{b}, \vec{c}\}$ компланарна. Это может осуществиться в трех случаях: (1) один из векторов есть нуль-вектор, (2) пара векторов коллинеарна, (3) векторы лежат или параллельны одной плоскости. Во всех трех случаях в соотношениях (8.1) либо $|\vec{x}| = 0$ либо $\text{Пр}_{\vec{x}}\, \vec{c} = 0$ и, следовательно, смешанное произведение равно нулю.

Достаточность. Пусть $(\vec{a}\,\vec{b}\,\vec{c}) = 0$. Это означает, что $|[\vec{a}\,\vec{b}]|\,\text{Пр}_{\vec{x}}\, \vec{c} = 0$. Если первый сомножитель равен нулю, то векторное произведение $[\vec{a}\,\vec{b}]$ равно нулю, векторы $\vec{a}, \vec{b}$ — коллинеарные, а, следовательно, тройка $\{\vec{a}, \vec{b}, \vec{c}\}$ — компланарная. Если $\text{Пр}_{\vec{x}}\, \vec{c} = 0$, то вектор $\vec{c}$ ортогонален к вектору $\vec{x} = [\vec{a}\,\vec{b}]$ и, следовательно, параллелен плоскости параллелограмма, построенного на векторах $\vec{a}$ и $\vec{b}$. Имеем компланарную тройку векторов. Теорема доказана.

Свойство 1. Операции скалярного и векторного умножений в смешанном произведении можно поменять местами, т. е.

$$([\vec{a}\,\vec{b}]\,\vec{c}) = (\vec{a}\,[\vec{b}\,\vec{c}]) = ([\vec{b}\,\vec{c}]\,\vec{a}). \tag{8.2}$$

Справедливость этого равенства следует из того, что ориентация троек $\{\vec{a}, \vec{b}, \vec{c}\}$ и $\{\vec{b}, \vec{c}, \vec{a}\}$ не меняется, а параллелепипед для обеих троек один и тот же.

Свойство 2. Круговая перестановка сомножителей не меняет величины смешанного произведения. Перестановка местами двух соседних сомножителей изменяет знак произведения на противоположный, т. е.

$$(\vec{a}\,\vec{b}\,\vec{c}) = (\vec{b}\,\vec{c}\,\vec{a}) = (\vec{c}\,\vec{a}\,\vec{b}) = -(\vec{b}\,\vec{a}\,\vec{c}) = -(\vec{a}\,\vec{c}\,\vec{b}) = -(\vec{c}\,\vec{b}\,\vec{a}). \tag{8.3}$$

В самом деле, в силу коммутативности скалярного произведения и свойства 1 имеем

$$\left.\begin{aligned} (\vec{a}\,\vec{b}\,\vec{c}) &= ([\vec{a}\,\vec{b}]\,\vec{c}) = (\vec{c}\,[\vec{a}\,\vec{b}]) = (\vec{c}\,\vec{a}\,\vec{b}) \\ (\vec{a}\,\vec{b}\,\vec{c}) &= (\vec{a}\,[\vec{b}\,\vec{c}]) = ([\vec{b}\,\vec{c}]\,\vec{a}) = (\vec{b}\,\vec{c}\,\vec{a}) \end{aligned}\right\}. \tag{8.4}$$

В силу антикоммутативности векторного произведения и равенств (8.4) получим

$$(\vec{a}\,\vec{b}\,\vec{c}) = ([\vec{a}\,\vec{b}]\,\vec{c}) = -([\vec{b}\,\vec{a}]\,\vec{c}) = -(\vec{b}\,\vec{a}\,\vec{c}). \tag{8.5}$$

Пусть выбрана прямоугольная система координат. В ортонормированном базисе $\{\vec{i}, \vec{j}, \vec{k}\}$ векторы $\vec{a}, \vec{b}, \vec{c}$ имеют координаты (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3). Согласно определению смешанного произведения как скалярного произведения векторов $[\vec{a}\,\vec{b}]$ и $\vec{c}$ и выражению (7.16) для $[\vec{a}\,\vec{b}]$ получим

$$(\vec{a}\,\vec{b}\,\vec{c}) = x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} - y_3 \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix} + z_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}. \tag{8.6}$$

Правая часть (8.6) с учетом свойств определителей представляет собой разложение определителя третьего порядка по элементам последней строки. Поэтому

$$(\vec{a}\,\vec{b}\,\vec{c}) = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}. \tag{8.7}$$

Получили компактное выражение смешанного произведения через координаты векторов-сомножителей.

Переходим к рассмотрению третьей возможности комбинированного произведения трех векторов.

Определение. *Двойным векторным произведением* векторов $\vec{a}, \vec{b}, \vec{c}$ называется выражение вида $[\vec{a}\,[\vec{b}\,\vec{c}]]$.

Рассмотрим это произведение в прямоугольной системе координат, когда векторы $\vec{a}, \vec{b}, \vec{c}$ заданы своими координатами: $\vec{a} = x_1\vec{i} + y_1\vec{j} + z_1\vec{k}$, $\vec{b} = x_2\vec{i} + y_2\vec{j} + z_2\vec{k}$, $\vec{c} = x_3\vec{i} + y_3\vec{j} + z_3\vec{k}$.

$$[\vec{a}\,[\vec{b}\,\vec{c}]] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ \begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} & -\begin{vmatrix} x_2 & z_2 \\ x_3 & z_3 \end{vmatrix} & \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} \end{vmatrix}.$$

Раскрывая определитель по элементам первой строки, вычисляя определители второго порядка и добавляя в сомножителях при $\vec{i}, \vec{j}, \vec{k}$ соответственно нули в виде $(x_1x_2x_3 - x_1x_2x_3)$, $(y_1y_2y_3 - y_1y_2y_3)$, $(z_1z_2z_3 - z_1z_2z_3)$, получим

$$[\vec{a}\,[\vec{b}\,\vec{c}]] = \vec{i}\,(y_1x_2y_3 - y_1x_3y_2 - z_1x_3z_2 + z_1x_2z_3 + x_1x_2x_3 - x_1x_2x_3) + \vec{j}\,(z_1y_2z_3 - z_1y_3z_2 - x_1x_2y_3 + x_1x_3y_2 + y_1y_2y_3 - y_1y_2y_3) + \vec{k}\,(x_1x_3z_2 - x_1x_2z_3 - y_1y_2z_3 + y_1y_3z_2 + z_1z_2z_3 - z_1z_2z_3) =$$

$$= (x_2\vec{i} + y_2\vec{j} + z_2\vec{k})(x_1x_3 + y_1y_3 + z_1z_3) - (x_3\vec{i} + y_3\vec{j} + z_3\vec{k})(x_1x_2 + y_1y_2 + z_1z_2).$$

Поскольку $(x_1x_3 + y_1y_3 + z_1z_3) = (\vec{a}\,\vec{c})$, $(x_1x_2 + y_1y_2 + z_1z_2) = (\vec{a}\,\vec{b})$, то окончательно имеем

$$[\vec{a}\,[\vec{b}\,\vec{c}]] = \vec{b}\,(\vec{a}\,\vec{c}) - \vec{c}\,(\vec{a}\,\vec{b}). \tag{8.8}$$

Формула (8.8) носит название *формулы раскрытия двойного векторного произведения по векторам-сомножителям внутреннего векторного произведения.*

Очевидно, что, используя определение смешанного произведения векторов и (8.8), можно рассматривать различные комбинированные произведения четырех и т. д. векторов.

Пример 1. Компланарны ли векторы $\vec{a} = (2, 3, -1)$, $\vec{b} = (1, -1, 3)$, $\vec{c} = (1, 9, -11)$?

Вычислим определитель

$$\begin{vmatrix} 2 & 3 & -1 \\ 1 & -1 & 3 \\ 1 & 9 & -11 \end{vmatrix} = 2(11-27) - 3(-11-3) - 1(9+1) = 0.$$

Векторы компланарны.

Пример 2. Проверить справедливость равенства

$$[[\vec{a}\,\vec{b}]\,\vec{c}] + [[\vec{b}\,\vec{c}]\,\vec{a}] + [[\vec{c}\,\vec{a}]\,\vec{b}] = \vec{0}.$$

Переставим сомножители во внешних векторных произведениях и воспользуемся формулой (8.8). Имеем

$$[\vec{c}\,[\vec{a}\,\vec{b}]] = \vec{a}\,(\vec{c}\,\vec{b}) - \vec{b}\,(\vec{c}\,\vec{a}), \quad [\vec{a}\,[\vec{b}\,\vec{c}]] = \vec{b}\,(\vec{a}\,\vec{c}) - \vec{c}\,(\vec{a}\,\vec{b}),$$

$$[\vec{b}\,[\vec{c}\,\vec{a}]] = \vec{c}\,(\vec{b}\,\vec{a}) - \vec{a}\,(\vec{b}\,\vec{c}).$$

Сложим все три равенства и учтем коммутативность скалярного произведения пары векторов. При сложении правые части взаимно уничтожаются и справедливость написанного равенства доказана.

В заключение лекции остановимся на вопросе о переходе от одной декартовой прямоугольной системы координат к другой. Сначала рассмотрение проведем для плоскости. Пусть $\{O, \vec{i}, \vec{j}\}$ — некоторая прямоугольная система координат на плоскости и пусть $\{O', \vec{i}', \vec{j}'\}$ — другая прямоугольная система координат (рис. 16а). Координаты точки M в первой системе координат есть координаты $\overrightarrow{OM}$ в базисе $\{\vec{i}, \vec{j}\}$ $(\overrightarrow{OM} = \vec{i}x + \vec{j}y)$. Координаты точки M во второй системе координат есть координаты вектора $\overrightarrow{O'M}$ в базисе $\{\vec{i}', \vec{j}'\}$ $(\overrightarrow{O'M} = x'\vec{i}' + y'\vec{j}')$. Установим связь координат $(x; y)$ с координатами $(x'; y')$ точки M. С этой целью заметим (рис. 16б), что

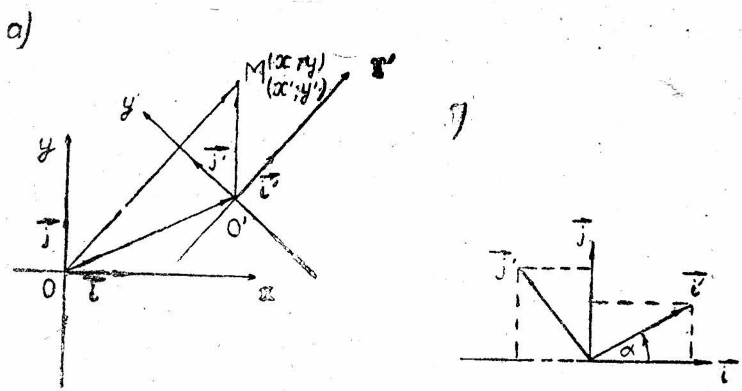


Рис. 16.

$$\left.\begin{aligned}\vec{i}' &= (\text{Пр}_{\vec{i}}\,\vec{i}')\,\vec{i} + (\text{Пр}_{\vec{j}}\,\vec{i}')\,\vec{j} = \vec{i}\cos(\widehat{\vec{i},\vec{i}'}) + \vec{j}\cos(\widehat{\vec{j},\vec{i}'})\\ \vec{j}' &= (\text{Пр}_{\vec{i}}\,\vec{j}')\,\vec{i} + (\text{Пр}_{\vec{j}}\,\vec{j}')\,\vec{j} = \vec{i}\cos(\widehat{\vec{i},\vec{j}'}) + \vec{j}\cos(\widehat{\vec{j},\vec{j}'})\end{aligned}\right\}. \tag{8.9}$$

Если обозначить через α угол между осью Ox и осью $O'x'$, то (8.9) примет вид

$$\left.\begin{aligned}\vec{i}' &= \vec{i}\cos\alpha + \vec{j}\sin\alpha\\ \vec{j}' &= -\vec{i}\sin\alpha + \vec{j}\cos\alpha\end{aligned}\right\}. \tag{8.10}$$

Радиусы-векторы точки M в первой и второй системе координат связаны соотношением

$$\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M}. \tag{8.11}$$

Если $(a; b)$ координаты нового начала O', то (8.11) примет вид

$$x\vec{i} + y\vec{j} = a\vec{i} + b\vec{j} + x'\vec{i}' + y'\vec{j}'. \tag{8.12}$$

Подставив в (8.12) вместо $\vec{i}'$, $\vec{j}'$ их выражения (8.10), окончательно получим

$$x\vec{i} + y\vec{j} = a\vec{i} + b\vec{j} + x'(\vec{i}\cos\alpha + \vec{j}\sin\alpha) + \\ + y'(-\vec{i}\sin\alpha + \vec{j}\cos\alpha). \tag{8.13}$$

В силу линейной независимости векторов $\vec{i}$ и $\vec{j}$ коэффициенты при них в левой и правой частях (8.13) равны между собой. Мы получаем связь координат не штрихованной системы координат с координатами штрихованной системы координат

$$\left.\begin{aligned} x &= x'\cos\alpha - y'\sin\alpha + a \\ y &= x'\sin\alpha + y'\cos\alpha + b \end{aligned}\right\}. \tag{8.14}$$

Если (8.14) разрешить относительно x', y', то получим выражение для новых (штрихованных) координат

$$\left.\begin{aligned} x' &= x\cos\alpha + y\sin\alpha + a' \\ y' &= -x\sin\alpha + y\cos\alpha + b' \end{aligned}\right\}, \tag{8.15}$$

где $a' = -a\cos\alpha - b\sin\alpha$, $b' = a\sin\alpha - b\cos\alpha$ есть координаты точки O относительно штрихованной системы координат.

Очевидно, что если $\alpha = 0$, то совершен лишь параллельный перенос начала координат без вращения осей координат. Полагая $\alpha = 0$ в (8.14), (8.15) и выражениях для a' и b', получим

$$\left.\begin{aligned} x &= x' + a \\ y &= y' + b \end{aligned}\right\} \quad \left.\begin{aligned} x' &= x - a \\ y' &= y - b \end{aligned}\right\}. \tag{8.16}$$

Если $a = b = 0$, $\alpha \neq 0$, то перенос начала координат не совершается, происходит лишь поворот осей. Для него имеем

$$\left.\begin{aligned} x &= x'\cos\alpha - y'\sin\alpha \\ y &= x'\sin\alpha + y'\cos\alpha \end{aligned}\right\} \quad \left.\begin{aligned} x' &= x\cos\alpha + y\sin\alpha \\ y' &= -x\sin\alpha + y\cos\alpha \end{aligned}\right\}. \tag{8.17}$$

Осуществим сейчас переход в пространстве от одной прямоугольной системы координат $\{O, \vec{i}, \vec{j}, \vec{k}\}$ к другой прямоугольной системе координат $\{O', \vec{i}', \vec{j}', \vec{k}'\}$ без переноса начала координат. $(\overrightarrow{OO'} = \vec{0})$ и с сохранением ориентации (обе тройки векторов — правые).

Очень часто в приложениях формулы перехода, связывающих штрихованные и нештрихованные координаты, требуется записать через три независимых параметра — углы Эйлера.

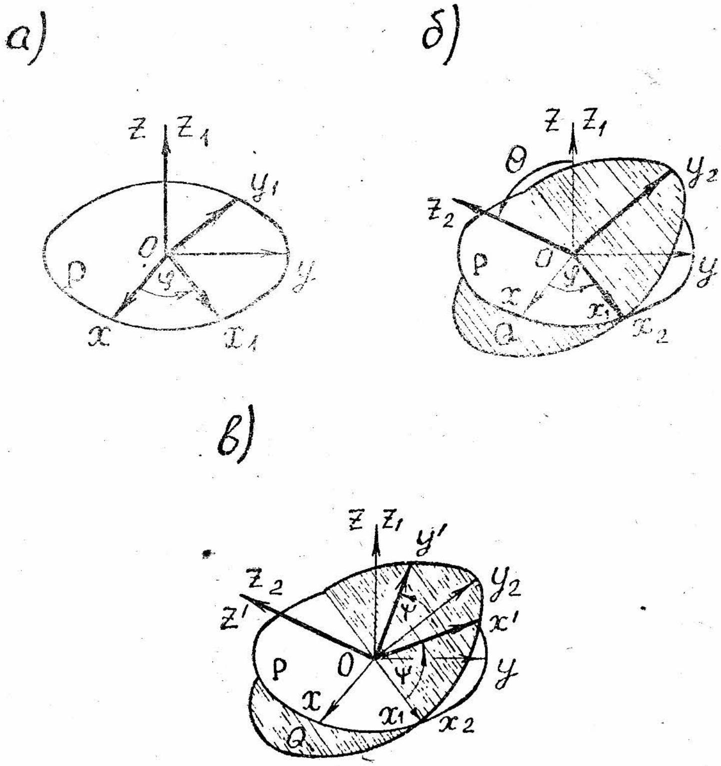


Рис. 17.

С этой целью переход от первого ортонормированного базиса ко второму разобьем на три этапа (рис. 17).

Первый этап (рис. 17а). Повернем вокруг оси Oz на угол φ оси Ox и Oy. Новые оси обозначим через Ox_1, Oy_1, Oz_1. Согласно формулам (8.17) поворота осей на плоскости P имеем

$$\left.\begin{aligned} x &= x_1 \cos\varphi - y_1 \sin\varphi \\ y &= x_1 \sin\varphi + y_1 \cos\varphi \\ z &= z_1 \end{aligned}\right\} . \qquad (8.18)$$

Второй этап (рис. 17б). Повернем вокруг оси Ox_1 оси Oy_1, Oz_1 на угол θ. Новые оси обозначим через Ox_2, Oy_2, Oz_2. Имеем

$$
\left.\begin{aligned}
x_1 &= x_2 \\
y_1 &= y_2 \cos\theta - z_2 \sin\theta \\
z_1 &= y_2 \sin\theta + z_2 \cos\theta
\end{aligned}\right\}. \tag{8.19}
$$

Третий этап (рис. 17в). Повернем вокруг оси Oz_2 в плоскости Q оси Ox_2, Oy_2 на угол ψ. Получим оси Ox', Oy', Oz'. Формулы перехода следующие:

$$
\left.\begin{aligned}
x_2 &= x' \cos\psi - y' \sin\psi \\
y_2 &= x' \sin\psi + y' \cos\psi \\
z_2 &= z'
\end{aligned}\right\}. \tag{8.20}
$$

Подставляя сейчас (8.20) в (8.19), а затем полученный результат в (8.18), получим следующий переход от одной прямоугольной системы координат в пространстве к другой прямоугольной системе без переноса начала координат:

$$
\left\{\begin{aligned}
x &= (\cos\varphi\cos\psi - \sin\varphi\cos\theta\sin\psi)\, x' - \\
&\quad - (\cos\varphi\sin\psi + \sin\varphi\cos\theta\cos\psi)\, y' + z' \sin\varphi\sin\theta, \\
y &= (\sin\varphi\cos\psi + \cos\varphi\cos\theta\sin\psi)\, x' - \\
&\quad - (\sin\varphi\sin\psi - \cos\varphi\cos\theta\cos\psi)\, y' - z' \cos\varphi\sin\theta, \\
z &= x' \sin\theta\sin\psi + y' \sin\theta\cos\psi + z' \cos\theta.
\end{aligned}\right. \tag{8.21}
$$

Чтобы получить выражение x', y', z' через „старые“ координаты x, y, z, необходимо в (8.21) заменить φ на $-\varphi$, θ на $-\theta$, ψ на $-\psi$, штрихованные координаты на нештрихованные.

В случае, если осуществлен и перенос начала координат ($\overrightarrow{OO'} = a\vec{i} + b\vec{j} + c\vec{k}$), то к правым частям соотношений (8.21) нужно добавить соответственно слагаемые a, b, c.

На этом подготовительный материал (играющий также значительную самостоятельную роль в приложениях совершенно независимо от излагаемого ниже) окончен, и мы можем перейти к непосредственным вопросам аналитической геометрии на плоскости и в пространстве.

Лекция 9

ПРОСТЕЙШИЕ ЗАДАЧИ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ НА ПЛОСКОСТИ. ЛИНИИ НА ПЛОСКОСТИ И ИХ УРАВНЕНИЯ. КАНОНИЧЕСКОЕ И ОБЩЕЕ УРАВНЕНИЕ ПРЯМОЙ НА ПЛОСКОСТИ

Как нами установлено ранее, при помощи системы координат устанавливается взаимно-однозначное соответствие между геометрическими образами — точками и алгебраическими объектами — числами. Показано, что каждой точке плоскости в прямоугольных и полярных координатах соответствует пара чисел, взятых в определенном порядке, и обратно, каждой паре чисел соответствует единственная точка плоскости. Взаимно-однозначное соответствие между точками и их координатами сводит изучение геометрических свойств различных объектов к изучению аналитических соотношений между координатами точек множеств, задающих рассматриваемые геометрические объекты.

Прежде чем переходить к установлению соответствия между линиями на плоскости и уравнениями с двумя переменными x и y или r и φ, остановимся на трех простейших фактах, знание которых необходимо при решении задач аналитической геометрии на плоскости.

Согласно определению 3 из 5-й лекции расстояние между точками $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$ равно модулю вектора $\overrightarrow{M_1M_2}$, т. е. $d(M_1, M_2) = d(M_2, M_1) = |\overrightarrow{M_1M_2}|$.

Поскольку

$$\overrightarrow{M_1M_2} = \overrightarrow{OM_2} - \overrightarrow{OM_1} = (x_2 - x_1)\vec{i} + (y_2 - y_1)\vec{j},$$

то согласно (7.5) имеем

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \qquad (9.1)$$

Пусть задан, далее, отрезок как упорядоченная пара точек $A(x_1; y_1)$, $B(x_2; y_2)$. Пусть задана также третья точка $C(x; y)$, лежащая на прямой, соединяющей точки A и B, и не совпадающая с концом B отрезка (точка C может находиться как внутри отрезка, так и вне его). Рассмотрим векторы $\overrightarrow{AC}$ и $\overrightarrow{CB}$. Они коллинеарны, так как находятся на одной прямой. Поэтому

$$\overrightarrow{AC} = \lambda\overrightarrow{CB}. \tag{9.2}$$

Определение. Число λ в (9.2) называется *отношением, в котором точка C делит отрезок AB*.

Отметим, что поскольку точки A и B не совпадают между собой, то $\lambda \neq -1$. В координатах (9.2) имеет вид $(x - x_1)\vec{i} + \vec{j}(y - y_1) = \lambda((x_2 - x)\vec{i} + (y_2 - y)\vec{j})$.

В силу линейной независимости векторов $\vec{i}$ и $\vec{j}$ получим

$$\lambda = \frac{x - x_1}{x_2 - x}, \quad \lambda = \frac{y - y_1}{y_2 - y}, \tag{9.3}$$

а также

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}, \quad y = \frac{y_1 + \lambda y_2}{1 + \lambda}. \tag{9.4}$$

При $\lambda = 1$ точка C делит отрезок пополам. Согласно (9.4) координаты середины отрезка имеют вид

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}. \tag{9.5}$$

Наконец, часто приходится иметь дело с формулой вычисления площади треугольника, заданного своими вершинами $A_1(x_1; y_1)$, $A_2(x_2; y_2)$, $A_3(x_3; y_3)$. Она была выведена нами в конце 7-й лекции. Приведем ее

$$S_\Delta = \frac{1}{2}\left|\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}\right|. \tag{9.6}$$

Используя свойства определителей, выражение (9.6) часто записывают в следующей эквивалентной форме:

$$S_\Delta = \frac{1}{2}\left|\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}\right|. \tag{9.7}$$

В самом деле, вычитая первую строку из второй и третьей строк и понижая порядок определителя, приходим к соотношению (9.6).

Следующим этапом является установление между линиями на плоскости и уравнениями с двумя переменными x и y (r и φ) определенного соответствия, что позволяет свести изучение геометрических свойств линий к изучению аналитических свойств соответствующих уравнений. В аналитической геометрии всякую линию рассматривают как геометрическое место точек, обладающих общим свойством. Так, например, пусть задана прямоугольная декартова система координат и требуется вывести уравнение окружности радиуса R, центр которой находится в точке $C(a; b)$.

Определение. Окружностью с центром в точке C называется геометрическое место точек (ГМТ), находящихся на одном и том же расстоянии R от центра.

Пусть x, y — координаты произвольной точки M окружности (*текущая точка M окружности*). Тогда $|\overrightarrow{CM}| = R$ и, следовательно, $|\overrightarrow{CM}|^2 = R^2$. Последнее равенство в координатах запишется в виде

$$(x-a)^2 + (y-b)^2 = R^2. \tag{9.8}$$

Ясно, что данное соотношение выполняется для всех точек окружности и только для них. Равенство (9.8) называется *уравнением окружности* в прямоугольной системе координат. Если центр окружности совпадает с началом координат ($a = b = 0$), то ее уравнение принимает вид

$$x^2 + y^2 = R^2. \tag{9.9}$$

В полярной системе координат уравнение окружности с центром в полюсе имеет наиболее простой вид: $r = R$.

В общем случае посредством уравнения, связывающего координаты x, y (или r, φ) переменной точки M, выражаем общее свойство точек линии.

Определение. Уравнение $F(x, y) = 0$ ($\Phi(r, \varphi) = 0$) в прямоугольной системе координат (в полярной системе координат) называется уравнением линии L, если этому уравнению удовлетворяют координаты любой точки линии и не удовлетворяют координаты ни одной точки, не лежащей на ней.

Если уравнение имеет вид $y = f(x)$, то оно называется *явным* в отличие от *неявного* уравнения $F(x, y) = 0$.

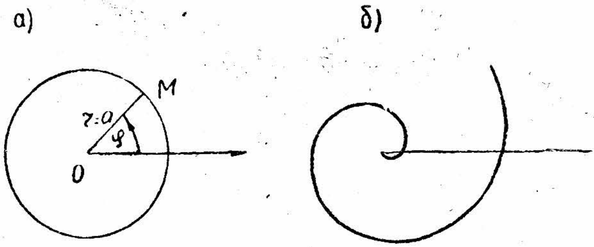


Рис. 18. а) — окружность $r = a$; б) — спираль Архимеда $r = a\varphi$.

Отметим, что множество точек плоскости, координаты которых удовлетворяют некоторому уравнению, может обладать особенностями и не составлять линии на плоскости в том смысле, который придается этому слову. Перечислим наиболее часто встречающиеся особенности: (1) уравнение $F(x, y) = 0$, ($\Phi(r, \varphi) = 0$) не определяет ни одной точки плоскости (например, $x^2 + y^2 + 1 = 0$); (2) уравнение задает одну или несколько точек (например, $x^2 + y^2 = 0$), (3) левая часть уравнения $F(x, y) = 0$ распадается на k штук сомножителей: $F(x, y) = f_1(x, y) f_2(x, y) \ldots$ $\ldots f_k(x, y)$ и тем самым выделяется k ветвей линии с уравнениями $f_1(x, y) = 0, \ldots, f_k(x, y) = 0$. Например, $x^2 - y^2 = 0$ представляется в виде двух ветвей $(x - y) = 0$ и $(x + y) = 0$.

Рассмотрим конкретные примеры часто встречающихся в приложениях уравнений линий в полярной системе координат.

Окружность: $r = a$ ($a > 0$ — const) (рис. 18а).

Спираль Архимеда: $r = a\varphi$ (a — положительная константа) $0 \leqslant \varphi < \infty$ (рис. 18б).

Логарифмическая спираль: $r = ae^{k\varphi}$ (a, k — положительные константы), $-\infty < \varphi < +\infty$ (рис. 19а).

Кардиоида: $r = a(1 + \cos\varphi)$ ($a > 0$ — const), $0 \leqslant \varphi < 2\pi$ (рис. 19б).

Пересечением двух (трех и т. д.) линий называется ГМТ, координаты которых одновременно удовлетворяют уравнениям обеих линий

$$\begin{cases} F_1(x, y) = 0 \\ F_2(x, y) = 0. \end{cases} \tag{9.10}$$

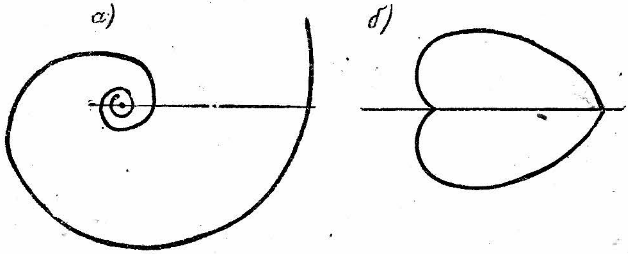


Рис. 19. а) — логарифмическая спираль $r = ae^{k\varphi}$;
б) кардиодида $r = a(1 + \cos\varphi)$.

Часто линию на плоскости удобно (особенно при рассмотрении движения материальных точек в механике) задавать системой уравнений, в которой каждая текущая координата есть некоторая функция одной переменной t (например, в механике времени t), называемой *параметром*. В данном случае уравнения вида

$$\left.\begin{aligned} x &= x(t) \\ y &= y(t) \end{aligned}\right\} \quad (t \in T) \tag{9.11}$$

называются *параметрическими уравнениями линии на плоскости*. При фиксированном значении $t = t_0$ (9.11) определяют координаты $x_0 = x(t_0)$, $y_0 = y(t_0)$ точки M_0 линии.

Исключив из (9.11) параметр t, мы получим уравнение $F(x, y) = 0$ (либо $y = f(x)$), задающее ту же линию на плоскости, что и уравнения (9.11).

Например, уравнения $\left\{\begin{aligned} x &= R\cos t \\ y &= R\sin t \end{aligned}\right.$ $(0 \leqslant t < 2\pi)$ определяют окружность с центром в $O(0; 0)$ и радиуса R, так как, возводя в квадрат уравнения и складывая, получим уравнение $x^2 + y^2 = R^2$, описывающее окружность с центром в начале координат. Параметр t в рассматриваемом случае есть угол между осью Ox и радиус-вектором текущей точки окружности.

В качестве еще одного примера выведем параметрические уравнения *циклоиды*, определяемой как траектория точки M окружности радиуса a при ее каче-

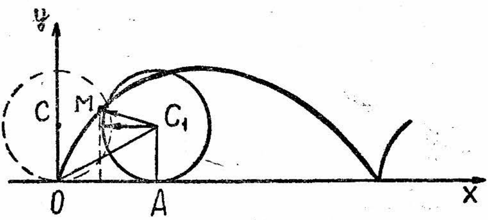


Рис 20.

нии по прямой линии без скольжения (рис. 20). За параметр t выберем увеличивающийся при качении угол $t = \angle AC_1M$.

Имеем $\overrightarrow{OM} = \overrightarrow{OA} + \overrightarrow{AC_1} + \overrightarrow{C_1M}$. Поэтому

$$\left.\begin{aligned} x &= \text{Пр}_{Ox}\,\overrightarrow{OA} + \text{Пр}_{Ox}\,\overrightarrow{AC_1} + \text{Пр}_{Ox}\,\overrightarrow{C_1M} = \\ &\qquad = a\,(t - \sin t) \\ y &= \text{Пр}_{Oy}\,\overrightarrow{OA} + \text{Пр}_{Oy}\,\overrightarrow{AC_1} + \text{Пр}_{Oy}\,\overrightarrow{C_1M} = \\ &\qquad = a\,(1 - \cos t) \end{aligned}\right\}. \quad (t \in [0, 2\pi])$$

Если линия задана в полярной системе координат явным способом $r = r(\varphi)$, то, используя связь декартовых и полярных координат (6.4), легко получить ее параметрическое представление

$$\left.\begin{aligned} x &= r(\varphi)\cos\varphi \\ y &= r(\varphi)\sin\varphi \end{aligned}\right\} \quad (\varphi \in \Phi). \tag{9.12}$$

Если задано неравенство $F(x, y) > 0$ или $F(x, y) < 0$, то его геометрический смысл легко устанавливается. Такие неравенства определяют в прямоугольной системе координат множество точек плоскости—область плоскости, границей которой является линия с уравнением $F(x, y) = 0$. Аналогично и в полярной системе координат. Так, например, неравенство $x^2 + y^2 < R^2$ определяет множество точек, находящихся внутри окружности $x^2 + y^2 = R^2$, называемое *открытым кругом радиуса* R.

В заключение лекции остановимся на выводе канонического и общего уравнения прямой линии на плоскости в прямоугольной системе координат.

Прямую на плоскости можно задать, если фиксировать некоторую точку $M_0(x_0; y_0)$ (*опорная точка*

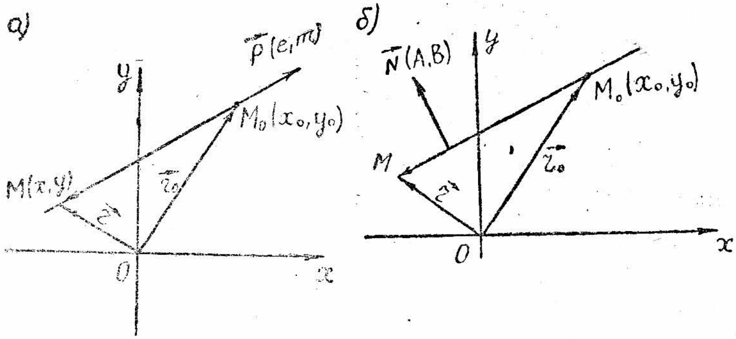


Рис. 21.

прямой) и *направляющий вектор* $\vec{p}(l, m)$ (рис. 21а). Пусть $\vec{r}(x, y)$—радиус-вектор текущей точки искомой прямой линии. Тогда векторы $\vec{r}-\vec{r}_0$ и $\vec{p}$ суть коллинеарные векторы ($\vec{r}-\vec{r}_0=t\vec{p}$), и мы имеем

$$\vec{r}=\vec{r}_0+t\vec{p} \quad (-\infty<t<+\infty). \tag{9.13}$$

Получили *параметрическое уравнение прямой в векторном виде.* В координатной записи (9.13) имеет вид

$$\left.\begin{aligned} x&=x_0+lt\\ y&=y_0+mt \end{aligned}\right\} \quad (-\infty<t<+\infty). \tag{9.14}$$

Коллинеарность векторов $\vec{r}-\vec{r}_0$ и $\vec{p}$ согласно (7.18) может быть выражена и так:

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}. \tag{9.15}$$

Имеем *каноническое уравнение* прямой на плоскости. Если $l=0$ ($m=0$), то соответствующее уравнение прямой имеет вид $x=x_0$ ($y=y_0$).

Прямую на плоскости можно задать другим способом. Зафиксируем некоторую точку прямой $M_0(x_0; y_0)$ и вектор $\vec{N}(A, B)$, *ортогональный* к искомой прямой (он носит название *нормального* вектора прямой). Пусть $\vec{r}(x; y)$—радиус-вектор текущей точки прямой. Тогда (рис. 21б)

$$(\vec{N}(\vec{r}-\vec{r}_0))=0 \tag{9.16}$$

суть *векторное уравнение прямой с нормальным вектором.*

В координатах имеем

$$A(x-x_0)+B(y-y_0)=0. \qquad (9.17)$$

Чтобы описать все множество прямых на плоскости, мы должны произвольным образом задавать $\vec{N}$ и M_0, т. е. считать A, B и $C=-Ax_0-By_0$ произвольными параметрами, пробегающими множество вещественных чисел.

Таким образом, уравнение

$$Ax+By+C=0 \qquad (9.18)$$

называется *общим уравнением* прямой на плоскости. Геометрический смысл коэффициентов при x и y следующий: *коэффициенты A и B суть координаты в ортонормированном базисе $\{\vec{i}, \vec{j}\}$ нормального вектора прямой.* Геометрический смысл параметра C будет выяснен ниже. Этот параметр связан с расстоянием от начала координат до прямой. В частности, при $C=0$ имеем $Ax+By=0$ — уравнение прямой, проходящей через начало координат.

Обратно, если в прямоугольной системе координат на плоскости имеем уравнение 1-й степени вида (9.18), то это будет уравнение прямой линии. Действительно, (9.18) как уравнение с двумя неизвестными всегда имеет решение. Пусть $(x_0; y_0)$ — некоторое решение, т. е. $Ax_0+By_0+C\equiv 0$. Вычитая это тождество из (9.18), получим эквивалентное ему уравнение $A(x-x_0)+$ $+B(y-y_0)=0$, которое может быть записано в виде $(\vec{N}(\vec{r}-\vec{r}_0))=0$, где $\vec{N}=A\vec{i}+B\vec{j}$, $\vec{r}=x\vec{i}+y\vec{j}$ — переменный радиус-вектор, $\vec{r}_0=x_0\vec{i}+y_0\vec{j}$ — фиксированный вектор, задающий точку $M_0(x_0; y_0)$. Очевидно, что данному уравнению удовлетворяют координаты любой точки прямой линии, проходящей через M_0 и имеющей нормальный вектор $\vec{N}$, и не удовлетворяют координаты ни одной точки с радиусом-вектором $\vec{r}'$, не лежащей на прямой, так как $\vec{r}'-\vec{r}_0$ и $\vec{N}$ не ортогональны между собой.

Лекция 10

РАЗЛИЧНЫЕ ВИДЫ УРАВНЕНИЙ ПРЯМОЙ НА ПЛОСКОСТИ И ИХ ГЕОМЕТРИЧЕСКИЕ ПРИЛОЖЕНИЯ. РАССТОЯНИЕ ОТ ТОЧКИ ДО ПРЯМОЙ. ВЗАИМНОЕ РАСПОЛОЖЕНИЕ ДВУХ ПРЯМЫХ НА ПЛОСКОСТИ. УРАВНЕНИЕ ПУЧКА ПРЯМЫХ

Остановимся на трех видах уравнения прямой, имеющих наибольшие приложения.

Уравнение в отрезках. Пусть в общем уравнении прямой $Ax+By+C=0$ коэффициенты $A\neq 0$, $B\neq 0$, $C\neq 0$. Тогда это уравнение можно переписать в виде $\dfrac{x}{\left(-\dfrac{C}{A}\right)}+\dfrac{y}{\left(-\dfrac{C}{B}\right)}=1$. Обозначив $a=-\dfrac{C}{A}$, $b=-\dfrac{C}{B}$, получим уравнение прямой в отрезках

$$\frac{x}{a}+\frac{y}{b}=1. \tag{10.1}$$

Видим, что точки с координатами $(a;0)$ и $(0;b)$ удовлетворяют уравнению (10.1), и, следовательно, уравнение (10.1) описывает прямые, отсекающие от осей координат отрезки длиной $|a|$ и $|b|$.

Уравнение с угловым коэффициентом. Каноническое уравнение (9.15) может быть записано в виде $y-y_0=\dfrac{m}{l}(x-x_0)$, где $l\neq 0$. Вспомним, что

$$l=\text{Пр}_{\vec{i}}\,\vec{p}=|\vec{p}|\cos\varphi,\ m=\text{Пр}_{\vec{j}}\,\vec{p}=|\vec{p}|\sin\varphi,$$

и поэтому коэффициент

$$k=\frac{m}{l}=\operatorname{tg}\varphi, \tag{10.2}$$

где φ (рассматриваемый в пределах $\varphi\in\left[0,\dfrac{\pi}{2}\right)\cup\left(\dfrac{\pi}{2},\pi\right]$), угол между векторами $\vec{i}$ и $\vec{p}$ (угол наклона прямой к оси Ox), называем ***угловым коэффициентом.*** Если

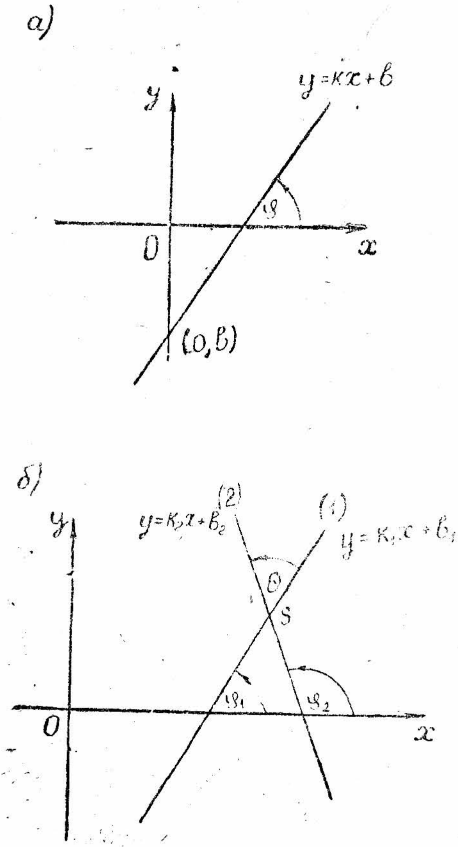


Рис. 22.

обозначить $b = y_0 - \frac{m}{l} x_0$, то уравнение прямой с угловым коэффициентом примет вид

$$y = kx + b. \qquad (10.3)$$

— Легко убедиться в том, что точка с координатами (0; b) удовлетворяет уравнению (10.3), и, следовательно, прямая (10.3) отсекает от оси Oy отрезок длиной $|b|$ (рис. 22а).

Если в общем уравнении прямой $B \neq 0$ (прямая не параллельна оси ординат), то его можно привести к виду $y=-\frac{A}{B}x-\frac{C}{B}$. Обозначив $k=-\frac{A}{B}$, $b=-\frac{C}{B}$, приходим к уравнению прямой с угловым коэффициентом.

Используя геометрический смысл углового коэффициента k, легко получить выражение для тангенса угла между двумя прямыми.

Определение. Углом θ между прямыми, заданными уравнениями с угловыми коэффициентами, будем называть угол, на который надо повернуть против часовой стрелки первую прямую вокруг точки S до совмещения со второй прямой ($0 \leqslant \theta \leqslant \pi$) (рис. 22б).

Пусть заданы прямые $y=k_1x+b_1$ и $y=k_2x+b_2$. Поскольку $\varphi_2=\varphi_1+\theta$ (рис. 22б), то $\operatorname{tg}\theta=\operatorname{tg}(\varphi_2-\varphi_1)$. Учитывая, что $\operatorname{tg}\varphi_1=k_1$, $\operatorname{tg}\varphi_2=k_2$, и применяя формулу для тангенса разности углов, получим

$$\operatorname{tg}\theta=\frac{k_2-k_1}{1+k_1k_2}, \quad (1+k_1k_2\neq 0). \tag{10.4}$$

Если прямые заданы общими уравнениями $A_1x+B_1y+C_1=0$, $A_2x+B_2y+C_2=0$ то, так как $k_1=-\frac{A_1}{B_1}$, $k_2=-\frac{A_2}{B_2}$, из (10.4) имеем

$$\operatorname{tg}\theta=\frac{A_1B_2-B_1A_2}{A_1A_2+B_1B_2}, \quad (A_1A_2+B_1B_2\neq 0). \tag{10.5}$$

Последняя формула годится и в том случае, когда одна из прямых параллельна оси ординат. Формула (10.4) этот случай не затрагивает.

Теорема 1. *Для того, чтобы прямые были параллельными, необходимо и достаточно, чтобы* $k_1=k_2$ $(A_1B_2=B_1A_2)$.

Пусть прямые параллельны. Угол $\theta=0$ (либо $\theta=\pi$), и, следовательно, $\operatorname{tg}\theta=0$. Из (10.4) вытекает, что $k_1=k_2$, а из (10.5) — $A_1B_2=B_1A_2$. Необходимость доказана.

Пусть обратно $k_1=k_2$ $(A_1B_2=B_1A_2)$. Тогда $\operatorname{tg}\theta=0$. Следовательно, прямые параллельны. Достаточность доказана.

Теорема 2. *Для того, чтобы прямые были перпендикулярны между собой, необходимо и достаточно, чтобы* $k_2=-\frac{1}{k_1}$ $(A_1A_2+B_1B_2=0)$.

В самом деле, если прямые ортогональны друг другу, то ортогональны и их нормальные векторы, имеющие для прямых с угловым коэффициентом вид: $\vec{N}_1 = k_1\vec{i} - \vec{j}$, $\vec{N}_2 = k_2\vec{i} - \vec{j}$, а для прямых с общим уравнением $\vec{N}_1 = A_1\vec{i} + B_1\vec{j}$, $\vec{N}_2 = A_2\vec{i} + B_2\vec{j}$. В силу равенства $(\vec{N}_1\vec{N}_2) = 0$ получим, что $k_1k_2 + 1 = 0$ и соответственно $A_1A_2 + B_1B_2 = 0$.

Пусть обратно, $k_2 = -\frac{1}{k_1}$. Это значит $k_1k_2 + 1 = 0$. Поскольку для прямой $y = k_1x + b_1$ нормальный вектор $\vec{N}_1$ имеет координаты $(k_1, -1)$, а для прямой $y = k_2x + b_2$ $\vec{N}_2 = k_2\vec{i} - \vec{j}$, то равенство $k_1k_2 + 1 = 0$ эквивалентно $(\vec{N}_1\vec{N}_2) = 0$. Следовательно, нормальные векторы у прямых ортогональны. Ортогональны и прямые. Аналогичное рассуждение проводится в случае выполнения условия $A_1A_2 + B_1B_2 = 0$, если вспомнить, что $\vec{N}_1 = A_1\vec{i} + B_1\vec{j}$, $\vec{N}_2 = A_2\vec{i} + B_2\vec{j}$.

Нормированное уравнение прямой. Рассмотрим прямую, не проходящую через начало координат, опустим из начала координат перпендикуляр на прямую (рис. 23а) и обозначим через Q точку пересечения перпендикуляра и прямой. Пусть $\vec{n}$ — орт вектора $\overrightarrow{OQ} = \vec{q}$. Для любой точки M прямой вектор $(\vec{r} - \vec{q})$ ортогонален к $\vec{n}$, т. е. $((\vec{r} - \vec{q})\vec{n}) = 0$. Поскольку $\vec{n}$ — орт-вектор, то $\vec{n} = \vec{i}\cos\alpha + \vec{j}\sin\alpha$, где $\alpha = (\widehat{\vec{i}, \vec{n}})$, и $(\vec{q}\,\vec{n}) = |\vec{q}\,|\,|\vec{n}|\cos(\widehat{\vec{q}, \vec{n}}) = p$ (длина перпендикуляра, равная расстоянию от начала координат до прямой). Поэтому равенство $((\vec{r} - \vec{q})\vec{n}) = (\vec{r}\,\vec{n}) - (\vec{q}\,\vec{n}) = 0$ в координатах имеет вид

$$x\cos\alpha + y\sin\alpha - p = 0. \qquad (10.6)$$

Уравнение (10.6) носит название *нормированного* (*нормального*) уравнения прямой.

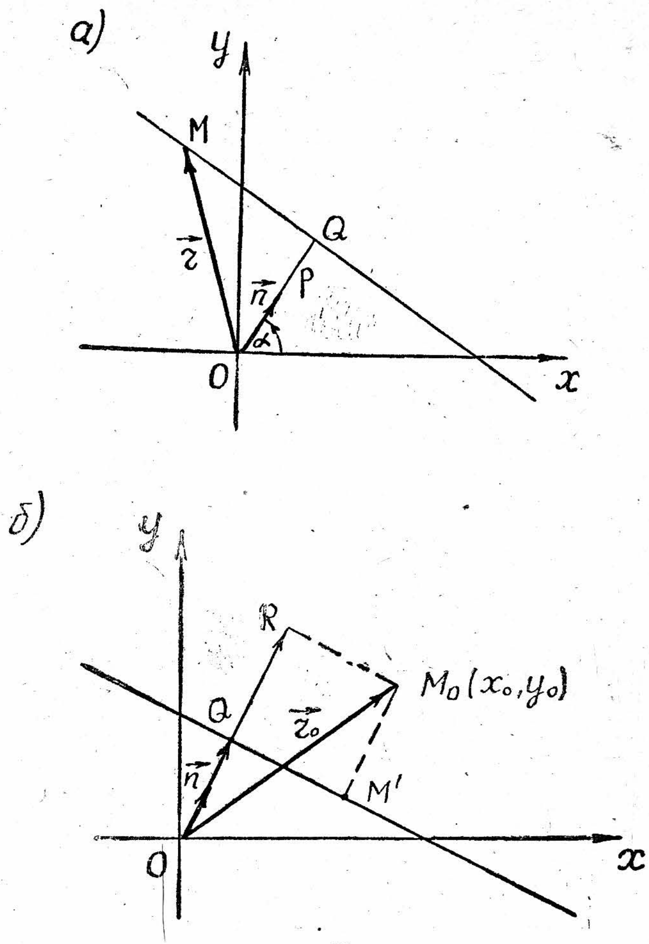


Рис. 23.

Пусть прямая задана общим уравнением $Ax + By + C = 0$ или $(\vec{r}\vec{N}) + C = 0$. Найдем выражение для единичного вектора $\vec{n}$, коллинеарного с $\vec{N}$. Получим, что $\vec{n} = \frac{\vec{N}}{\varepsilon|\vec{N}|} = \frac{A}{\varepsilon\sqrt{A^2 + B^2}}\vec{i} + \frac{B}{\varepsilon\sqrt{A^2 + B^2}}\vec{j}$. Поэтому нормированное уравнение имеет следующий вид:

$$\frac{A}{\varepsilon\sqrt{A^2+B^2}}x+\frac{B}{\varepsilon\sqrt{A^2+B^2}}y+\frac{C}{\varepsilon\sqrt{A^2+B^2}}=0, \qquad (10.7)$$

где $\varepsilon=+1$, если свободный член $C<0$ и $\varepsilon=-1$, если $C>0$. Такой выбор знака перед корнем определяется тем, что в (10.6) свободный член $-p$ всегда отрицателен.

Пример. Пронормировать уравнение прямой $x-y+1=0$. Так как свободный член положителен, то нормирующий множитель будет равен

$$-\frac{1}{\sqrt{1^2+(-1)^2}}=-\frac{1}{\sqrt{2}}.$$

Поэтому нормированное уравнение имеет вид:

$$-\frac{1}{\sqrt{2}}x+\frac{1}{\sqrt{2}}y-\frac{1}{\sqrt{2}}=0.$$

Используя нормированное уравнение прямой, легко вывести формулу вычисления расстояния d от точки до прямой.

Пусть $M_0(x_0;y_0)$ — точка вне прямой $x\cos\alpha+y\sin\alpha-p=0$.

Определение. Отклонением δ точки M_0 от прямой называется число, равное $+d$, если M_0 и начало координат находятся по разные стороны прямой, и равное $-d$, если M_0 и начало координат находятся по одну сторону от прямой.

Теорема. *Для вычисления отклонения δ точки $M_0(x_0;y_0)$ от прямой $x\cos\alpha+y\sin\alpha-p=0$ нужно в левую часть уравнения вместо x и y подставить координаты x_0 и y_0 точки M_0, т. е.*

$$\delta=x_0\cos\alpha+y_0\sin\alpha-p. \qquad (10.8)$$

Для вычисления расстояния от точки до прямой получаем формулу

$$d=|x_0\cos\alpha+y_0\sin\alpha-p|. \qquad (10.9)$$

Докажем теорему. Доказательство проведем для случая, когда начало координат и M_0 находятся по разные стороны прямой (рис. 23б). В другом случае все повторяется буквально. Для рассматриваемой ситуации $\delta=\text{Пр}_{\vec{n}}\overrightarrow{M'M_0}=\text{Пр}_{\vec{n}}\overrightarrow{QR}$. Но $\overrightarrow{QR}=(\text{Пр}_{\vec{n}}\overrightarrow{OM_0})\vec{n}-p\vec{n}=(\vec{r}_0\vec{n})\vec{n}-p\vec{n}$. Поэтому $\delta=(\overrightarrow{QR}\cdot\vec{n})=(\vec{r}_0\vec{n})-p$, так как $(\vec{n}\,\vec{n}) = 1$. Формулы (10.8), а, следовательно, и (10.9) доказаны.

Пример. Найти расстояние от точки P (1, —2) до прямой $2x + y - 3 = 0$.

Нормируем согласно (10.7) данное уравнение:

$$\frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y - \frac{3}{\sqrt{5}} = 0.$$

Подставляем координаты точки P (1, —2). Получим, что

$$\delta = \frac{2}{\sqrt{5}} + \frac{1}{\sqrt{5}}(-2) - \frac{3}{\sqrt{5}} = -\frac{3}{\sqrt{5}},$$

$$d = \left| -\frac{3}{\sqrt{5}} \right| = \frac{3}{\sqrt{5}}.$$

В качестве приложения формулы для отклонения (10.8) выступает вывод *уравнений биссектрис* углов, образованных при пересечении прямых $A_1x + B_1y + C_1 = 0$ и $A_2x + B_2y + C_2 = 0$. Так как биссектриса есть ГМТ равноудаленных от сторон угла, то, учитывая знак отклонения текущей точки биссектрисы $M(X; Y)$ от первой и второй прямой, получим уравнения биссектрис. А именно

$$\left(\frac{A_1}{\varepsilon_1\sqrt{A_1^2 + B_1^2}} - \frac{A_2}{\varepsilon_2\sqrt{A_2^2 + B_2^2}}\right)X +$$
$$+ \left(\frac{B_1}{\varepsilon_1\sqrt{A_1^2 + B_1^2}} - \frac{B_2}{\varepsilon_2\sqrt{A_2^2 + B_2^2}}\right)Y +$$
$$+ \left(\frac{C_1}{\varepsilon_1\sqrt{A_1^2 + B_1^2}} - \frac{C_2}{\varepsilon_2\sqrt{A_2^2 + B_2^2}}\right) = 0$$

для угла, где отклонения совпадают по знаку, а также

$$\left(\frac{A_1}{\varepsilon_1\sqrt{A_1^2 + B_1^2}} + \frac{A_2}{\varepsilon_2\sqrt{A_2^2 + B_2^2}}\right)X +$$
$$+ \left(\frac{B_1}{\varepsilon_1\sqrt{A_1^2 + B_1^2}} + \frac{B_2}{\varepsilon_2\sqrt{A_2^2 + B_2^2}}\right)Y +$$
$$+ \left(\frac{C_1}{\varepsilon_1\sqrt{A_1^2 + B_1^2}} + \frac{C_2}{\varepsilon_2\sqrt{A_2^2 + B_2^2}}\right) = 0$$

для угла, где отклонения разнятся знаком.

П р и м е р. Найти уравнения биссектрис углов, образованных пересечением прямых $x-2y+1=0$ и $2x+y-2=0$.

Р е ш е н и е. Нормируем прямые. Имеем $-\frac{1}{\sqrt{5}}x+\frac{2}{\sqrt{5}}y-\frac{1}{\sqrt{5}}=0$ и $\frac{2}{\sqrt{5}}x+\frac{1}{\sqrt{5}}y-\frac{2}{\sqrt{5}}=0$. Для одинаковых знаков δ получим, что $-\frac{3}{\sqrt{5}}X+\frac{1}{\sqrt{5}}Y+\frac{1}{\sqrt{5}}=0$ или $Y=3X-1$.

Для разных знаков δ получим, что $\frac{1}{\sqrt{5}}X+\frac{3}{\sqrt{5}}Y-\frac{3}{\sqrt{5}}=0$, т. е. $X+3Y-3=0$ или $Y=-\frac{1}{3}X+1$.

Взаимное расположение двух прямых $A_1x+B_1y+C_1=0$ и $A_2x+B_2y+C_2=0$ на плоскости описывается тремя ситуациями: (1) прямые пересекаются; (2) прямые параллельны, но не совпадают; (3) прямые совпадают. Рассмотрение системы уравнений

$$\begin{cases} A_1x+B_1y+C_1=0 \\ A_2x+B_2y+C_2=0 \end{cases} \tag{10.10}$$

с привлечением теории систем линейных уравнений приводит нас к следующим выводам: (1) *если определитель* $\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0$, *то система* (10.10) *имеет единственное решение — прямые пересекаются*; (2) *если* $\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}=0$, *а ранг матрицы* $\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{bmatrix}$ *равен двум, то система не совместна — прямые параллельны и не совпадают*; (3) *если ранг* $\begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{bmatrix}$ *равен* 1, *то прямые совпадают.*

Перейдем к описанию *пучка прямых.*

О п р е д е л е н и е. Множество всех прямых плоскости, проходящих через точку $M_0(x_0;y_0)$ плоскости (называемую *центром*), называется *пучком прямых.*

Как задать пучок прямых? Первый способ состоит в задании центра пучка $M_0(x_0;y_0)$. В таком случае множество всех прямых, проходящих через эту точку, опишется соотношениями

$$\left.\begin{cases} y-y_0=k(x-x_0) \\ x=x_0 \end{cases}\right\} \quad (-\infty<k<+\infty). \qquad (10.11)$$

Второй способ состоит в задании двух фиксированных пересекающихся прямых: $A_1x+B_1y+C_1=0$ и $A_2x+B_2y+C_2=0$ $(A_1B_2-A_2B_1\neq 0)$.

Теорема. *Уравнение пучка прямых с центром в точке пересечения прямых* $A_1x+B_1y+C_1=0$ *и* $A_2x+B_2y+C_2=0$ *(без определения координат точки пересечения) описывается соотношением вида*

$$\lambda(A_1x+B_1y+C_1)+\mu(A_2x+B_2y+C_2)=0, \qquad (10.12)$$

где λ, μ *— произвольные вещественные числа.*

Действительно, (10.12) относится к уравнению вида $Ax+By+C=0$ и, следовательно, есть уравнение прямой (при $\mu=0$ имеет первую *базовую прямую пучка*, при $\lambda=0$ имеет вторую *базовую прямую*). Прямая, описываемая уравнением (10.12), проходит через точку пересечения базовых прямых, ибо координаты точки пересечения обращают в нуль как первую скобку, так и вторую и, следовательно, удовлетворяют (10.12). Наконец, (10.12) описывает все прямые пучка, т. е. какова бы ни была наперед заданная прямая, проходящая через центр пучка, она будет описываться уравнением (10.12) при некоторых значениях параметров λ и μ. В самом деле, наперед заданная прямая пучка однозначно определяется заданием еще одной точки $P(x_1; y_1)$ (через 2 точки проходит единственная прямая!). Если мы покажем, что найдутся такие λ и μ, что координаты точки P удовлетворяют (10.12), то доказательство теоремы будет закончено. Все эти условия соблюдаются, так как, взяв

$$\lambda=-\frac{A_2x_1+B_2y_1+C_2}{A_1x_1+B_1y_1+C_1}\cdot\mu,$$

мы получим

$$\lambda(A_1x_1+B_1y_1+C_1)+\mu(A_2x_1+B_2y_1+C_2)\equiv 0.$$

Пример. Через точку пересечения прямых $x-y+1=0$ и $2x+y+1=0$ проходит прямая, параллельная прямой $3x+y+1=0$. Написать ее уравнение.

Решение. Имеем $\lambda(x-y+1)+\mu(2x+y+1)=0$ или $(\lambda+2\mu)x+(-\lambda+\mu)y+\lambda+\mu=0$. Условие параллельности: $\lambda+2\mu=3$ $-\lambda+\mu=1$, т. е. $\lambda=\frac{1}{3}$, $\mu=\frac{4}{3}$.

Ответ: $3x+y+\frac{5}{3}=0$.

Лекция 11

ПРИВЕДЕНИЕ ОБЩЕГО УРАВНЕНИЯ КРИВОЙ ВТОРОГО ПОРЯДКА К КАНОНИЧЕСКОМУ ВИДУ. ЭЛЛИПС, ЕГО ФОРМА И ГЕОМЕТРИЧЕСКИЕ СВОЙСТВА

В предыдущей лекции мы изучали на плоскости геометрический образ, описываемый в декартовой прямоугольной системе координат *уравнением первой степени*: $Ax+By+C=0$ — *прямую линию*. Сейчас мы перейдем к исследованию линий на плоскости, координаты точек которых в прямоугольной системе координат удовлетворяют уравнению второй степени.

Мы отступаем от традиционного изложения теории кривых второго порядка, как ГМТ, обладающих определенным свойством, и начинаем с классификации кривых второго порядка, приведя общее уравнение к каноническим видам. Полное и последовательное изложение этого вопроса для n-мерного евклидова пространства будет представлено в линейной алгебре (лекция 25). Здесь же проводится хотя и строгое, но с учетом двумерного случая специфическое исследование без привлечения аппарата линейной алгебры. Затем с использованием канонических уравнений устанавливаются геометрические свойства кривых, положенные в основу их определения как геометрического места точек.

Определение. Линия, определяемая в прямоугольной системе координат уравнением

$$Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0, \qquad (11.1)$$

где A, B, C одновременно не обращаются в нуль, называется *кривой второго порядка*, а уравнение (11.1) — ее *общим уравнением*.

Совершенно очевидно, что всякая окружность $(x-a)^2+(y-b)^2=R^2$ есть кривая второго порядка (11.1) при $A=C$, $B=0$, так как после возведения в квадрат имеем $x^2+y^2-2ax-2by+a^2+b^2-R^2=0$.

Обратно, *если в* (11.1) $A=C$, $B=0$, $\frac{F}{A}<0$, *то* (11.1) *есть уравнение окружности*. В самом деле, умножив (11.1) на $\frac{1}{A}$, получим $x^2+y^2-2ax-2by-$ $-c=0 \left(a=-\frac{D}{A},\ b=-\frac{E}{A},\ c=-\frac{F}{A}>0\right)$, которое может быть записано в виде $(x-a)^2+(y-b)^2=c+a^2+b^2$.

Нашей целью является выбор такой новой прямоугольной системы координат $\{O', \vec{i}', \vec{j}', \vec{k}'\}$, относительно которой *уравнение кривой* (11.1) *имеет наименьшее число числовых параметров* (*каноническое уравнение*). Переход осуществляется по формулам (8.14)

$$\left.\begin{aligned} x &= x'\cos\alpha - y'\sin\alpha + a \\ y &= x'\sin\alpha + y'\cos\alpha + b \end{aligned}\right\}. \qquad (11.2)$$

Подставив вместо x и y в (11.1) их выражения (11.2), получим

$$A(x'^2\cos^2\alpha + y'^2\sin^2\alpha + a^2 - 2x'y'\cos\alpha\sin\alpha +$$
$$+2ax'\cos\alpha - 2ay'\sin\alpha) + 2B(x'^2\cos\alpha\sin\alpha + x'y'\cos^2\alpha +$$
$$+ bx'\cos\alpha - x'y'\sin^2\alpha - y'^2\sin\alpha\cos\alpha - by'\sin\alpha +$$
$$+ ax'\sin\alpha + ay'\cos\alpha + ab) + C(x'^2\sin^2\alpha + y'^2\cos^2\alpha +$$
$$+ b^2 + 2x'y'\cos\alpha\sin\alpha + 2bx'\sin\alpha + 2by'\cos\alpha) +$$
$$+ 2D(x'\cos\alpha - y'\sin\alpha + a) + 2E(x'\sin\alpha +$$
$$+ y'\cos\alpha + b) + F = 0.$$

Приведя подобные члены, имеем

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0, \qquad (11.3)$$

где новые коэффициенты выражаются следующим образом:

$$\left.\begin{aligned} A' &= A\cos^2\alpha + 2B\cos\alpha\sin\alpha + C\sin^2\alpha \\ B' &= B\cos 2\alpha + \frac{1}{2}(C-A)\sin 2\alpha \\ C' &= A\sin^2\alpha - 2B\sin\alpha\cos\alpha + C\cos^2\alpha \\ D' &= a(A\cos\alpha + B\sin\alpha) + b(B\cos\alpha + C\sin\alpha) + \\ &\qquad + D\cos\alpha + E\sin\alpha \\ E' &= a(B\cos\alpha - A\sin\alpha) + b(C\cos\alpha - B\sin\alpha) - \\ &\qquad - D\sin\alpha + E\cos\alpha \\ F' &= Aa^2 + 2Bab + Cb^2 + 2Da + 2Eb + F \end{aligned}\right\}. \qquad (11.4)$$

Отметим следующий факт: *определитель* $\Delta=\begin{vmatrix} A & B \\ B & C \end{vmatrix}$, составленный из коэффициентов при второй степени, *есть инвариант относительно преобразований координат* (11.2). В самом деле, подставив в выражение $\Delta' = A'C' - B'^2$ значения A', B', C' из (11.4), получим

$$\Delta' = A'C' - B'^2 = (AC - B^2)\sin^4\alpha + (AC - B^2)\cos^4\alpha + \\ + 2(AC - B^2)\sin^2\alpha\cos^2\alpha = (AC - B^2)(\sin^2\alpha + \cos^2\alpha)^2 = \Delta. \tag{11.5}$$

Если в уравнении $B \neq 0$, то в преобразованиях (11.2) угол α подберем таким, чтобы коэффициент B' в (11.3) обратился в нуль. Для этого достаточно, чтобы $B\cos 2\alpha + \frac{1}{2}(C - A)\sin 2\alpha = 0$, т. е.

$$\operatorname{ctg} 2\alpha = \frac{1}{2B}(A - C). \tag{11.6}$$

Распорядимся сейчас выбором параметров a и b для упрощения вида уравнения (11.3), где $B' = 0$. С этой целью дальнейшее рассмотрение разобьем на 2 случая в зависимости от того, отличен от нуля определитель $\Delta = AC - B^2$ в уравнениях (11.1) или он равен нулю.

(*A*) Пусть $\Delta = AC - B^2 \neq 0$. Заметив, что определитель, составленный из коэффициентов при a и b для D' и E' в (11.4)

$$\begin{vmatrix} A\cos\alpha + B\sin\alpha & B\cos\alpha + C\sin\alpha \\ B\cos\alpha - A\sin\alpha & C\cos\alpha - B\sin\alpha \end{vmatrix} = AC - B^2 \neq 0, \tag{11.7}$$

выберем a и b таким образом, чтобы $D' = 0$ и $E' = 0$. Для этого, согласно (11.4), необходимо решить систему уравнений относительно неизвестных a и b (значение α фиксировано равенством (11.6))

$$\left.\begin{aligned} a(A\cos\alpha + B\sin\alpha) + b(B\cos\alpha + C\sin\alpha) &= \\ = -D\cos\alpha - E\sin\alpha& \\ a(B\cos\alpha - A\sin\alpha) + b(C\cos\alpha - B\sin\alpha) &= \\ = D\sin\alpha - E\cos\alpha& \end{aligned}\right\}.$$

В силу (11.7) для a и b по формулам Крамера имеем единственное решение.

Подставив найденные значения α, a, b в A', C', F' из (11.4), мы получим в штрихованной системе координат с началом в точке $O'(a; b)$ следующее уравнение:

$$px'^2 + qy'^2 = r. \tag{11.8}$$

Покажем, что начало координат $O'(a; b)$ является центром симметрии кривой (11.8), а штрихованные оси координат — осями симметрии. В самом деле, если в уравнение (11.8) вместо координаты x' подставить $-x'$, то уравнение не изменится. Не изменится оно и после замены y' на $-y'$. Это значит, точка O' (пересечение осей $O'x'$ и $O'y'$) есть центр симметрии кривой (11.8). Линию второго порядка, имеющую единственный центр симметрии, называют *центральной*, остальные носят название *нецентральных*.

Если $r \neq 0$, то центральная кривая называется *невырожденной*, при $r = 0$ — *вырожденной*.

(A.1) *Невырожденные центральные кривые*. Представив уравнение (11.8) в виде $\frac{x^2}{\frac{r}{p}} + \frac{y^2}{\frac{r}{q}} = 1$ (ради упрощения записи штрихи у переменных опустим) и обозначив $\frac{r}{p} = \varepsilon_1 a^2$, $\frac{r}{q} = \varepsilon_2 b^2$ $(a, b > 0, \varepsilon_1 = \pm 1, \varepsilon_2 = \pm 1)$, получим

$$\frac{x^2}{\varepsilon_1 a^2} + \frac{y^2}{\varepsilon_2 b^2} = 1. \tag{11.9}$$

При $\varepsilon_1 = 1$, $\varepsilon_2 = 1$ имеем *каноническое уравнение эллипса*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{11.10}$$

Когда $a = b$, эллипс представляет собой окружность. При $\varepsilon_1 = 1$, $\varepsilon_2 = -1$ имеем *каноническое уравнение гиперболы*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{11.11}$$

При $\varepsilon_1 = -1$, $\varepsilon_2 = 1$ опять имеем уравнение гиперболы, в котором по сравнению с (11.11) роль переменной x

играет y и наоборот. Этот случай принципиально не отличается от предыдущего.

При $\varepsilon_1 = -1$, $\varepsilon_2 = -1$ вещественных точек не существует (*мнимый эллипс*). Очевидно, что первый и четвертый случай соответствуют ситуации, когда $\Delta > 0$, второй и третий, когда $\Delta < 0$.

(А.2) *Вырожденные центральные кривые.* При $r = 0$ уравнение (11.8) записывается в виде

$$\varepsilon_1 \frac{x^2}{a^2} + \varepsilon_2 \frac{y^2}{b^2} = 0 \left(\varepsilon_1 a^2 = \frac{1}{p}, \ \varepsilon_2 b^2 = \frac{1}{q}\right). \qquad (11.12)$$

При $\varepsilon_1 = 1$, $\varepsilon_2 = 1$ $(\Delta > 0)$ имеем одну точку O' $(0, 0)$. При $\varepsilon_1 = 1$, $\varepsilon_2 = -1$ $(\Delta < 0)$ левая часть (11.12) распадается на два сомножителя $\left(\frac{x}{a} + \frac{y}{b}\right)\left(\frac{x}{a} - \frac{y}{b}\right) = 0$. Поэтому имеем пару пересекающихся прямых

$$\frac{x}{a} + \frac{y}{b} = 0, \ \frac{x}{b} - \frac{y}{b} = 0. \qquad (11.13)$$

При $\varepsilon_1 = -1$, $\varepsilon_2 = 1$ $(\Delta < 0)$ ситуация та же самая. Этот случай ничем принципиальным от предыдущего не отличается.

Наконец, при $\varepsilon_1 = -1$, $\varepsilon_2 = -1$ $(\Delta > 0)$ возвращаемся к первой ситуации.

(В). Пусть $\Delta = AC - B^2 = 0$. Взяв $a = 0$, $b = 0$ в (11.2), после выбора угла поворота осей α в виде (11.6) мы получим, что $B' = 0$. Из-за инвариантности Δ имеем $\Delta' = A'C' = 0$. Пусть $A' = 0$ (если $C' = 0$, то все рассуждения повторяются буквально, только роль x будет играть y), и, следовательно, $C' \neq 0$. Уравнение (11.3) принимает после умножения на $\frac{1}{C'}$ следующий вид:

$$y'^2 + 2px' + 2qy' + r = 0. \qquad (11.14)$$

После переноса начала координат в точку

$$O''\left(\frac{q^2 - r}{2p}; -q\right) \ x'' = x' + \frac{r - q^2}{2p}, \ y'' = y' + q$$

(случай $p = 0$ будет рассмотрен ниже) имеем

$$y''^2 = 2px''. \qquad (11.15)$$

Получили *каноническое уравнение не центральной невырожденной* кривой второго порядка — *параболы.*

Если в (11.14) $p=0$, то после переноса начала координат в точку O'' $(0;\ -q): x''=x',\ y''=y'+q$ имеем *каноническое уравнение вырожденной нецентральной кривой второго порядка*

$$(y'')^2=\varepsilon a^2\ \ (\varepsilon a^2=q^2-r,\ \varepsilon=\pm 1). \qquad (11.16)$$

При $\varepsilon=1$ уравнение (11.6) описывает *пару параллельных* прямых (при $a=0$ *слившихся*), при $\varepsilon=-1$ — *пару мнимых параллельных прямых.*

На этом классификация всех возможных алгебраических линий второго порядка закончена.

Перейдем к рассмотрению свойств и установлению формы *эллипса* (11.10). При этом мы будем предполагать, что $b<a$ (при $b>a$ роли переменных x и y меняются местами), и называть a — *большой полуосью* эллипса, b — *малой полуосью* эллипса.

Введем параметр $c>0$, определив его из равенства $c^2=a^2-b^2$, и рассмотрим на оси Ox две точки $F_1(-c;\ 0)$ и $F_2(c;\ 0)$ так, что начало координат делит отрезок F_1F_2 пополам. Вычислим сейчас расстояния r_1 и r_2 от точек F_1 и F_2 до произвольной точки эллипса $M(x;\ y): r_1=\sqrt{(x+c)^2+y^2},\ \ r_2=\sqrt{(x-c)^2+y^2}$. Из уравнения (11.10), подставив вместо b^2 его выражение $b^2=a^2-c^2$, получим $y^2=a^2-c^2-x^2+\dfrac{c^2}{a^2}x^2$ и внесем его под корни для r_1 и r_2

$$\left.\begin{aligned} r_1&=\sqrt{2cx+a^2+\frac{c^2}{a^2}x^2}=\sqrt{\left(a+\frac{c}{a}x\right)^2}=\\ &\qquad\qquad=\left|a+\frac{c}{a}x\right|\\ r_2&=\sqrt{-2cx+a^2+\frac{c^2}{a^2}x^2}=\sqrt{\left(a-\frac{c}{a}x\right)^2}=\\ &\qquad\qquad=\left|a-\frac{c}{a}x\right| \end{aligned}\right\}.$$

Из уравнения (11.10) следует, что $|x|\leqslant a$, и так как $c<a$, то $\left|\dfrac{c}{a}x\right|<a$. Таким образом, $a+\dfrac{c}{a}x>0$ и

$a-\frac{c}{a}x>0$. Поэтому для r_1 и r_2 окончательно имеем

$$r_1=a+ex,\ r_2=a-ex, \tag{11.17}$$

где параметр $e=\frac{c}{a}<1$ называют *эксцентриситетом* эллипса. Складывая равенства (11.17), получим, что $r_1+r_2=2a$ для *любой* точки эллипса. Это свойство позволяет определить эллипс как *геометрическое место точек плоскости, сумма расстояний от которых до двух данных точек F_1 и F_2* (называемых *фокусами* эллипса) *есть величина постоянная* (большая чем расстояние между фокусами).

Числа r_1 и r_2 из (11.17) часто называют *фокальными радиусами* эллипса.

Исследуем форму эллипса по его каноническому уравнению. Нами уже установлено, что оси координат являются осями симметрии, а начало координат — центром симметрии. Поскольку $\frac{x^2}{a^2}\leqslant 1$ и $\frac{y^2}{b^2}\leqslant 1$, то эллипс расположен в прямоугольнике $|x|\leqslant a$, $|y|\leqslant b$.

Для исследования формы эллипса в силу симметрии достаточно рассмотреть ту его часть, что лежит в первой координатной четверти: $y=\frac{b}{a}\sqrt{a^2-x^2}$. Получив график этой линии и достроив ее симметрично в остальных четвертях, получим линию эллипса (рис. 24а).

При изучении эллипса особую роль играют две прямые, перпендикулярные к оси Ox, с уравнениями $x=\pm\frac{a}{e}$ (рис. 24б) — *директрисы* эллипса. Для эллипса $e<1$, $\frac{a}{e}>a$, и, следовательно, директрисы расположены вне эллипса. Для директрис имеет место теорема, которую можно взять за новое определение эллипса.

Теорема. *Отношение расстояния от любой точки эллипса до фокуса к расстоянию ее до соответствующей директрисы есть величина постоянная, равная эксцентриситету эллипса.*

Действительно, пусть $M(x;y)$ — точка эллипса, d_1 и d_2 — расстояния до соответствующих директрис. В силу симметрии достаточно доказать теорему для одного из фокусов (например, для F_2).

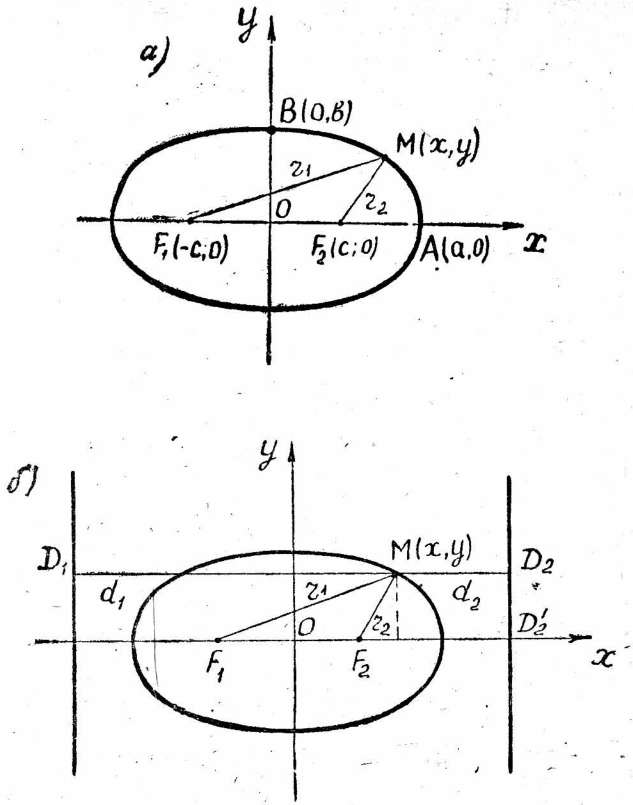


Рис. 24.

Точка M' имеет координаты $(x;\ 0)$, а $D_2' - \left(\frac{a}{e};\ 0\right)$. Поэтому $d_2 = |\overrightarrow{M'D_2'}| = \frac{a}{e} - x$. Поскольку $r_2 = a - ex$, то $\frac{r_2}{d_2} = \frac{(a-ex)\,e}{a-ex} = e$. Теорема доказана.

В заключение остановимся на параметрическом уравнении эллипса. С этой целью рассмотрим окружность $x^2 + y^2 = a^2$ и произведем точечное сжатие плоскости вдоль ординат точек

$$x = X,\ \ y = \frac{a}{b}\,Y. \tag{11.18}$$

Подставив в уравнение окружности, получим уравнение эллипса $\frac{X^2}{a^2}+\frac{Y^2}{b^2}=1$. Пусть окружность задана в параметрической форме: $x=a\cos t$, $y=a\sin t$ $(t\in[0, 2\pi))$. Из (11.18) находим $X=a\cos t$, $Y=\frac{b}{a}y=$ $=b\cos t$ $(t\in[0, 2\pi))$. Поэтому параметрическое уравнение эллипса с полуосями a и b имеют вид

$$\left.\begin{aligned} x&=a\cos t\\ y&=b\sin t\end{aligned}\right\}\ (0\leqslant t<2\pi). \tag{11.19}$$

В следующей лекции мы изучим гиперболу и параболу.

Лекция 12

ГИПЕРБОЛА, ЕЕ СВОЙСТВА И ФОРМА. ПАРАБОЛА, ЕЕ СВОЙСТВА И ФОРМА. ПОЛЯРНОЕ УРАВНЕНИЕ ЭЛЛИПСА, ГИПЕРБОЛЫ И ПАРАБОЛЫ. УСЛОВИЯ КАСАНИЯ ПРЯМОЙ И КРИВОЙ ВТОРОГО ПОРЯДКА

Перейдем к установлению геометрических свойств и формы гиперболы

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1\ (a,\ b>0). \tag{12.1}$$

Введем параметр $c>0$ с помощью равенства $c^2=a^2+b^2$, так что $c>a$, и рассмотрим на оси Ox две точки $F_1(-c;\ 0)$ и $F_2(c;\ 0)$, которые назовем *левым и правым фокусами* гиперболы. Вычислим фокальные расстояния $|\overrightarrow{F_1M}|$, $|\overrightarrow{F_2M}|$ до произвольной точки гиперболы $M(x;\ y):|\overrightarrow{F_1M}|=r_1=\sqrt{(x+c)^2+y^2}$, $|\overrightarrow{F_2M}|=r_2=$ $=\sqrt{(x-c)^2+y^2}$. Из (12.1) находим $y^2=\frac{c^2}{a^2}x^2-x^2-$ $-c^2+a^2$ и подставляем в выражения для r_1 и r_2. Имеем

$$\left.\begin{aligned} r_1 &= \sqrt{\left(a+\frac{c}{a}x\right)^2} = |a+ex| \\ r_2 &= \sqrt{\left(a-\frac{c}{a}x\right)^2} = |a-ex| \end{aligned}\right\}, \qquad (12.2)$$

где $e = \frac{c}{a}$ — параметр, называемый *эксцентриситетом гиперболы*. Заметим, что для гиперболы $e > 1$, так как $c > a$.

Исследуем различные возможности, представляемые равенствами (12.2). Из (12.1) следует $\frac{x^2}{a^2} \geqslant 1$, т. е. $|x| \geqslant a$. Поэтому в зависимости от того, лежит ли $M(x; y)$ в правой полуплоскости ($x > 0$) или в левой ($x < 0$), выражения для (12.2) различные.

Если $x > 0$, то $a + ex > 0$ и $a - ex < 0$, так как $e > 1$, $x > a$, $ex > a$. Поэтому *в правой полуплоскости* выполнено

$$r_1 = a + ex, \quad r_2 = -a + ex. \qquad (12.3)$$

Если $x < 0$, то $a + ex < 0$, так как $|ex| > a$, и $a - ex > 0$. Поэтому *в левой полуплоскости* выполнено

$$r_1 = -a - ex, \; r_2 = a - ex. \qquad (12.4)$$

Из этих рассуждений вытекает, что гипербола состоит из двух симметричных ветвей, расположенных соответственно в правой и левой полуплоскостях. Из (12.3), (12.4) для обеих ветвей выводим инвариантное равенство

$$|r_1 - r_2| = 2a. \qquad (12.5)$$

На основании (12.5) можно дать новое определение гиперболы как *ГМТ плоскости, модуль разности расстояний от которых до двух фиксированных точек F_1, F_2, называемых фокусами гиперболы, есть величина постоянная* (не равная нулю и меньшая расстояния между фокусами).

При установлении формы гиперболы отметим, что оси Ox и Oy являются *осями симметрии*, а их пересечение — *центром симметрии* (центр гиперболы). Поскольку $\frac{x^2}{a^2} \geqslant 1$ ($|x| \geqslant a$), то ветви гиперболы лежат на плоскости вне полосы $-a < x < a$. Точки пересечения ветвей с осью Ox ($y=0$, $A_1(-a; 0)$, $A_2(a; 0)$) называются *вершинами* гиперболы. Для гиперболы, заданной уравнением (12.1), ось Ox называется *действительной осью* гиперболы, а ось Oy — *мнимой осью* гиперболы $\left(\text{пересечения нет из-за равенства } -\frac{y^2}{b^2}=1\right)$.

В силу симметрий достаточно исследовать форму гиперболы в первой четверти, $y=\frac{b}{a}\sqrt{x^2-a^2}$, а затем распространить на всю плоскость. Наряду с указанной ветвью гиперболы целесообразно рассмотреть луч, исходящий из начала координат, с уравнением $Y=\frac{b}{a}x$.

Пусть M — точка гиперболы с абсциссой x, N — точка указанной прямой с той же абсциссой. Для разности ординат имеем

$$Y-y=\frac{b}{a}(x-\sqrt{x^2-a^2}).$$

При неограниченном возрастании x данная разность монотонно стремится к нулю, так как

$$\lim_{x\to\infty}(Y-y)=\frac{b}{a}\lim_{x\to\infty}(x-\sqrt{x^2-a^2})=$$

$$=\frac{b}{a}\lim_{x\to\infty}\frac{x^2-x^2+a^2}{(x+\sqrt{x^2-a^2})}=0.$$

Поэтому прямые $y=\pm\frac{b}{a}x$ являются *асимптотами* гиперболы $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$. Для построения асимптот строим прямоугольник со сторонами $x=\pm a$, $y=\pm b$. Асимптоты — прямые, содержащие диагонали прямоугольника. При построении гиперболы удобно построить сначала асимптоты, а затем ветви гиперболы (рис. 25а).

Для гиперболы вводятся две прямые, перпендикулярные к оси Ox, с уравнениями $x=\pm\frac{a}{e}$, и называемые *директрисами гиперболы*. В силу неравенства $e>1$ директрисы лежат между вершинами гиперболы.

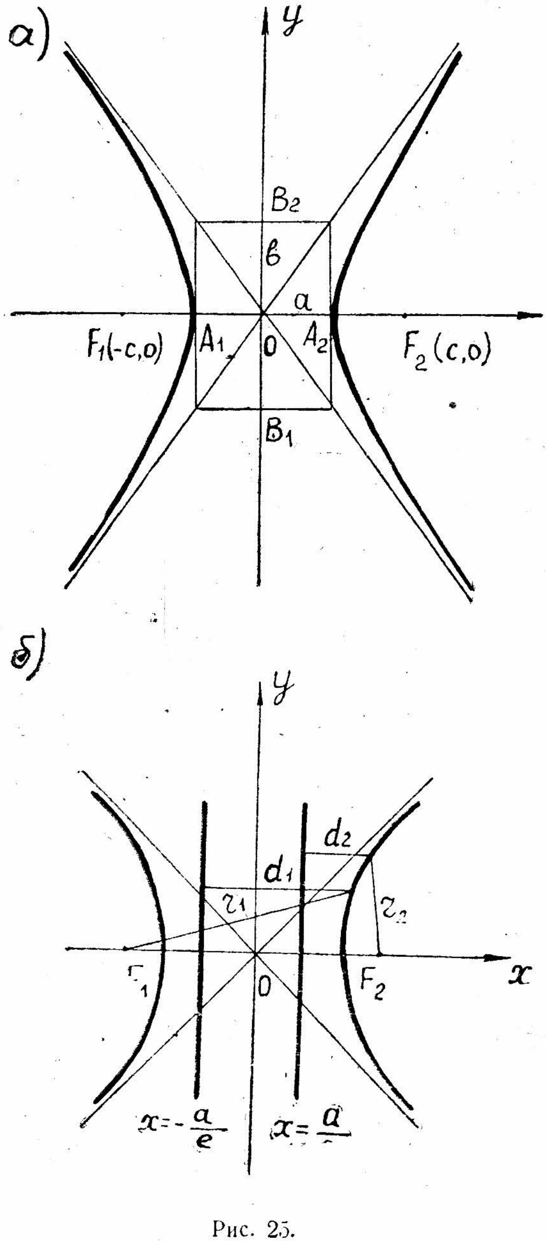


Рис. 25.

Так же, как и для эллипса, имеет место утверждение: $\frac{r_1}{d_1} = \frac{r_2}{d_2} = e$, где d_1 и d_2 — расстояния от точки гиперболы до соответствующих директрис (рис. 25б). Доказательство совершенно аналогично тому, что проведено для эллипса, и поэтому мы приводить его не будем. Приведем лишь формулировку теоремы, которая является фактически еще одним определением гиперболы.

Теорема. *Отношение расстояния от любой точки гиперболы до фокуса к расстоянию ее до соответствующей директрисы есть величина постоянная, равная эксцентриситету гиперболы* $e > 1$.

Отметим, что гипербола $\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1$ называется *сопряженной* по отношению к гиперболе (12.1). Действительной осью сопряженной гиперболы является ось Oy, асимптоты остаются прежними.

Перейдем к установлению геометрических свойств и вида формы *параболы*

$$y^2 = 2px \quad (p > 0). \tag{12.6}$$

Прежде всего отметим, что кривая лежит в правой полуплоскости ($x \geqslant 0$) и проходит через начало координат. Рассмотрим точку $F\left(\frac{p}{2};\ 0\right)$, называемую *фокусом параболы*, и прямую, перпендикулярную к оси Ox, с уравнением $x = -\frac{p}{2}$, называемую *директрисой параболы*. Пусть $M(x;\ y)$ — произвольная точка параболы. Вычислим расстояние

$$|\overrightarrow{FM}| = r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2},$$

подставив y^2 из соотношения (12.6). Получим

$$r = \sqrt{\left(x + \frac{p}{2}\right)^2} = \left|x + \frac{p}{2}\right| = x + \frac{p}{2}. \tag{12.7}$$

Расстояние d от M до директрисы $x = -\frac{p}{2}$ также равно $x + \frac{p}{2}$. Поэтому для параболы $\frac{r}{d} = 1$. Таким образом, *парабола есть ГМТ, для которых расстояние до фиксированной точки F* (фокуса параболы) *равно расстоянию до фиксированной прямой* (директрисы параболы).

Так как для эллипса и гиперболы отношение $\frac{r}{d}$ есть постоянное число, равное e — эксцентриситету, то и для параболы отношение $\frac{r}{d}=e$ называют *эксцентриситетом*, т. е. для параболы $e=1$.

Из приведенного рассуждения следует, что эллипс, гипербола и парабола обладают *общим* геометрическим свойством: *отношение расстояния r от любой точки M каждой из этих кривых до фокуса к расстоянию d от этой точки M до соответствующей директрисы есть величина постоянная, равная эксцентриситету e кривой.*

Форма параболы легко устанавливается с использованием симметрии относительно оси Ox и монотонного возрастания функции $y=\sqrt{2px}$ (рис. 26а).

Для вывода единого уравнения эллипса, гиперболы и параболы *в полярной системе координат* воспользуемся установленным единым для этих кривых свойством $\frac{r}{d}=e$.

Пусть задана какая-либо из перечисленных кривых. Поместим полюс полярной системы координат в фокус F, а полярную ось — по перпендикуляру от директрисы к фокусу (рис. 26б).

Пусть B — точка пересечения кривой с перпендикуляром к полярной оси, исходящим из фокуса F, а p — длина вектора $\overrightarrow{FB}$, называемая *фокальным параметром*. Согласно свойству $\frac{r}{d}=e$ имеем, что $\frac{p}{|\overrightarrow{CB}|}=\frac{p}{|\overrightarrow{EF}|}=e$. Поэтому $|\overrightarrow{EF}|=\frac{p}{e}$, а также

$$d=|\overrightarrow{DM}|=|\overrightarrow{EN}|=|\overrightarrow{EF}|+r\cos\varphi=\frac{p}{e}+r\cos\varphi.$$

Подставляя данное выражение в отношение $\frac{r}{d}=e$, получим

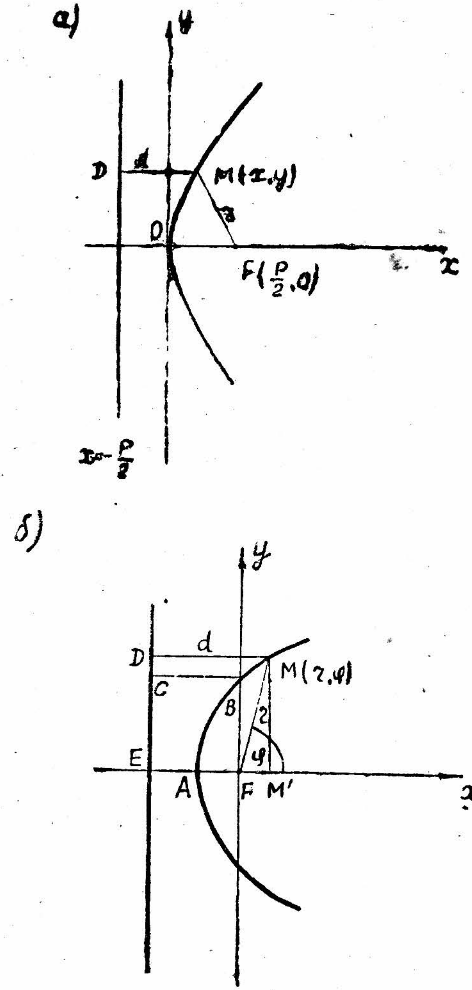


Рис. 26.

$$\frac{r}{\frac{p}{e} + r \cos \varphi} = e,$$

откуда следует полярное уравнение для рассматриваемых кривых второго порядка ($e<1$—эллипс, $e>1$—гипербола, $e=1$ — парабола):

$$r = \frac{p}{1 - e\cos\varphi}. \tag{12.8}$$

Отметим, что при $e > 1$ из-за условия $r > 0$ уравнение (12.8) описывает только одну ветвь гиперболы, внутри которой лежит фокус.

Эллипс, гиперболу и параболу называют *коническими сечениями*, потому что их можно получить путем сечения плоскостями кругового конуса. Если плоскость пересекает только одну полость конуса и параллельна одной из образующих конуса, то получим параболу. Если секущая плоскость пересекает только одну полость конуса и не параллельна ни одной образующей конуса, то коническое сечение будет эллипсом. Когда секущая плоскость перпендикулярна оси конуса, получаем окружность.

Наконец, если секущая плоскость не проходит через вершину и параллельна оси конуса, то получим гиперболу.

В заключение лекции остановимся на условиях касания прямой и кривой второго порядка, заданной своим каноническим уравнением. Рассмотрение подробно проведем для эллипса. Для гиперболы и параболы все повторяется буквально, и для них мы приведем лишь результат.

Пусть заданы прямая $Ax + By + C = 0$ и эллипс $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Считаем $B \neq 0$, так как прямые, перпендикулярные к оси Ox ($B = 0$) и касающиеся эллипса, известны ($x = \pm a$). Совместное рассмотрение системы уравнений

$$\begin{cases} y = -\frac{A}{B}x - \frac{C}{B} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \end{cases}$$

приводит к квадратному уравнению

$$\frac{x^2}{a^2} + \frac{1}{b^2}\left(\frac{A}{B}x + \frac{C}{B}\right)^2 = 1,$$

корни которого должны совпадать (прямая $Ax + By + C = 0$ касается кривой). Поэтому дискриминант

$D_1 = \frac{A^2C^2}{B^4b^4} - \left(\frac{1}{a^2} + \frac{A^2}{B^2b^2}\right)\left(\frac{C^2}{B^2b^2} - 1\right) = 0$, откуда следует *условие касания прямой* $Ax + By + C = 0$ *и эллипса* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

$$A^2a^2 + B^2b^2 = C^2. \qquad (12.9)$$

Аналогично рассуждая, получим *условие касания прямой и гиперболы* (12.1)

$$A^2a^2 - B^2b^2 = C^2, \qquad (12.10)$$

а также *прямой* $y = kx + b$ *и параболы* $y^2 = 2px$

$$(k - p)^2 = k^2b^2.$$

Пример. Написать уравнение касательной к эллипсу $\frac{x^2}{25} + \frac{y^2}{16} = 1$, которая была бы перпендикулярна к прямой $2x - 3y + 1 = 0$.

Решение. Условие перпендикулярности прямой $Ax + By + C = 0$ и прямой $2x - 3y + 1 = 0$ записывается в виде $2A - 3B = 0$ $\left(A = \frac{3}{2}B\right)$, а условие касания: $25A^2 + 16B^2 = C^2$. Решая совместно данные условия, получим $\left(25 \cdot \frac{9}{4} + 16\right)B^2 = C^2$, т. е. $C = \pm\frac{17}{2}B$. Подставляя в уравнение прямой значения параметров $A = \frac{3}{2}B$, $C = \pm\frac{17}{2}B$ и сокращая на общий множитель $\frac{1}{2}B$, получим два решения задачи: $3x + 2y \pm 17 = 0$.

Лекция 13

ПРОСТЕЙШИЕ ЗАДАЧИ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ В ПРОСТРАНСТВЕ. УРАВНЕНИЕ ПОВЕРХНОСТИ И УРАВНЕНИЕ ЛИНИИ В ПРОСТРАНСТВЕ. РАЗЛИЧНЫЕ ВИДЫ УРАВНЕНИЯ ПЛОСКОСТИ В ПРОСТРАНСТВЕ

Переходим к аналитическому рассмотрению геометрических образов в пространстве относительно прямоугольной системы координат $\{0, \vec{i}, \vec{j}, \vec{k}\}$. Предварительно остановимся на трех простейших понятиях аналитической геометрии в пространстве, лежащих в основе решений прикладных задач.

Определение. Под расстоянием $d(M_1, M_2)$ между двумя точками $M_1(x_1; y_1; z_1)$, $M_2(x_2; y_2; z_2)$ понимается модуль вектора $\overrightarrow{M_1M_2}$, т. е.

$$d(M_1, M_2) = d(M_2, M_1) = |\overrightarrow{M_1M_2}| = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}. \qquad (13.1)$$

Пусть задана упорядоченная пара точек

$$A(x_1; y_1; z_1),\ B(x_2; y_2; z_2).$$

Проведем прямую линию, проходящую через точки A и B, и выберем на прямой точку $C(x; y; z)$, не совпадающую с точкой B. Рассмотрим векторы $\overrightarrow{AC}$ и $\overrightarrow{CB}$. Они коллинеарны, и мы можем записать

$$\overrightarrow{AC} = \lambda\overrightarrow{CB}. \qquad (13.2)$$

Число λ в (13.2) называется *отношением, в котором точка C делит отрезок AB*.

В координатах (13.2) имеет вид

$$x - x_1 = \lambda(x_2 - x),\ y - y_1 = \lambda(y_2 - y),\ z - z_1 = \lambda(z_2 - z),$$

откуда находим

$$\lambda = \frac{x-x_1}{x_2-x} = \frac{y-y_1}{y_2-y} = \frac{z-z_1}{z_2-z},$$

а также

$$x = \frac{x_1+\lambda x_2}{1+\lambda},\quad y = \frac{y_1+\lambda y_2}{1+\lambda},\quad z = \frac{z_1+\lambda z_2}{1+\lambda}. \qquad (13.3)$$

В частности, при $\lambda = 1$ точка C делит отрезок пополам. Координаты середины отрезка будут иметь вид

$$x = \frac{x_1+x_2}{2},\quad y = \frac{y_1+y_2}{2},\quad z = \frac{z_1+z_2}{2}. \qquad (13.4)$$

Часто в приложениях участвует формула для площади треугольника, заданного своими вершинами $A_1(x_1; y_1; z_1)$, $A_2(x_2; y_2; z_2)$, $A_3(x_3; y_3; z_3)$. Напомним ее (см. лекцию 7)

$$S_\Delta = \frac{1}{2}\left| \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} \right|. \qquad (13.5)$$

Перейдем к выяснению геометрического смысла уравнения с тремя переменными. Целесообразно начать с примера. Пусть нам задана сфера с центром в точке $C(a; b; c)$ *и радиуса* R. Сферу можно определить как *ГМТ пространства, отстоящих от центра сферы на одном и том же расстоянии* R. Если $M(x; y; z)$ — произвольная точка сферы, то $|\overrightarrow{CM}| = R$ или $|\overrightarrow{CM}|^2 = R^2$. В координатах это равенство называется уравнением сферы и выглядит следующим образом:

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = R^2. \qquad (13.6)$$

В частности, если центр сферы находится в начале координат, уравнение будет иметь вид

$$x^2 + y^2 + z^2 = R^2. \qquad (13.7)$$

В аналитической геометрии поверхность рассматривают как множество точек пространства, обладающее определенным свойством. Поэтому, если $M(x; y; z)$ — произвольная точка поверхности, то посредством соотношения, связывающего 3 переменные x, y, z, выражают их общее свойство.

Определение. Соотношение $F(x, y, z) = 0$ называют *неявным уравнением поверхности* в пространстве ($z = f(x, y)$ — *явным уравнением*), если ему удовлетворяют координаты любой точки поверхности и не удовлетворяют координаты ни одной точки, не принадлежащей поверхности.

Заметим, что уравнение $F(x, y, z) = 0$ может иметь особенности и множество точек, координаты которых удовлетворяют ему, может быть пустым, составлять конечное или счетное множество, распадаться на ветви и т. д.

Например, множество действительных точек, удовлетворяющих уравнению $x^2 + y^2 + z^2 + 1 = 0$, есть пустое множество, уравнение $x^2 + y^2 + z^2 = 0$ определяет одну точку $O(0; 0; 0)$, уравнение $x^2 - z^2 = 0$ может быть представлено в виде $(x-z)(x+z) = 0$ и поверхность распадается на две ветви: $x - z = 0$, $x + z = 0$.

Пусть в уравнение поверхности не входит одна из переменных. Какую поверхность определяет такое уравнение? Чтобы ответить на этот вопрос, дадим определение цилиндрической поверхности.

Определение. *Цилиндрической поверхностью* (*цилиндром*) называется поверхность, полученная движением прямой, перемещающейся параллельно некоторому вектору и пересекающей во время движения фиксированную линию. Движущаяся прямая называется образующей цилиндрической поверхности (рис. 27а).

Пусть, например, уравнение не содержит z и имеет вид $F(x, y) = 0$. На плоскости xOy уравнение определяет линию (рис. 27б). Но этому уравнению удовлетворяют также координаты точек пространства, у которых первые две координаты совпадают с координатами точек L, а координата z произвольна, т. е. координаты точек всех прямых, перпендикулярных к плоскости xOy в точках кривой L. Данная поверхность получена движением прямой вдоль L параллельно вектору $\vec{k}$. Итак, *уравнение* $F(x, y) = 0$ *определяет цилиндрическую поверхность с образующей параллельной оси* Oz. Аналогично, $F(x, z) = 0$ задает цилиндрическую поверхность с образующей параллельной оси Oy, а $F(y, z) = 0$—цилиндр, образующие которого параллельны оси Ox.

Методы аналитической геометрии и линейной алгебры позволяют выяснить свойства поверхностей, левая часть уравнения которых представляет полиномы первой и второй степени (поверхности, описываемые более сложными уравнениями, изучаются методами математического анализа и дифференциальной геометрии). В связи с этим дадим определение алгебраической поверхности.

Определение. Алгебраической поверхностью m-го порядка называется поверхность, левая часть уравнения которой представляет полином m-й степени относительно переменных x, y, z, т. е. сумму слагаемых вида $Ax^p y^q z^r$, где $p + q + r \leqslant m$ $(p, q, r \geqslant 0)$.

Например, $Ax + By + Cz + D = 0$ — общее уравнение алгебраической поверхности первого порядка, а $Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2 + 2Gx + 2Hy + 2Kz + L = 0$ — общее уравнение алгебраической поверхности второго порядка.

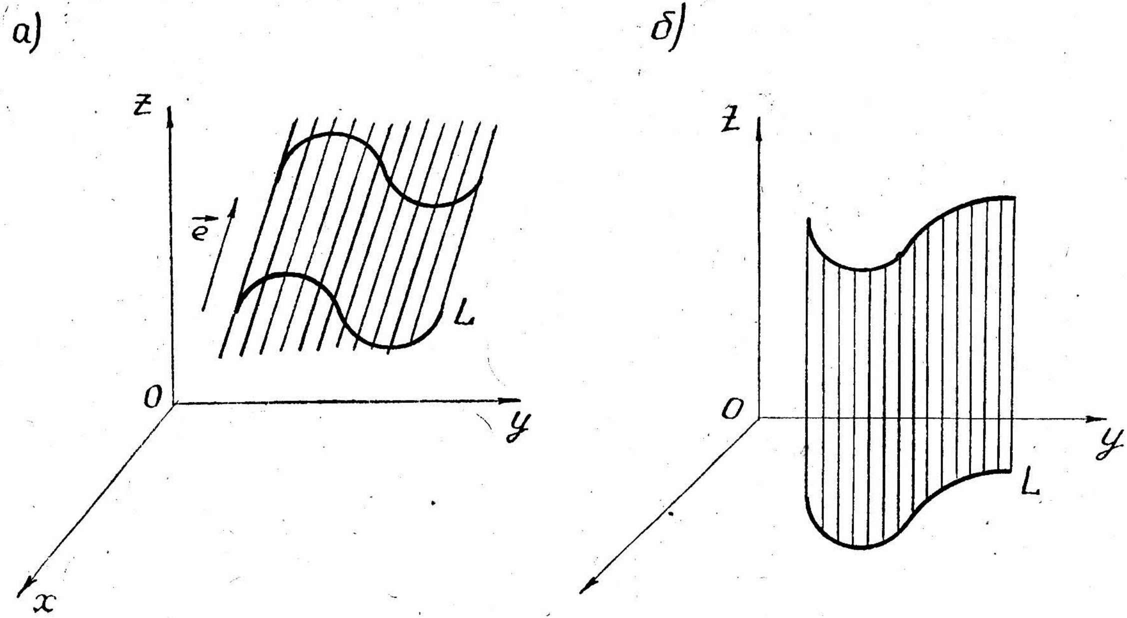


Рис. 27.

Линию в пространстве можно рассматривать как пересечение двух поверхностей

$$\left.\begin{matrix} F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0 \end{matrix}\right\} \quad (13.8)$$

$$\left(\text{или} \begin{cases} z = f_1(x, y) \\ z = f_2(x, y) \end{cases}\right),$$

т. е. *линия есть ГМТ, координаты которых удовлетворяют системе двух уравнений с тремя неизвестными.*

Линию в пространстве в классической механике рассматривают как след движущейся материальной точки, координаты которой есть функции времени t (параметра t)

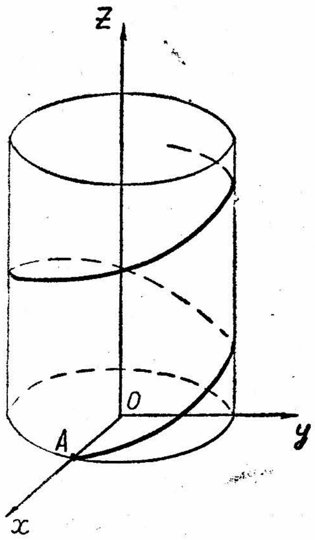


Рис. 28.

$$\left.\begin{matrix} x = x(t) \\ y = y(t) \\ z = z(t) \end{matrix}\right\} \quad (t \in T). \qquad (13.9)$$

В аналитической геометрии от временного смысла параметра отказываются, и t может иметь смысл переменного угла и др.

Уравнения (13.9) носят название параметрических уравнений линии в пространстве.

Пример. Пусть отрезок OA длины a равномерно вращается и равномерно движется вдоль оси Oz (рис. 28). Точка A описывает линию, находящуюся на поверхности цилиндра. Она носит название *винтовой линии.* Если за параметр t выбрать угол поворота отрезка, то параметрические уравнения винтовой линии имеют вид

$$\left.\begin{matrix} x = a\cos t \\ y = a\sin t \\ z = bt \end{matrix}\right\} \quad (-\infty < t < +\infty). \qquad (13.10)$$

Обратимся сейчас к исследованию алгебраической поверхности первой степени. С этой целью рассмотрим некоторую плоскость в пространстве. Ее можно зафиксировать, если задать точку $M_0(x_0; y_0; z_0)$ (*опорная точка плоскости*) и *нормальный* вектор $\vec{N}(A, B, C) \neq \vec{0}$ (вектор, ортогональный к плоскости).

Пусть точка $M(x; y; z)$ — текущая точка плоскости (рис. 29а). Вектор $\vec{r} - \vec{r}_0$ ортогонален к $\vec{N}$. Поэтому уравнение

$$(\vec{N}(\vec{r} - \vec{r}_0)) = 0 \qquad (13.11)$$

есть *векторное уравнение плоскости с опорной точкой и нормальным вектором* $\vec{N}$. В координатной записи имеем

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \qquad (13.12)$$

Если раскрыть скобки и обозначить $D = -Ax_0 - By_0 - Cz_0$, то приходим к алгебраическому уравнению первого порядка

$$Ax + By + Cz + D = 0. \qquad (13.13)$$

Обратно, если в прямоугольной системе координат имеем уравнение (13.13), то это будет уравнение некоторой плоскости. В самом деле, уравнение (13.13) есть линейное уравнение с тремя неизвестными. Оно всегда имеет решение. Пусть (x_0, y_0, z_0) — некоторое решение (13.13), т. е. $Ax_0 + By_0 + Cz_0 + D \equiv 0$. Вычитая это тождество из (13.13), получим $A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$. На левую часть этого соотношения можно смотреть как на скалярное произведение вектора $\vec{N}(A, B, C)$ и вектора $\vec{r} - \vec{r}_0$, где $\vec{r} = \vec{i}x + \vec{j}y + \vec{k}z$ — переменный радиус-вектор, $\vec{r}_0 = \vec{i}x_0 + \vec{j}y_0 + \vec{k}z_0$ — фиксированный вектор, задающий точку $M_0(x_0; y_0; z_0)$, т. е. $(\vec{N}(\vec{r} - \vec{r}_0)) = 0$. Очевидно, что данному уравнению удовлетворяют координаты *любой точки* плоскости, проходящей через M_0 и имеющей нормальный вектор $\vec{N}$, и не удовлетворяют координаты ни одной точки P с радиус-вектором $\vec{r}'$, не лежащей на пло-

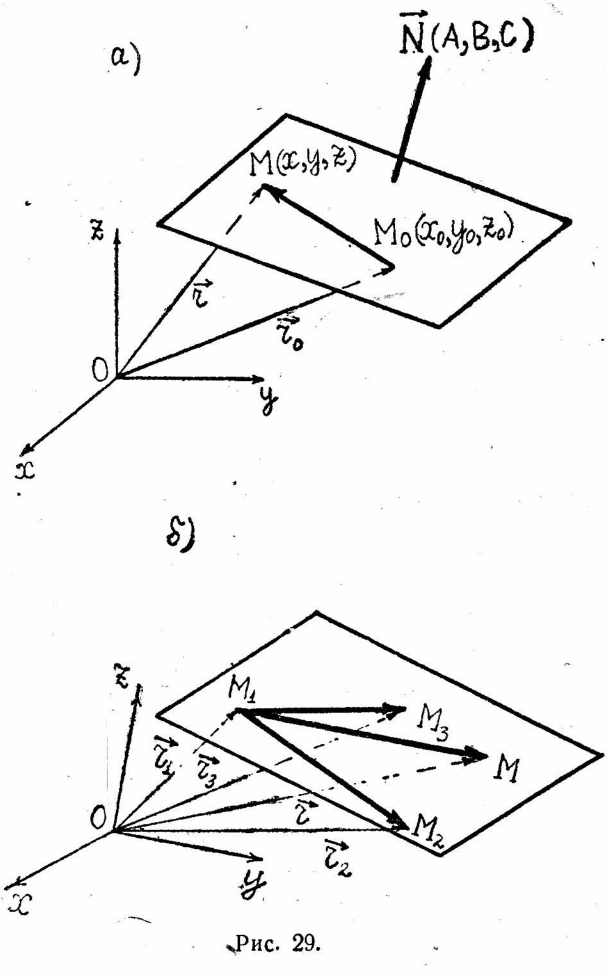


Рис. 29.

скости, так как $\vec{r}' - \vec{r}_0$ не ортогонален $\vec{N}$ и, следовательно, $(\vec{N}(\vec{r}' - \vec{r}_0)) \neq 0$.

Определение. Уравнение (13.13) называется *общим уравнением* плоскости в пространстве.

В зависимости от того, какой из параметров A, B, C, D равен нулю или отличен от нуля, расположение плоскости относительно системы координат будет обладать определенным свойством (параллельность координатной оси, координатной плоскости и т. д.). Например, $Ax+By+Cz=0$ есть уравнение плоскости, проходящей через начало координат. Исследование других частных случаев предоставим читателю и остановимся лишь на видах уравнения плоскости, имеющих наибольшие приложения.

Уравнение в отрезках. Пусть $A\neq 0$, $B\neq 0$, $C\neq 0$, $D\neq 0$. Тогда общее уравнение (13.13) можно переписать в виде $\dfrac{x}{\left(-\dfrac{D}{A}\right)}+\dfrac{y}{\left(-\dfrac{D}{B}\right)}+\dfrac{z}{\left(-\dfrac{D}{C}\right)}=1$. Обозначив

$a=-\dfrac{D}{A}$, $b=-\dfrac{D}{B}$, $c=-\dfrac{D}{C}$, имеем уравнение плоскости в отрезках

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1. \tag{13.14}$$

Плоскость отсекает от осей координат отрезки длиной $|a|$, $|b|$, $|c|$, поскольку уравнению (13.14) удовлетворяют координаты точек $P_1(a;0;0)$, $P_2(0;b;0)$, $P_3(0;0;c)$.

Уравнение плоскости, проходящей через 3 точки. Заданы три точки $M_i(x_i;y_i;z_i)$, не лежащие на одной прямой. Пусть $M(x;y;z)$ — текущая точка плоскости. Векторы $\overrightarrow{M_1M}$, $\overrightarrow{M_2M_1}$, $\overrightarrow{M_1M_3}$ компланарны, и, следовательно, смешанное произведение этих векторов равно нулю (рис. 29б)

$$((\vec{r}-\vec{r}_1)(\vec{r}_2-\vec{r}_1)(\vec{r}_3-\vec{r}_1))=0. \tag{13.15}$$

Имеем *векторное уравнение плоскости, проходящей через 3 точки*. В координатной записи (13.15) имеет вид

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix}=0. \tag{13.16}$$

Нормированное (нормальное) уравнение плоскости. Пусть плоскость не проходит через начало координат. Опустим перпендикуляр из начала координат на плоскость и рассмотрим вектор $\overrightarrow{OP}=\vec{p}$. Пусть $\vec{n}=\vec{i}\cos\alpha+$

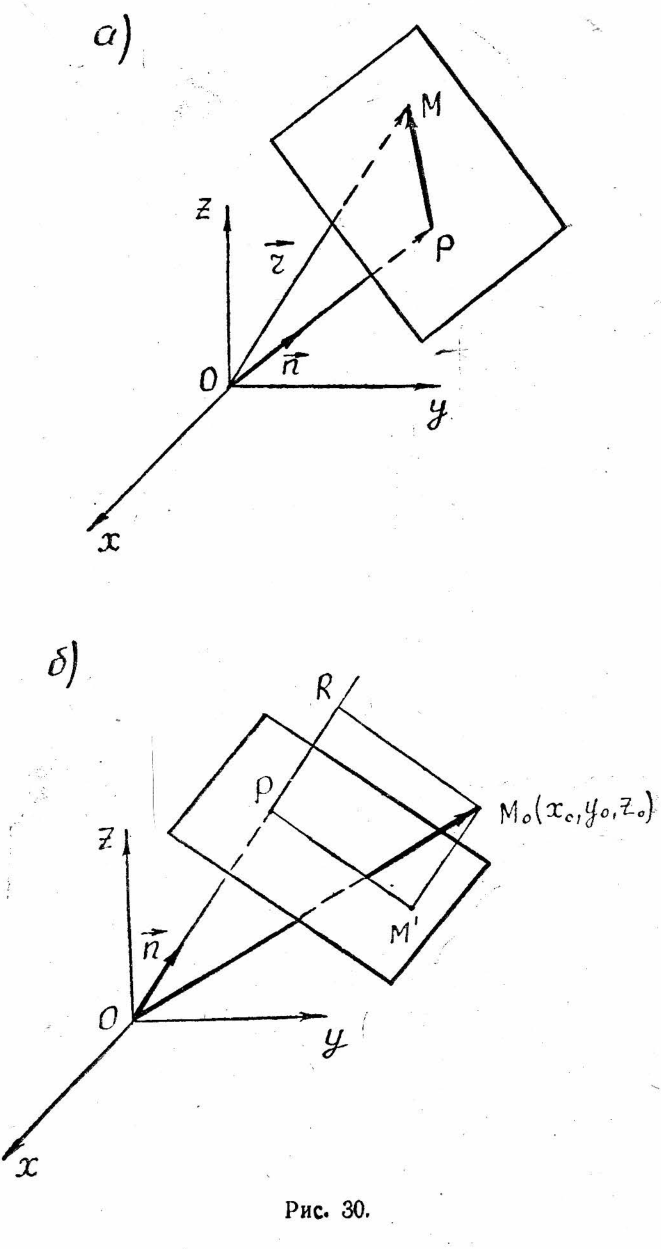


Рис. 30.

$+\vec{j}\cos\beta+\vec{k}\cos\gamma$ — орт вектора $\vec{p}$, $M(x;y;z)$ — текущая точка плоскости (рис. 30а). Видим, что $(\vec{r}-p\vec{n})$ ортогонален вектору $\vec{n}$ и, следовательно, $((\vec{r}-p\vec{n})\vec{n})=0$. Раскрывая скобки, имеем

$$(\vec{r}\,\vec{n})-p=0 \quad (p>0). \tag{13.17}$$

Получили *нормированное* (*нормальное*) *уравнение плоскости в векторном виде.* В координатной записи (13.17) имеет вид

$$x\cos\alpha+y\cos\beta+z\cos\gamma-p=0 \quad (p>0). \tag{13.18}$$

Если уравнение плоскости задано общим уравнением $Ax+By+Cz+D=0$, то его можно пронормировать, учитывая, что

$$\vec{n}=\frac{\vec{N}}{\varepsilon|\vec{N}|}=\frac{A}{\varepsilon\sqrt{A^2+B^2+C^2}}\vec{i}+\frac{B}{\varepsilon\sqrt{A^2+B^2+C^2}}\vec{j}+\frac{C}{\varepsilon\sqrt{A^2+B^2+C^2}}\vec{k},$$

где $\varepsilon=+1$, когда $D<0$, и $\varepsilon=-1$, когда $D>0$.

Итак,

$$\frac{Ax}{\varepsilon\sqrt{A^2+B^2+C^2}}+\frac{By}{\varepsilon\sqrt{A^2+B^2+C^2}}+\frac{Cz}{\varepsilon\sqrt{A^2+B^2+C^2}}+\frac{D}{\varepsilon\sqrt{A^2+B^2+C^2}}=0 \tag{13.19}$$

— *нормированное уравнение плоскости, полученное из общего уравнения.*

Определение. Отклонением δ точки $M_0(x_0;y_0;z_0)$ от плоскости называется число, равное $+d$, где d — расстояние от точки M_0 до плоскости, если M_0 и начало координат находятся по разные стороны плоскости, и равное $-d$, если M_0 и начало координат находятся по одну сторону плоскости.

Поскольку $\delta=\text{Пр}_{\vec{n}}\overrightarrow{OM_0}-p$ (рис. 30б), то

$$\delta=(\overrightarrow{OM_0}\cdot\vec{n})-p=(\vec{r}_0\,\vec{n})-p.$$

В координатной записи имеем

$$\delta = x_0 \cos\alpha + y_0 \cos\beta + z_0 \cos\gamma - p. \qquad (13.20)$$

Из (13.20) следует, что для расстояния от точки до плоскости имеет место формула

$$d = |x_0 \cos\alpha + y_0 \cos\beta + z_0 \cos\gamma - p|. \qquad (13.21)$$

Пример. Найти расстояние от точки M (1, —1, 2) до плоскости $3x - y + z + 1 = 0$.

Решение. Нормируем уравнение плоскости:

$$-\frac{3x}{\sqrt{11}} + \frac{y}{\sqrt{11}} - \frac{z}{\sqrt{11}} - \frac{1}{\sqrt{11}} = 0.$$

По (13.21) получаем

$$d = \left| -\frac{3}{\sqrt{11}} - \frac{1}{\sqrt{11}} - \frac{2}{\sqrt{11}} - \frac{1}{\sqrt{11}} \right| = \frac{7}{\sqrt{11}}.$$

Наконец, остановимся на условиях параллельности и перпендикулярности двух плоскостей $A_1x + B_1y + C_1z + D_1 = 0$, $A_2x + B_2y + C_2z + D_2 = 0$. Поскольку параллельность плоскостей означает коллинеарность нормальных векторов плоскостей, то в силу необходимых и достаточных условий коллинеарности двух векторов мы имеем следующее утверждение.

Теорема 1. *Необходимым и достаточным условием параллельности двух плоскостей является условие*

$$[\vec{N}_1 \vec{N}_2] = 0 \Leftrightarrow \text{ранг} \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{bmatrix} = 1. \qquad (13.22)$$

Перпендикулярность двух плоскостей эквивалентна ортогональности их нормальных векторов. Поэтому имеем еще одно утверждение.

Теорема 2. *Необходимым и достаточным условием перпендикулярности двух плоскостей является условие*

$$(\vec{N}_1 \vec{N}_2) = 0 \Leftrightarrow A_1A_2 + B_1B_2 + C_1C_2 = 0. \qquad (13.23)$$

Пример 1. Найти уравнение плоскости, проходящей через точку M_0 (1, 2, —2) и параллельно плоскости $3x + 2y - z + 1 = 0$.

Решение. Нормальный вектор $\vec{N}_1$ имеет координаты (3, 2, —1). За нормальный вектор искомой плоскости можно взять (ради простоты решения) вектор, совпадающий с $\vec{N}_1$. Поэтому имеем $3(x-1)+ + 2(y-2)-(z+2)=0$ или $3x+2y-z-9=0$.

Пример 2. Найти уравнение множества плоскостей, проходящих через точку M_0 (1; 2; —2) и перпендикулярных к плоскости $3x+2y-z+1=0$.

Решение. Из (13.23) следует $3A_2+2B_2-C_2=0$. Находим $C_2=3A_2+2B_2$ и подставим в уравнение $A_2(x-1)+B_2(y-2)+C_2(z+2)=0$. Окончательно имеем: $A_2x+B_2y+(3A_2+2B_2)z=5A_2-2B_2$—параметрическое семейство плоскостей.

Лекция 14

ПРЯМАЯ В ПРОСТРАНСТВЕ. ВЗАИМНОЕ РАСПОЛОЖЕНИЕ ПРЯМОЙ И ПЛОСКОСТИ В ПРОСТРАНСТВЕ. ИССЛЕДОВАНИЕ ФОРМЫ ПОВЕРХНОСТЕЙ ВТОРОГО ПОРЯДКА ПО КАНОНИЧЕСКИМ УРАВНЕНИЯМ

Прямая в пространстве полностью определена, если на ней задана точка $M_0(x_0; y_0; z_0)$ (*опорная точка прямой*) и *направляющий вектор* $\vec{p}(l, m, n)$. Пусть $M(x; y; z)$ — *текущая точка* прямой. Тогда векторы $\overrightarrow{M_0M}$ и $\vec{p}$ коллинеарны, и мы можем записать $\vec{r}-\vec{r}_0=t\vec{p}$ $(-\infty<t<+\infty)$ (рис. 31а), т. е. для радиус-вектора $\vec{r}$ текущей точки имеем

$$\vec{r}=\vec{r}_0+t\vec{p} \quad (-\infty<t<+\infty). \tag{14.1}$$

Уравнение (14.1) есть *векторное параметрическое уравнение прямой в пространстве*. В координатной записи (14.1) выглядит следующим образом:

$$\left.\begin{aligned} x&=x_0+lt\\ y&=y_0+mt\\ z&=z_0+nt \end{aligned}\right\} \quad (-\infty<t<+\infty). \tag{14.2}$$

Исключим t из уравнений (14.2), разрешив их сначала относительно t, а затем приравняв правые части равенств, имеем:

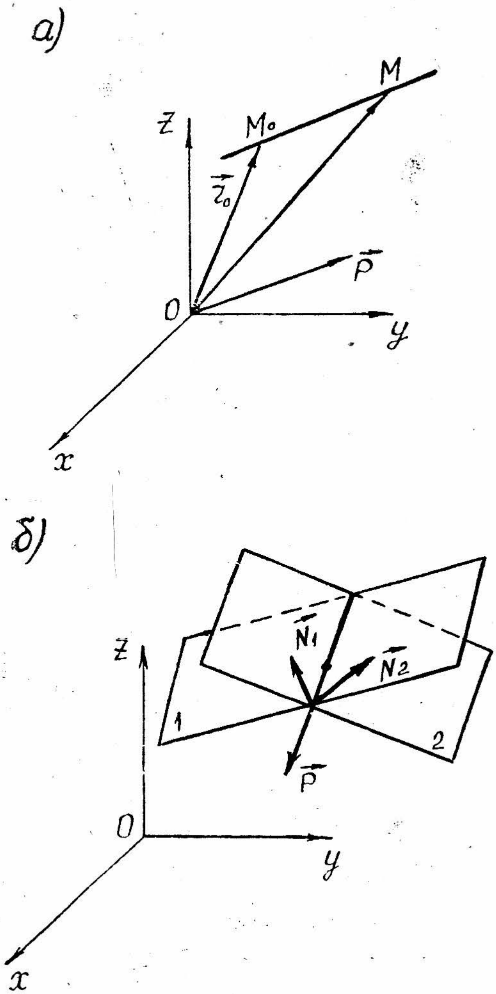


Рис. 31.

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}. \qquad (14.3)$$

Уравнения (14.3) носят название *канонических уравнений прямой в пространстве*. Заметим, что если какая-либо координата направляющего вектора равна нулю, то равен нулю и числитель дроби. Например, выражение $\frac{z-z_0}{0}$ понимают условно, оно означает равенство $z-z_0=0$.

На основании (14.3) легко написать уравнение прямой, проходящей через 2 точки $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$. Первую точку, например, мы принимаем за опорную, а за направляющий вектор $\vec{p}\neq 0$ возьмем вектор $\overrightarrow{M_1M_2}=(x_2-x_1,\ y_2-y_1,\ z_2-z_1)$. Поэтому *уравнение прямой через 2 точки* примет вид

$$\frac{x-x_1}{x_2-x_1}=\frac{y-y_1}{y_2-y_1}=\frac{z-z_1}{z_2-z_1}. \tag{14.4}$$

Прямая в пространстве может быть задана также как пересечение двух плоскостей

$$\begin{cases} A_1x+B_1y+C_1z+D_1=0 \\ A_2x+B_2y+C_2z+D_2=0. \end{cases} \tag{14.5}$$

Плоскости пересекаются, если их нормальные векторы не коллинеарные векторы (иначе плоскости параллельны). Поэтому за направляющий вектор $\vec{p}$ прямой может быть взят вектор, полученный от перемножения $\vec{N}_1$ и $\vec{N}_2$ векторным способом (рис. 31б). Итак, $\vec{p}=[\vec{N}_1\vec{N}_2]$. За опорную точку искомой прямой можно взять любую тройку чисел, удовлетворяющую системе (14.5). Если $\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}\neq 0$, то положив в (14.5) $z=0$ и решая оставшуюся систему двух уравнений с двумя неизвестными, находим x_0 и y_0. Точка с координатами $(x_0; y_0; 0)$ и будет опорной точкой прямой.

Тем самым от уравнения прямой в виде (14.5) мы переходим к каноническим (или параметрическим) уравнениям той же прямой.

Пример. Прямая задана в виде $\begin{cases} x-y+2z+1=0 \\ x+y+3z+2=0. \end{cases}$ Написать ее канонические уравнения.

Решение. Положив $z=0$, решаем систему

$$\begin{cases} x-y=-1 \\ x+y=-2. \end{cases}$$

Находим, что $x=-\frac{3}{2}$, $y=-\frac{1}{2}$. Итак, опорная точка имеет координаты $\left(-\frac{3}{2};\ -\frac{1}{2};\ 0\right)$. Найдем направляющий вектор $\vec{p}$

$$\vec{p}=[\vec{N}_1\vec{N}_2]=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ 1 & -1 & 2\\ 1 & 1 & 3\end{vmatrix}=-5\vec{i}-\vec{j}+2\vec{k}.$$

Канонические уравнения прямой имеют вид

$$\frac{x+\frac{3}{2}}{-5}=\frac{y+\frac{1}{2}}{-1}=\frac{z}{2}.$$

В механике часто уравнение прямой записывают в виде момента вектора $\vec{p}$. Поскольку вектор $\vec{r}-\vec{r}_0$ коллинеарен с $\vec{p}$, то их векторное произведение $[(\vec{r}-\vec{r}_0)\,\vec{p}]=0$. Обозначив $[\vec{r}_0\vec{p}]=\vec{a}$, получим векторное уравнение прямой в виде „момента"

$$[\vec{r}\,\vec{p}]=\vec{a}. \tag{14.6}$$

Используя сведения из векторной алгебры, легко вывести значение косинуса угла между прямыми, подразумевая под ним угол (в пределах от нуля до π) между направляющими векторами прямых $\varphi=(\widehat{\vec{p}_1,\vec{p}_2})$. Поэтому

$$\cos\varphi=\frac{(\vec{p}_1\vec{p}_2)}{|\vec{p}_1|\,|\vec{p}_2|}=\frac{l_1l_2+m_1m_2+n_1n_2}{\sqrt{l_1^2+m_1^2+n_1^2}\sqrt{l_2^2+m_2^2+n_2^2}}. \tag{14.7}$$

Теорема 1. *Необходимым и достаточным условием параллельности прямых в пространстве является условие*

$$[\vec{p}_1\vec{p}_2]=0\ \left(\frac{l_1}{l_2}=\frac{m_1}{m_2}=\frac{n_1}{n_2}\right).$$

Доказательство теоремы совпадает дословно с доказательством теоремы о необходимых и достаточных условиях коллинеарности двух векторов. Аналогично имеет место также следующая теорема.

Теорема 2. *Необходимым и достаточным условием перпендикулярности двух прямых является условие* $(\vec{p}_1 \vec{p}_2) = 0$ $(l_1 l_2 + m_1 m_2 + n_1 n_2 = 0)$.

Как узнать, являются ли две данные прямые скрещивающимися? Если прямые не являются скрещивающимися (т. е. пересекаются или параллельны), то они лежат в одной плоскости и векторы $\overrightarrow{M_1M_2}$, $\vec{p}_1$, $\vec{p}_2$, где M_1 и M_2 — опорные точки прямых, компланарны, т. е.

$$\Delta = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{vmatrix} = 0. \qquad (14.8)$$

Если $\Delta \neq 0$, то прямые являются скрещивающимися.

Пример. Лежат ли прямые $\frac{x+1}{2} = \frac{y-1}{0} = \frac{z+2}{-1}$, $\frac{x-2}{-1} = \frac{y+1}{2} = \frac{z-5}{0}$ в одной плоскости?

Решение. Проверим равенство (14.8). Вычисляем определитель

$$\Delta = \begin{vmatrix} 3 & -2 & 7 \\ 2 & 0 & -1 \\ -1 & 2 & 0 \end{vmatrix} = 3 \cdot 2 + 2 \cdot 1 + 7 \cdot 4 = 36 \neq 0.$$

Прямые являются скрещивающимися и в одной плоскости не лежат.

Перейдем к рассмотрению ряда задач на взаимное расположение прямой и плоскости. Используя доказанные в векторной алгебре теоремы, приведем 3 утверждения.

Теорема 3. *Необходимым и достаточным условием перпендикулярности прямой и плоскости является условие* $[\vec{N}\vec{p}] = \vec{0}$ $\left(\text{или } \frac{A}{l} = \frac{B}{m} = \frac{C}{n}\right)$.

Теорема 4. *Необходимым и достаточным условием параллельности прямой и плоскости является условие* $(\vec{N}\vec{p}) = 0$ (или $Al + Bm + Cn = 0$).

Теорема 5. *Необходимым и достаточным условием принадлежности прямой плоскости является одновременное выполнение условий:* $Al + Bm + Cn = 0$, $Ax_0 + By_0 + Cz_0 + D = 0$.

В самом деле, первое условие выполнено в силу теоремы 4, а второе условие означает, что опорная точка прямой принадлежит и плоскости и, следовательно, ее координаты удовлетворяют уравнению плоскости.

Пример. Лежит ли прямая $\frac{x-1}{-1}=\frac{y+1}{0}=\frac{z-5}{2}$ на плоскости $2x+3y+z-4=0$?

Решение. Проверяем первое условие: $2\cdot(-1)+{}$ $+3\cdot 0+1\cdot 2\equiv 0$. Проверяем второе условие: $2\cdot 1+{}$ $+3(-1)+5-4\equiv 0$. Прямая принадлежит плоскости.

Задача 1. Написать уравнение плоскости, проходящей через прямую $\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}$ и точку $M_1(x_1; y_1; z_1)$, не лежащую на прямой.

Пусть $M(x; y; z)$ — текущая точка плоскости. В таком случае векторы $\overrightarrow{M_0M}$ (M_0 — опорная точка прямой), $\overrightarrow{M_0M_1}$ и $\vec{p}$ компланарны. Это значит

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ x_1-x_0 & y_1-y_0 & z_1-z_0 \\ l & m & n \end{vmatrix}=0. \qquad (14.9)$$

Это и есть уравнение искомой плоскости.

Задача 2. Написать уравнение плоскости, проходящей через прямую $\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}$ и параллельной прямой $\frac{x-x_1}{l_1}=\frac{y-y_1}{m_1}=\frac{z-z_1}{n_1}$ $([\vec{p}\,\vec{p}_1]\neq\vec{0})$.

За нормальный вектор плоскости можно взять вектор, равный векторному произведению $\vec{p}$ и $\vec{p}_1$: $\vec{N}=[\vec{p}\,\vec{p}_1]$. За опорную точку плоскости можно взять опорную точку прямой, через которую проходит плоскость. Поэтому искомое уравнение имеет вид $([\vec{p}\,\vec{p}_1](\vec{r}-\vec{r}_0))=0$. В координатах имеем

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ l & m & n \\ l_1 & m_1 & n_1 \end{vmatrix}=0. \qquad (14.10)$$

Задача 3. Написать уравнение плоскости, которая перпендикулярна к плоскости $A_1x + B_1y + C_1z + D_1 = 0$ и проходит через прямую $\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$.

За нормальный вектор искомой плоскости возьмем $\vec{N} = [\vec{p}\vec{N}_1]$, а за опорную точку искомой плоскости — опорную точку прямой, так как плоскость проходит через нее. Имеем, что $([\vec{p}\vec{N}_1](\vec{r} - \vec{r}_0)) = 0$ или в координатах

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ l & m & n \\ A_1 & B_1 & C_1 \end{vmatrix} = 0. \qquad (14.11)$$

Задача 4. Найти координаты точки пересечения прямой $\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$ и плоскости $Ax + Bx + Cz + D = 0$.

Решение. Запишем параметрические уравнения прямой: $x = x_0 + lt$, $y = y_0 + mt$, $z = z_0 + nt$. Подставив в уравнение плоскости эти значения, получим $A(x_0 + lt) + B(y_0 + mt) + C(z_0 + nt) + D = 0$, откуда при $Al + Bm + Cn \neq 0$ найдем значение параметра $t = t'$, которое удовлетворяет данному уравнению. Тогда $x' = x_0 + lt'$, $y' = y_0 + mt'$, $z' = z_0 + nt'$ есть координаты точки пересечения. Если $Al + Bm + Cn = 0$, то прямая параллельна плоскости.

Задача 5. Найти расстояние от точки до прямой в пространстве.

Решение. Пусть прямая задана векторным параметрическим уравнением $\vec{r} = \vec{r}_0 + t\vec{p}$ с опорной точкой $M_0(x_0; y_0; z_0)$ и точка $M_1(x_1; y_1; z_1)$ вне прямой. Опустим из точки M_1 перпендикуляр M_1M_1' (рис. 32а). Совместим начало направляющего вектора с опорной точкой M_0 и построим параллелограмм на векторах $\vec{p}$ и $\overrightarrow{M_0M_1}$. Тогда расстояние d от точки M_1 до прямой мы можем найти, разделив площадь параллелограмма на длину основания $|\vec{p}|$.

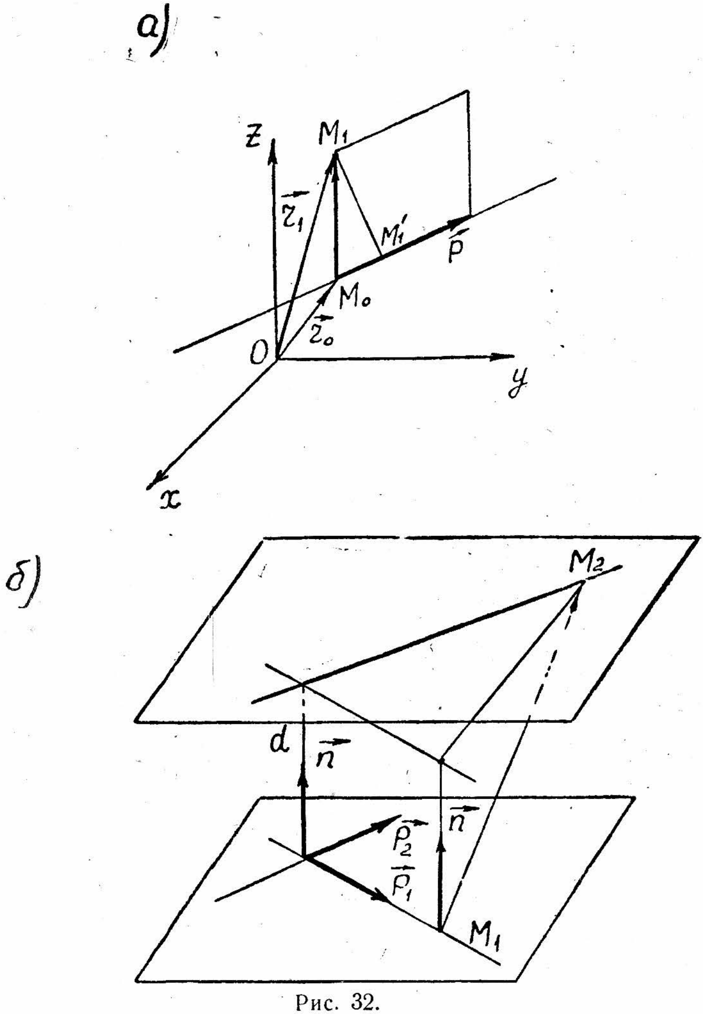


Рис. 32.

Площадь параллелограмма, как известно, совпадает с модулем векторного произведения $\overrightarrow{M_0M_1} \times \vec{p}$. Поэтому

$$d = \frac{|[(\vec{r}_1 - \vec{r}_0)\,\vec{p}]|}{|\vec{p}|}. \tag{14.12}$$

Пример. Требуется найти расстояние от точки M_1 $(1, -1, 2)$ до прямой $\frac{x+1}{0}=\frac{y-1}{1}=\frac{z-3}{2}$.

Имеем

1) $$[(\vec{r}_1-\vec{r}_0)\vec{p}]=\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2 & -2 & -1 \\ 0 & 1 & 2 \end{vmatrix}=-3\vec{i}-4\vec{j}+2\vec{k};$$

2) $|[(\vec{r}_1-r_0)\vec{p}]|=\sqrt{9+16+4}=\sqrt{29};$

3) $|\vec{p}|=\sqrt{0^2+1^2+2^2}=\sqrt{5};$

4) $d=\sqrt{\frac{29}{5}}.$

Задача 6. Найти расстояние между прямыми в пространстве. Если прямые пересекаются или параллельны (условие (14.8)), то в первом случае расстояние равно нулю, а во втором—расстояние между прямыми равно расстоянию от опорной точки первой прямой до второй прямой и вычисляется по формуле (14.12).

Пусть прямые

$$\frac{x-x_1}{l_1}=\frac{y-y_1}{m_1}=\frac{z-z_1}{n_1} \text{ и } \frac{x-x_1}{l_2}=\frac{y-y_2}{m_2}=\frac{z-z_2}{n_2} -$$

скрещивающиеся. Для них определитель (14.8) отличен от нуля. Кратчайшее расстояние между ними равно расстоянию между параллельными плоскостями, проходящими соответственно через первую и вторую прямые. Очевидно, за нормальный вектор $\vec{N}$ к плоскостям может быть выбран вектор $[\vec{p}_1\vec{p}_2]$. Опорные точки $M_1(x_1; y_1; z_1)$, $M_2(x_2; y_2; z_2)$ прямых могут быть взяты за опорные точки плоскостей. Уравнения плоскостей будут иметь вид

$$([\vec{p}_1\vec{p}_2](\vec{r}-r_1))=0, \quad ([\vec{p}_1\vec{p}_2](\vec{r}-r_2))=0.$$

Искомое расстояние будет равно модулю проекции вектора $\overrightarrow{M_1M_2}$ на $\vec{n}=\frac{\vec{N}}{|\vec{N}|}$ (рис. 32б). Окончательно имеем

$$d = |\text{Пр}_{\vec{n}}\, \overrightarrow{M_1M_2}| = |((\vec{r}_2 - \vec{r}_1)\,\vec{n})| = \frac{|((\vec{r}_2 - \vec{r}_1)\,\vec{p}_1\,\vec{p}_2)|}{|[\vec{p}_1\,\vec{p}_2]|}. \tag{14.13}$$

Заметим, что в числителе дроби (14.13) стоит объем параллелепипеда, построенного на векторах $\vec{p}_1\,\vec{p}_2$, $\overrightarrow{M_1M_2}$, а в знаменателе — площадь параллелограмма, лежащего в основании параллелепипеда. Поэтому кратчайшее расстояние совпадает с длиной перпендикуляра, опущенного из вершины параллелепипеда $\{\vec{p}_1, \vec{p}_2, \overrightarrow{M_1M_2}\}$ на основание $\{\vec{p}_1, \vec{p}_2\}$.

Материал по аналитической геометрии в пространстве мы закончим исследованием формы поверхностей второго порядка методом параллельных сечений по их каноническим уравнениям. Изложение этого материала вызвано потребностями математического анализа. Строгая классификация поверхностей второго порядка будет дана ниже с использованием методов линейной алгебры.

1. *Эллипсоид*:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a > 0,\ b > 0,\ c > 0). \tag{14.14}$$

Поскольку из (14.14) следует, что $|x| \leqslant a$, $|y| \leqslant b$, $|z| \leqslant c$, то эллипсоид заключен в прямоугольном параллелепипеде со сторонами $2a$, $2b$, $2c$. Координатные плоскости являются плоскостями симметрии. Рассечем поверхность плоскостями $z = h$ $(|h| \leqslant c)$. Получим кривые с уравнениями $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{h^2}{c^2}$, $z = h$. Это эллипсы с полуосями $a\sqrt{1 - \frac{h^2}{c^2}}$ и $b\sqrt{1 - \frac{h^2}{c^2}}$. При $h = 0$ эллипс имеет наибольшие полуоси. Когда $|h|$ растет, то полуоси эллипсов уменьшаются и при $|h| = c$ эллипсы вырождаются в точки $A_1(0; 0; c)$ и $A_2(0; 0; -c)$ (рис. 33а). Аналогичная картина возникает, если рассекать эллипсоид плоскостями, перпендикулярными к оси Ox и Oy.

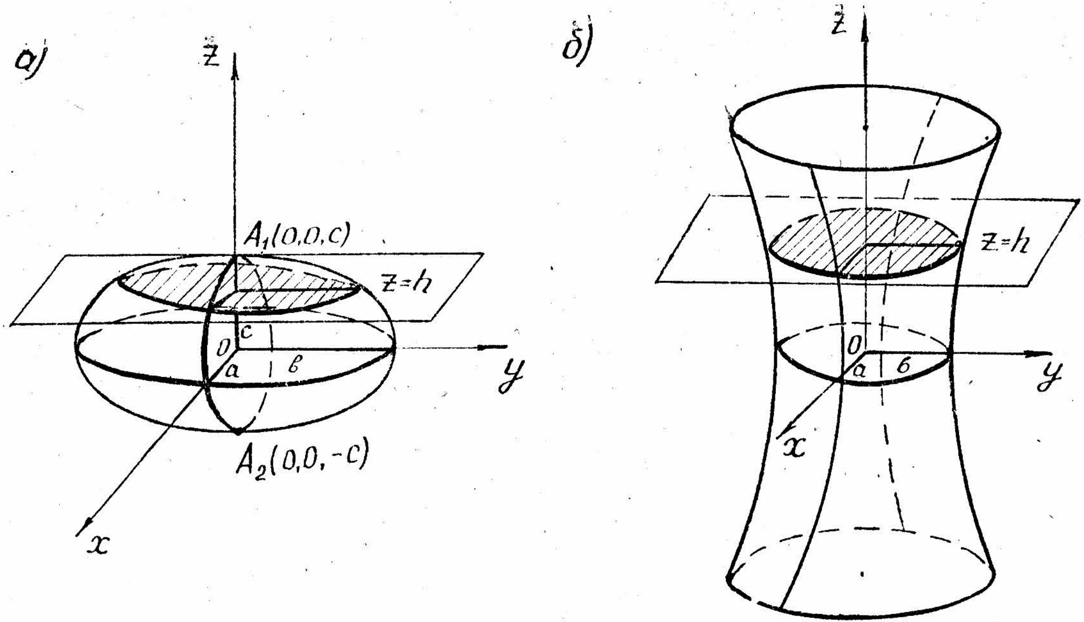


Рис. 33.

2. *Однополостной гиперболоид*: $\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$.

Координатные плоскости являются плоскостями симметрии (при замене x на $-x$, y на $-y$, z на $-z$ уравнение не меняется). Рассекаем плоскостями, параллельными плоскости $xOy: z=h$. Получим линии сечения с уравнениями $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1+\frac{h^2}{c^2}$, $z=h$ $(-\infty<$ $<h<+\infty)$. Это эллипсы, полуоси которых

$$a\sqrt{1+\frac{h^2}{c^2}},\quad b\sqrt{1+\frac{h^2}{c^2}}$$

растут вместе с ростом $|h|$. Наименьший эллипс имеем при $h=0$ (плоскость xOy). Он носит название *горлового* эллипса (рис. 33б). Сечения плоскостями $x=0$, $y=0$ дают нам гиперболы $\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$, $\frac{x^2}{a^2}-\frac{z^2}{c^2}=1$. Точки этих гипербол являются вершинами эллипсов, получаемых при сечении поверхности плоскостями $z=h$.

3. *Двуполостной гиперболоид*:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1.$$

Координатные плоскости являются плоскостями симметрии. Сечения $z=h$ дают линии с уравнениями $\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{h^2}{c^2}-1$, $z=h$. При $|h|<c$ имеем мнимые эллипсы. Следовательно, поверхность расположена выше плоскости $z=c$ и ниже плоскости $z=-c$. При $|h|\geqslant c$ получаем вещественные эллипсы, полуоси которых $a\sqrt{\frac{h^2}{c^2}-1}$, $b\sqrt{\frac{h^2}{c^2}-1}$ растут вместе с ростом $|h|$. При сечении плоскостями $x=0$ и $y=0$ имеем сопряженные гиперболы $\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1$, $\frac{x^2}{a^2}-\frac{z^2}{c^2}=-1$, в точках которых лежат вершины эллипсов, получаемых сечениями $z=h$ $(|h|\geqslant c)$ (рис. 34а).

4. *Эллиптический параболоид*: $z=\frac{x^2}{a^2}+\frac{y^2}{b^2}$.

Плоскости yOz и xOz являются плоскостями симметрии. Поверхность проходит через начало коорди-

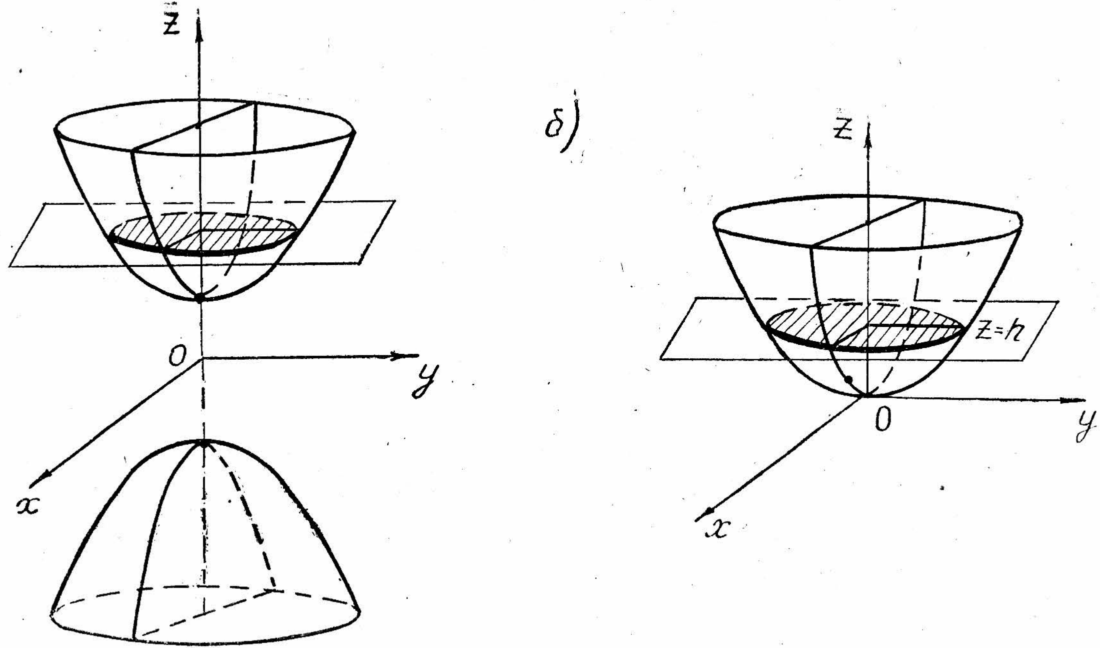


Рис. 34.

нат и расположена над плоскостью xOy, так как $z \geqslant 0$. Сечения $z = h$ дают эллипсы $\frac{x^2}{a^2} + \frac{y^2}{b^2} = h$, полуоси которых $a\sqrt{h}$, $b\sqrt{h}$ растут с ростом h (рис. 34б). При сечении плоскостями $x = 0$ и $y = 0$ имеем параболы $z = \frac{y^2}{b^2}$, $z = \frac{x^2}{a^2}$, точки которых являются вершинами указанных выше эллипсов.

5. *Гиперболический параболоид*: $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$.

Плоскости yOz и xOz являются плоскостями симметрии. Рассечем поверхность плоскостью $y = 0$. В плоскости xOz получим параболу $z = \frac{x^2}{a^2}$. Сечение плоскостью $x = 0$ дает нам в плоскости yOz параболу $z = -\frac{y^2}{b^2}$, ветви которой направлены вниз. Сечение координатной плоскостью $z = 0$ есть пара пересекающихся прямых $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ $\left(y = \pm \frac{b}{a} x\right)$. Сечение плоскостями $z = h$ дает гиперболы с уравнениями $\frac{x^2}{ha^2} - \frac{y^2}{hb^2} = 1$, причем при $h > 0$ ветви расположены вдоль оси Ox, а при $h < 0$ ветви расположены вдоль оси Oy. Построив соответствующие сечения, мы получим вид гиперболического параболоида (рис. 35а).

6. *Эллиптический конус*: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$.

Поверхность симметрична относительно всех координатных плоскостей. Сечения плоскостями $z = h$ дают нам эллипсы $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{h^2}{c^2}$ с полуосями $\frac{|h|\,a}{c}$, $\frac{|h|\,b}{c}$. Полуоси эллипсов растут с ростом $|h|$. При $h = 0$ имеем точку (начало координат). Сечения плоскостями $x = 0$ и $y = 0$ дают соответственно пары прямых $z = \pm \frac{c}{b} y$ и $z = \pm \frac{c}{a} x$. Вид эллиптического конуса дан на рис. 35б.

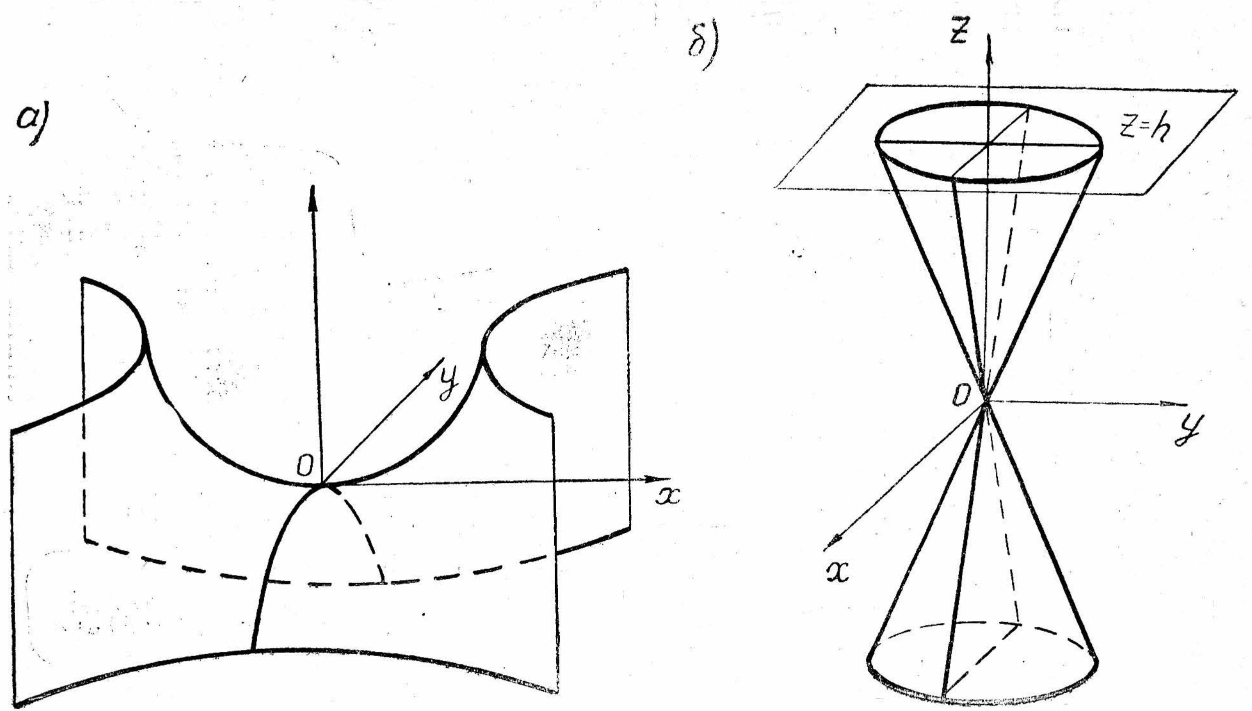


Рис. 35.

7. *Цилиндры второго порядка.* Для примера приведем лишь цилиндры с образующими, параллельными оси Oz:

(а) *эллиптический цилиндр*: $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$;

(б) *гиперболический цилиндр*: $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$;

(в) *параболический цилиндр*: $y^2=2px$.

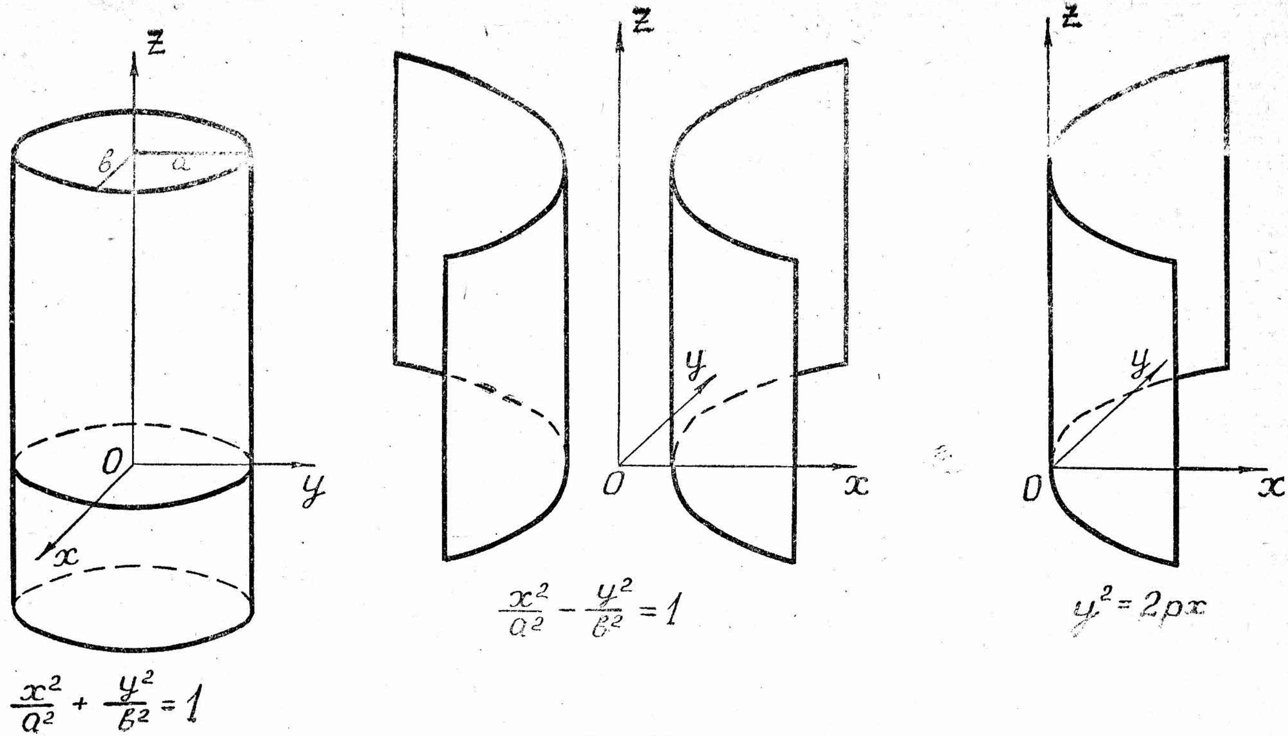
Поскольку кривые второго порядка изучены нами ранее и в лекции 13 подробно рассмотрен вопрос о цилиндрических поверхностях с образующими, параллельными оси Oz, то вид цилиндров устанавливается легко (рис. 36). Аналогично обстоит дело, когда отсутствует какая-либо другая переменная.

Если нужно получить уравнение общей цилиндрической поверхности с образующими, параллельными вектору $\vec{p}\,(l, m, n)$ и проходящими через точки линии L: $\begin{cases} F_1(x, y, z)=0 \\ F_2(x, y, z)=0 \end{cases}$, то поступаем следующим образом. Напишем уравнение образующей $\frac{X-x}{l}=\frac{Y-y}{m}=\frac{Z-z}{n}$, где x, y, z — координаты переменной точки линии L. Затем из четырех уравнений

$$\begin{cases} F_1(x, y, z)=0 \\ F_2(x, y, z)=0 \\ \frac{X-x}{l}=\frac{Y-y}{m}=\frac{Z-z}{n} \end{cases}$$

исключим переменные x, y, z. В результате получим уравнение искомой цилиндрической поверхности $\Phi(X, Y, Z)=0$.

На этом материал по аналитической геометрии исчерпан, и мы переходим к курсу линейной алгебры.



z
b
a
O
y
x
$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$
z
y
O
x
$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$
z
y
O
x
$y^2=2px$

Рис. 36.

Лекция 15

ОПРЕДЕЛЕНИЕ ЛИНЕЙНОГО ПРОСТРАНСТВА. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ ВЕКТОРОВ. БАЗИС И РАЗМЕРНОСТЬ

Пусть задано некоторое *множество* L, элементы которого будем называть *векторами* (независимо от природы элементов множества). В каждом конкретном случае они приобретают определенный смысл. Например, направленные отрезки, если речь идет о векторах евклидова пространства, матрицы-столбцы (матрицы-строки), если речь идет о множестве решений однородной системы линейных уравнений, или ψ-функции в квантовой механике и т. д.

Наряду с множеством векторов будем рассматривать числовое поле $\mathbf{K}$, под которым подразумевается поле комплексных чисел $\mathbf{C}$ либо поле вещественных чисел $\mathbf{R}$, о чем в случае необходимости будем оговаривать особо. Элементы L будем обозначать в дальнейшем латинскими малыми буквами, а элементы множества $\mathbf{K}$ — греческими малыми буквами.

Определение. Пара $\mathcal{L}=(L, \mathbf{K})$ называется *линейным пространством**, если (A) задан закон, по которому любой паре векторов $x, y \in L$ сопоставлен вектор, называемый их *суммой* и обозначаемый символом $x+y$, причем для любых $x, y, z \in L$ выполнено: (A_1) $x+y=y+x$; (A_2) $(x+y)+z=x+(y+z)$; (A_3) для любого $x \in L$ существует *нуль-вектор* 0, что $x+0=x$; (A_4) для любого $x \in L$ существует *противоположный* вектор x', что $x+x'=0$; (B) задан закон, по которому для любого $x \in L$ и любого числа $\lambda \in \mathbf{K}$ сопоставлен вектор λx, называемый *произведением* числа λ на вектор x, причем выполнено: (B_1) $1 \cdot x=x$; (B_2) $\lambda(\mu x)=(\lambda\mu) x$; (B_3) $(\lambda+\mu) x=\lambda x+\mu x$; (B_4) $\lambda(x+y)=\lambda x+\lambda y$.

Когда речь идет о конкретных линейных пространствах, то операции сложения векторов и умножения вектора на число приобретают свой конкретный смысл. Например, если в качестве L рассмотреть множество всех матриц-строк вида $a=[\alpha_1, \ldots, \alpha_n]$, $b=[\beta_1, \ldots, \beta_n]$,

---

* Наряду с термином „линейное пространство" употребителен и термин „векторное пространство".

то $a+b=[\alpha_1+\beta_1, \ldots, \alpha_n+\beta_n]$, $\lambda a=[\lambda\alpha_1, \ldots, \lambda\alpha_n]$. Легко убедиться в том, что все аксиомы (A_i), (B_i) выполнены, если за нуль-вектор взять $0=[0, 0, \ldots, 0]$, а за противоположный $a'=[-\alpha_1, -\alpha_2, \ldots, -\alpha_n]$.

На основании определения линейного пространства следуют 3 утверждения.

Теорема 1. *Нуль-вектор единственен.*

В самом деле, если существуют два нуль-вектора 0_1 и 0_2, то по (A_3) $0_1+0_2=0_1$, $0_2+0_1=0_2$. Но $0_1+0_2=0_2+0_1$ (по (A_1)) и, следовательно, $0_1=0_2$.

Теорема 2. *Если любой вектор x умножить на нуль-число*, то получим нуль-вектор.

Действительно, рассмотрим $x+0\cdot x=(1+0)x$ (по (B_3)). Поэтому $x+0\cdot x=1\cdot x=x$ (по (B_1)). Сравнивая равенства $x+0=x$ и $x+0\cdot x=x$, выводим равенство $0\cdot x=0$.

Теорема 3. *От умножения вектора x на -1 получим противоположный вектор a'.*

Доказательство: $x+(-1)x=(1+(-1))x=$ $=0\cdot x=0$. Сравнивая равенства $x+x'=0$ и $x+(-1)x=0$, выводим равенство $x'=-1\cdot x$.

Операции сложения векторов и умножения на число позволяют ввести в рассмотрение *линейные комбинации* векторов

$$a=\lambda_1 x_1+\lambda_2 x_2+\ldots+\lambda_k x_k. \tag{15.1}$$

Определение. Система векторов $\{a_1, \ldots, a_s\}$ называется *линейно-зависимой*, если существуют числа $\lambda_1, \ldots, \lambda_s$, среди которых хотя бы одно отлично от нуля, что выполнено

$$\sum_{k=1}^{s}\lambda_k a_k=\lambda_1 a_1+\lambda_2 a_2+\ldots+\lambda_s a_s=0. \tag{15.2}$$

Если соотношение вида (15.2) выполнено тогда и только тогда, когда все $\lambda_k=0$, то система векторов называется *линейно-независимой*.

Пример. В пространстве матриц-строк векторы $a_k=[0, \ldots \overset{k}{\widehat{1}} \ldots, 0]$ $(k=\overline{1, n})$ образуют линейно-независимую систему векторов, так как, составив линейную комбинацию векторов a_k с произвольными (априори) числами λ_k и приравнивая нулю, в подробной записи получаем

$$\left.\begin{array}{l}\lambda_1\cdot 1+\lambda_2\cdot 0+\ldots+\lambda_n\cdot 0=0\\ \lambda_1\cdot 0+\lambda_2\cdot 1+\ldots+\lambda_n\cdot 0=0\\ \cdots\cdots\cdots\cdots\cdots\\ \lambda_1\cdot 0+\lambda_2\cdot 0+\ldots+\lambda_n\cdot 1=0\end{array}\right\},$$

откуда следует равенство нулю всех коэффициентов λ_k $(k=\overline{1, n})$.

Докажем ряд утверждений, касающихся линейной зависимости системы векторов.

Теорема 4. *Если какая-либо подсистема системы векторов линейно-зависима, то и вся система линейно-зависима.*

Действительно, если подсистема $\{a_{j_1}, a_{j_2}, \ldots, a^{j_p}\}$, где $j_1, j_2, \ldots, j_p$ — фиксированные индексы из набора $1, 2, \ldots, s$, линейно-зависимая, то существуют числа $\lambda_{j_1}, \ldots, \lambda_{j_p}$, среди которых, например, $\lambda_{j_1}\neq 0$, что выполнено $\lambda_{j_1}a_{j_1}+\ldots+\lambda_{j_p}a_{j_p}=0$. Очевидно, выполнено также $0\cdot a_1+\ldots+0\cdot a_{j_1-1}+\lambda_{j_1}a_{j_1}+\ldots+\lambda_{j_p}a_{j_p}+0\cdot a_{j_p+1}+\ldots+0\cdot a_s=0$, где $\lambda_{j_1}\neq 0$. Система векторов

$$\{a_1, \ldots, a_{j_1}, \ldots, a_{j_p}, \ldots, a_s\}$$

линейно-зависима.

Теорема 5. *Для того, чтобы система* $\{a_1, \ldots, a_s\}$ *была линейно-зависимой, необходимо и достаточно, чтобы хотя бы один из векторов был линейной комбинацией других.*

В самом деле, если система векторов линейно-зависима, то выполнено условие (15.2), где, например, $\lambda_1\neq 0$. Поэтому из (15.2) следует, что

$$a_1=\left(-\frac{\lambda_2}{\lambda_1}\right)a_2+\ldots+\left(-\frac{\lambda_s}{\lambda_1}\right)a_s.$$

Необходимость доказана.

Пусть, обратно, какой-либо вектор из системы $\{a_1, \ldots, a_s\}$ есть линейная комбинация остальных. Например, первый вектор $a_1=\sigma_2a_2+\ldots+\sigma_sa_s$. Это равенство эквивалентно следующему равенству $(-1)a_1+\sigma_2a_2+\ldots+\sigma_sa_s=0$, где первый коэффициент отличен от нуля.

Определение. Система векторов $\{e_1, e_2, \ldots, e_n\}$ называется *базисом* в пространстве $\mathcal{L}$, если она линейно-независима и любой вектор x пространства представим в виде

$$x = \sum_{k=1}^{n} \xi^k e_k = \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n. \qquad (15.3)$$

Числа $\xi^1, \ldots, \xi^n$ называются *координатами* (компонентами) вектора x в данном базисе.

В матричном обозначении это выглядит в задании матрицы-столбца $[x]$:

$$[x] = \begin{bmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \xi^n \end{bmatrix}. \qquad (15.4)$$

Теорема 6. *Координаты вектора в заданном базисе определяются однозначно.*

Действительно, если предположить, что это не так, т. е. одновременно $x = \sum_{k=1}^{n} \xi^k e_k$ и $x = \sum_{k=1}^{n} \tilde{\xi}^k e_k$, то из равенства $\sum_{k=1}^{n} \xi^k e_k = \sum_{k=1}^{n} \tilde{\xi}^k e_k$ $\left(\text{или} \sum_{k=1}^{n} (\xi^k - \tilde{\xi}^k)\, e_k = 0\right)$ и линейной независимости векторов $e_1, \ldots, e_n$ следует $\xi^k = \tilde{\xi}^k$ $(k = \overline{1, n})$.

Очевидным является утверждение, что линейной комбинации векторов x_1 и x_2 соответствует линейная комбинация с теми же коэффициентами матриц-столбцов из координат этих векторов, так как

$$\alpha x_1 + \beta x_2 = \alpha \sum_{k=1}^{n} \xi_1^k e_k + \beta \sum_{k=1}^{n} \xi_2^k e_k = \sum_{k=1}^{n} (\alpha \xi_1^k + \beta \xi_2^k)\, e_k.$$

Определение. Если в $\mathcal{L}$ существует n линейно-независимых векторов, а любая система из $(n+1)$-го вектора линейно-зависимая, то натуральное число n называется *размерностью* пространства ($\dim \mathcal{L} = n$). Само пространство $\mathcal{L}$ называется *конечномерным* и обозначается символом $\mathcal{L}_n$. Если же в $\mathcal{L}$ можно указать систему из произвольного конечного числа линейно-независимых векторов, то говорят, что пространство $\mathcal{L}$ *бесконечномерно.*

Понятия базиса пространства и размерности пространства тесно связаны между собой.

Теорема 7. *В пространстве $\mathcal{L}_n$ любая совокупность из n линейно-независимых векторов есть базис пространства.*

Доказательство. Пусть $\{a_1, \ldots, a_n\}$ — некоторая линейно-независимая система векторов в пространстве $\mathcal{L}_n$. Согласно определению размерности пространства система векторов $\{x, a_1, \ldots, a_n\}$ при *любом* векторе x есть линейно-зависимая система и, следовательно, выполнено условие

$$\alpha x + \beta_1 a_1 + \ldots + \beta_n a_n = 0. \tag{15.5}$$

Число α необходимо отлично от нуля, так как при $\alpha = 0$ в силу линейной независимости векторов $a_1, \ldots, a_n$ последовало бы $\beta_1 = 0, \ldots, \beta_n = 0$, что привело бы к противоречию о линейной зависимости векторов $\{x, a_1, \ldots, a_n\}$. Поэтому из (15.5) для любого вектора x следует

$$x = \sum_{k=1}^{n} \left(-\frac{\beta_k}{\alpha}\right) a_k.$$

Это означает, что система линейно-независимых векторов $\{a_1, \ldots, a_n\}$ есть базис пространства $\mathcal{L}_n$.

Теорема 8. *Если в пространстве $\mathcal{L}$ существует базис из n векторов, то* $\dim \mathcal{L} = n$.

Доказательство. Пусть $\{e_1, \ldots, e_n\}$ — базис в $\mathcal{L}$. Рассмотрим *произвольную* систему из $(n+1)$-го вектора $\{x_1, \ldots, x_n, x_{n+1}\}$, где $x_i = \sum\limits_{k=1}^{n} \xi_i^k e_k$ $(i = \overline{1, n+1})$.

Построим матрицу из $(n+1)$ столбцов-координат векторов x_i

$$\begin{bmatrix} \xi_1^1 \ldots \xi_n^1 & \xi_{n+1}^1 \\ \xi_1^2 \ldots \xi_n^2 & \xi_{n+1}^2 \\ \vdots \quad\ \vdots & \vdots \\ \xi_1^n \quad\ \xi_n^n & \xi_{n+1}^n \end{bmatrix}.$$

Ранг этой матрицы $r \leqslant n$, так как число строк матрицы равно n. Поскольку столбцов $n+1$, то в силу теоремы о базисном миноре по крайней мере один из столбцов является линейной комбинацией базисных (следовательно, и остальных) столбцов рассматриваемой матрицы. Это означает, что по крайней мере один из векторов есть линейная комбинация других. На основании теоремы 5, доказанной выше, система из векторов $\{x_1, \ldots, x_n, x_{n+1}\}$ линейно-зависима. В силу произвольности рассматриваемой системы векторов теорема доказана.

Пример 1. Задано множество положительных вещественных чисел, которые будем называть векторами. Под „суммой" двух векторов понимаем обычное произведение чисел ab. Под „произведением" вектора a на число $\lambda \in \mathbf{R}$ понимаем возведение в степень λ числа a. Будет ли данное множество с введенными операциями линейным пространством?

Решение. Проверяем аксиомы группы (A). Поскольку $ab = ba$, $(ab)c = a(bc)$, то (A_1) и (A_2) выполнены. В качестве „нуль-вектора" нужно взять 1, так как $a \cdot 1 = a$. В качестве противоположного вектора необходимо взять $\frac{1}{a}$, так как $a \cdot \frac{1}{a} = 1$. После этого аксиомы (A_3) и (A_4) также будут выполнены. Перейдем к группе (B). Так как $a^1 = a$, $(a^\mu)^\lambda = a^{\mu\lambda}$, $a^{\lambda+\mu} = a^\lambda \cdot a^\mu$, $(ab)^\lambda = a^\lambda b^\lambda$, то все аксиомы группы (B) выполнены. Искомое множество с введенными операциями есть линейное пространство. За базис в введенном линейном пространстве может быть взято любое фиксированное число $p \neq 1$, так как любое положительное число может быть представлено в виде: $a = p^\lambda$, где $\lambda = \frac{\ln a}{\ln p}$. Пространство является одномерным (теорема 8).

Пример 2. Множество решений однородной системы линейных уравнений $\sum\limits_{k=1}^{n} \alpha_k^i \xi^k = 0$ $(i = \overline{1, m})$ с введенными в лекции 1 операциями сложения столбцов и умножения на число есть линейное пространство размерности $n - r$, где r — ранг матрицы системы уравнений.

В самом деле, на основании того, что линейная комбинация решений есть снова решение системы уравнений, все восемь аксиом для матриц-столбцов являющихся решениями однородной системы, выполнены. В силу теоремы, что общее решение (любой вектор множества решений) есть линейная комбинация нормальной фундаментальной системы решений, которые являются линейно-независимыми, заключаем: *пространство решений однородной системы есть линейное $(n-r)$-мерное пространство.*

Лекция 16

ПОДПРОСТРАНСТВА И ЛИНЕЙНЫЕ ОБОЛОЧКИ. ТЕОРЕМА О ПОПОЛНЕНИИ БАЗИСА. ПЕРЕСЕЧЕНИЕ И СУММА ПОДПРОСТРАНСТВ. ПРОСТРАНСТВО РЕШЕНИЙ ОДНОРОДНОЙ СИСТЕМЫ УРАВНЕНИЙ КАК ПОДПРОСТРАНСТВО ПРОСТРАНСТВА МАТРИЦ-СТОЛБЦОВ

Определение. Множество векторов $U \subset \mathscr{L}$, для которых при любых $x, y \in U$ и любых $\alpha, \beta \in \mathbf{K}$ выполнено $(\alpha x + \beta y) \in U$, называется *подпространством* линейного пространства $\mathscr{L}$. Само пространство $\mathscr{L}$ и множество, состоящее только из нуль-вектора, являются *тривиальными* подпространствами пространства $\mathscr{L}$. Примером нетривиального подпространства является множество всех векторов, коллинеарных данному вектору $\vec{e}$ в трехмерном евклидовом пространстве, так как $\vec{x} = \xi \vec{e}$, $\vec{y} = \eta \vec{e}$ и, следовательно, $\alpha \vec{x} + \beta \vec{y} = \alpha \xi \vec{e} + \beta \eta \vec{e} = (\alpha \xi + \beta \eta) \vec{e}$ — снова коллинеарный вектор при любых $\alpha, \beta \in \mathbf{R}$. Вектор $\vec{e}$ в данном примере играет роль базиса в рассматриваемом подпространстве, размерность которого, согласно теореме 8 из предыдущей лекции, равна единице.

Другим примером подпространства $\mathscr{L}$ является *линейная оболочка*, натянутая на систему векторов.

Определение. Линейной оболочкой, натянутой на систему векторов $\{a_1, \ldots, a_s\}$ (s — любое фиксированное натуральное число), называется множество векторов вида $l = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s$, где $\alpha_1, \ldots, \alpha_s$ — числовые параметры, пробегающие независимо друг от друга все числовое поле $\mathbf{K}$.

Пусть $l_1 = \sum_{k=1}^{s} \alpha_k a_k$, $l_2 = \sum_{k=1}^{s} \beta_k a_k$ — произвольные векторы оболочки. Видим, что $\alpha l_1 + \beta l_2 = \sum_{k=1}^{s} \alpha\alpha_k a_k + \sum_{k=1}^{s} \beta\beta_k a_k = \sum_{k=1}^{s} (\alpha\alpha_k + \beta\beta_k) a_k$ — снова вектор линейной оболочки при любых α и β из $\mathbf{K}$. Таким образом, *каждая линейная оболочка есть линейное подпространство.* Обратно, если U — линейное подпространство размерности $m < n$ и $\{h_1, \ldots, h_m\}$ — базис в подпространстве U, то любой вектор $x \in U$ представим в виде $x = \xi^1 h_1 + \ldots + \xi^m h_m$. Таким образом, *каждое подпространство есть линейная оболочка, натянутая на систему своих базисных векторов.*

В определении линейной оболочки векторы $a_1, \ldots, a_s$ не обязательно линейно-независимы (s может быть и больше n — размерности пространства $\mathscr{L}$). Поэтому необходимо решить вопрос о базисе и размерности линейной оболочки. С этой целью из системы $\{a_1, \ldots, a_s\}$ выделим *максимальное число* линейно-независимых векторов и обозначим их $e_1, \ldots, e_r$. За $e_1, \ldots, e_r$ можно взять r базисных столбцов матрицы, которую можно построить из матриц-столбцов координат векторов $a_1, \ldots, a_s$. Векторы $e_1, \ldots, e_r$ и являются базисными векторами оболочки, так как любой вектор оболочки есть линейная комбинация векторов $a_1, \ldots, a_s$, которые, в свою очередь, линейным образом выражаются через линейно-независимые векторы $e_1, \ldots, e_r$.

Итак, *размерность линейной оболочки равна максимальному числу линейно-независимых векторов, содержащихся среди векторов* $a_1, \ldots, a_s$ (равна рангу матрицы из s координатных столбцов векторов $a_1, \ldots, a_s$).

Пусть задано подпространство U и базис на нем $\{f_1, \ldots, f_m\}$.

Теорема 1. (О пополнении базиса подпространства до базиса всего пространства). *Векторы базиса подпространства всегда можно пополнить в $\mathscr{L}_n$ век-*

торами $e_1, e_2, \ldots, e_{n-m}$ так, *что система* $\{f_1, \ldots, f_m, e_1, \ldots, e_{n-m}\}$ *образует базис пространства* $\mathscr{L}_n$.

Действительно, в $\mathscr{L}_n$ существуют векторы, которые не являются линейной комбинацией векторов $f_1, \ldots f_m$, так как в противном случае $f_1, \ldots, f_m$ образовывали бы базис $\mathscr{L}_n$ и тогда $m = n$. Обозначим через e_1 любой вектор, не принадлежащий U. Ясно, что система векторов $\{f_1, \ldots, f_m, e_1\}$ есть линейно-независимая система векторов, потому что, составив линейную комбинацию их с произвольными коэффициентами и приравняв ее нулю: $\lambda_1 f_1 + \ldots + \lambda_m f_m + \mu_1 e_1 = 0$, получим, что $\mu_1 = 0$, иначе вектор e_1 можно было бы выразить как линейную комбинацию векторов $f_1, \ldots, f_m$. Но если $\mu_1 = 0$, то и все $\lambda_i = 0$ $(i = \overline{1, m})$ в силу линейной независимости векторов базиса подпространства. Итак, $\{f_1, \ldots, f_m, e_1\}$ — линейно-независимая система векторов. Если любой вектор $x \in \mathscr{L}_n$ есть линейная комбинация данных векторов, то они составляют базис $\mathscr{L}_n$, и, следовательно, $n = m + 1$. Наше построение на этом будет закончено.

Если же $m + 1 < n$, то, рассуждая подобным же образом, найдем вектор e_2 такой, что $\{f_1, \ldots, f_m, e_1, e_2\}$ будет линейно-независимой системой векторов. При $m + 2 < n$ продолжаем наш процесс. Через $(n - m)$ шагов мы получим систему векторов $\{f_1, \ldots, f_m, e_1, e_2, \ldots, e_{n-m}\}$, составляющую базис $\mathscr{L}_n$.

Практически пополнение осуществляется следующим образом. Матрицу, состоящую из m координатных столбцов векторов $f_1, \ldots, f_m$, расширяем новыми столбцами так, чтобы ее ранг каждый раз увеличивался на единицу.

Определение. Совокупность векторов, принадлежащих одновременно подпространствам U_1 и U_2, называется *пересечением* подпространств и обозначается символом $U_1 \cap U_2$.

Покажем, что $U_1 \cap U_2$ — линейное подпространство. Действительно, если $x, y \in U_1 \cap U_2$, то $(\alpha x + \beta y) \in U_1$ (поскольку $x, y \in U_1$) и $(\alpha x + \beta y) \in U_2$ (поскольку $x, y \in U_2$). Следовательно, $(\alpha x + \beta y) \in U_1 \cap U_2$ при любых $\alpha, \beta \in \mathbf{K}$. Отметим, что пересечение—линейное подпространство подпространства U_1 и подпространства U_2.

Определение. Совокупность векторов вида $x = x_1 + x_2$, где x_1 пробегает все множество векторов подпространства U_1, x_2 — все множество векторов подпространства U_2, называется *суммой* подпространств и обозначается символом U_1+U_2. Если при этом пересечение подпространств $U_1 \cap U_2$ состоит только из нуль-вектора, то сумма называется *прямой* и обозначается символом $U_1 \oplus U_2$.

Покажем, что U_1+U_2—линейное подпространство (может быть совпадающее со всем пространством $\mathscr{L}_n$). В самом деле, пусть $x=x_1+x_2$, $y=y_1+y_2$—векторы из суммы подпространств. Но если $x_1, y_1 \in U_1$, то $(\alpha x_1+\beta y_1) \in U_1$, и если $x_2, y_2 \in U_2$, то $(\alpha x_2+\beta y_2) \in U_2$. В таком случае $\alpha x+\beta y=\alpha(x_1+x_2)+\beta(y_1+y_2)=$ $=(\alpha x_1+\beta y_1)+(\alpha x_2+\beta y_2)$ принадлежит сумме подпространств при любых $\alpha, \beta \in \mathbf{K}$.

Теорема 2. *Размерности подпространств, пересечения и суммы связаны соотношением*

$$\dim(U_1+U_2)=\dim U_1+\dim U_2-\dim(U_1 \cap U_2). \quad (16.1)$$

Доказательство. Пусть $\dim U_1=m_1$, $\dim U_2=m_2$, $\dim(U_1 \cap U_2)=m_3$. Пусть $\{h_1, \ldots, h_{m_3}\}$ — базис пересечения $U_1 \cap U_2$. Поскольку пересечение является подпространством подпространства U_1, то по теореме о пополнении базиса базис пересечения можно дополнить до базиса U_1 и векторы $\{f_1, \ldots, f_{m_1-m_3}, h_1, \ldots, h_{m_3}\}$ составят базис U_1. Аналогично найдутся векторы $g_1, \ldots, g_{m_2-m_3}$, что $\{h_1, \ldots, h_{m_3}, g_1, \ldots, g_{m_2-m_3}\}$ составят базис U_2.

Если мы сейчас покажем, что система векторов

$$\{f_1, \ldots, f_{m_1-m_3}, h_1, \ldots, h_{m_3}, g_1, \ldots, g_{m_2-m_3}\} \quad (16.2)$$

образует базис суммы U_1+U_2, то теорема будет доказана, так как число векторов в системе (16.2) равно $m_3+(m_1-m_3)+(m_2-m_3)=m_1+m_2-m_3$. Для того, чтобы выяснить, является ли система векторов (16.2) базисом суммы, надо прежде всего показать, что любой вектор суммы есть линейная комбинация векторов (16.2), и затем установить линейную независимость векторов (16.2).

Пусть $z=x_1+x_2$ — произвольный вектор суммы U_1+U_2. Заметив, что $x_1=\alpha_1 f_1+\ldots+\alpha_{m_1-m_3} f_{m_1-m_3}+$ $+\beta_1 h_1+\ldots+\beta_{m_3} h_{m_3}$ (поскольку x_1 вектор из U_1) и

$$x_2=\gamma_1 h_1+\ldots+\gamma_{m_3} h_{m_3}+\delta_1 g_1+\ldots+\delta_{m_2-m_3} g_{m_2-m_3}$$

(поскольку $x_2 \in U_2$), получим

$$z = \alpha_1 f_1 + \ldots + \alpha_{m_1-m_3} f_{m_1-m_3} + (\beta_1 + \gamma_1) h_1 + \ldots + \\ + (\beta_{m_3} + \gamma_{m_3}) h_{m_3} + \delta_1 g_1 + \ldots + \delta_{m_2-m_3} g_{m_2-m_3}.$$

Итак, *любой* вектор из суммы $U_1 + U_2$ представлен линейной комбинацией векторов (16.2). Осталось доказать, что (16.2) линейно-независимая система векторов. Допустим противное, что существует линейная комбинация

$$\lambda_1 f_1 + \ldots + \lambda_{m_1-m_3} f_{m_1-m_3} + \mu_1 h_1 + \ldots + \mu_{m_3} h_{m_3} + \\ + \nu_1 g_1 + \ldots + \nu_{m_2-m_3} g_{m_2-m_3} = 0, \quad (16.3)$$

где по крайней мере один из коэффициентов отличен от нуля. Вектор $t = \nu_1 g_1 + \ldots + \nu_{m_2-m_3} g_{m_2-m_3}$ принадлежит U_2. Но из (16.3) вытекает, что $t = -\lambda_1 f_1 - \ldots - \lambda_{m_1-m_3} f_{m_1-m_3} - \mu_1 h_1 - \ldots - \mu_{m_3} h_{m_3}$ и, следовательно, $t \in U_1$. Это означает, что $t \in U_1 \cap U_2$ и может быть представлен в виде линейной комбинации лишь векторов $h_1, \ldots, h_{m_3}$

$$t = \theta_1 h_1 + \ldots + \theta_{m_3} h_{m_3}.$$

Поэтому имеем

$$\nu_1 g_1 + \ldots + \nu_{m_2-m_3} g_{m_2-m_3} = \theta_1 h_1 + \ldots + \theta_{m_3} h_{m_3}$$

или

$$\nu_1 g_1 + \ldots + \nu_{m_2-m_3} g_{m_2-m_3} - \theta_1 h_1 - \ldots - \theta_{m_3} h_{m_3} = 0. \quad (16.4)$$

В силу линейной независимости векторов базиса U_2 из (16.4) следует, что $\nu_1 = 0, \ldots, \nu_{m_2-m_3} = 0$, $\theta_1 = 0, \ldots$ $\ldots, \theta_{m_3} = 0$. Условия (16.3) принимают вид

$$\lambda_1 f_1 + \ldots + \lambda_{m_1-m_3} f_{m_1-m_3} + \mu_1 h_1 + \ldots + \mu_{m_3} h_{m_3} = 0. \quad (16.5)$$

В силу линейной независимости векторов базиса U_1 получим, что наряду с коэффициентами ν_s все коэффициенты λ_p $(p = \overline{1, m_1 - m_3})$, μ_q $(q = \overline{1, m_3})$ также нулевые. Пришли к противоречию. Таким образом, система векторов (16.2) линейно-независима и образует базис суммы подпространств. Теорема доказана.

Обратимся к пространству решений однородной системы линейных уравнений $\sum_{k=1}^{n} a_k^i \xi^k = 0$ $(i = \overline{1, m})$ и выясним геометрический смысл этого пространства. С этой целью рассмотрим множество матриц-столбцов, составленных из n чисел. Если ввести операции сложения матриц-столбцов, умножения на число матрицы-столбца, нуль-столбец и противоположный столбец, как это сделано в лекции 1, то мы получим выполнение всех аксиом (A), (B). Тем самым мы устанавливаем, что множество матриц-столбцов с введенными операциями образует линейное пространство. Базисом его будет совокупность линейно-независимых матриц-столбцов вида

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

так как любая матрица-столбец есть их линейная комбинация

$$a = \begin{bmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{bmatrix} = \alpha^1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \alpha^2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + \alpha^n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Полученное n-мерное линейное пространство обозначают обычно символом T_n, а линейное пространство матриц-строк — символом T_n^*. В предыдущей лекции (пример 2) нами показано, что пространство решений однородной системы уравнений образует $(n-r)$-мерное линейное пространство, векторами которого являются матрицы-столбцы, а базисом служит нормальная фундаментальная (либо просто фундаментальная) система решений. Таким образом, можно заключить, что *пространство решений однородной системы линейных уравнений есть подпространство пространства* T_n. *Его размерность равна* $n-r$, *где* r — *ранг матрицы системы уравнений.* Вспоминая определение линейной оболочки, мы можем также сказать, что *пространство решений однородной системы линейных уравнений есть линейная оболочка в пространстве* T_n, *натянутая на векторы фундаментальной системы решений.*

Если решения системы уравнений записывать в виде матриц-строк, то в предыдущем абзаце необходимо символ T_n заменить на символ T_n^*.

Обратимся к иллюстрирующим материал лекции примерам.

П р и м е р 1. Найти размерность и указать базис линейной оболочки, натянутой на векторы

$$a_1 = [1, 0, -1, 2], \; a_2 = [2, 1, 1, 1], \; a_3 = [4, 1, -1, 5].$$

Составляем матрицу из матриц-строк, находим ее ранг

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

Ранг $A = 2$. Первая и вторая строка являются базисными. Следовательно, размерность оболочки равна двум, базис составляют векторы $\{a_1, a_2\}$.

П р и м е р 2. Найти размерность и указать базис пересечения подпространств U_1, U_2, являющихся линейными оболочками, натянутыми на следующие системы векторов:

(1) $a_1 = [1, 0, -1, 2], \; a_2 = [2, 1, 1, 1], \; a_3 = [4, 1, -1, 5]$;

(2) $b_1 = [-2, 0, 2, -4], b_2 = [0, 0, 1, 1], b_3 = [-2, 0, 3, -3],$

$$b_4 = [-4, 0, 5, -7].$$

Р е ш е н и е. Согласно примеру 1 $\dim U_1 = 2$ и базис $\{a_1, a_2\}$. Поступаем так же, как и в примере 1, с системой векторов $b_1, \ldots, b_4$.

$$B = \begin{bmatrix} -2 & 0 & 2 & -4 \\ 0 & 0 & 1 & 1 \\ -2 & 0 & 3 & -3 \\ -4 & 0 & 5 & -7 \end{bmatrix} \sim \begin{bmatrix} -2 & 0 & 2 & -4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \sim$$

$$\sim \begin{bmatrix} -2 & 0 & 2 & -4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Заключаем, что $\dim U_2 = 2$. Его базис — $\{b_1, b_2\}$.

Ищем векторы, принадлежащие одновременно как U_1, так и U_2

$$\alpha_1 [1 \;\; 0 \;\; -1 \;\; 2] + \alpha_2 [2 \;\; 1 \;\; 1 \;\; 1] = \alpha_3 [-2 \;\; 0 \;\; 2 \;\; -4] + \\ + \alpha_4 [0 \;\; 0 \;\; 1 \;\; 1].$$

Получим систему линейных уравнений с неизвестными $\alpha_1, \alpha_2, \alpha_3, \alpha_4$

$$\left.\begin{array}{r} \alpha_1+2\alpha_2+2\alpha_3=0 \\ \alpha_2=0 \\ -\alpha_1+\alpha_2-2\alpha_3-\alpha_4=0 \\ 2\alpha_1+\alpha_2+4\alpha_3-\alpha_4=0 \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \alpha_1=-2c \\ \alpha_2=0 \\ \alpha_3=c \\ \alpha_4=0 \end{array}\right. \quad \text{(}c\text{ — произвольная постоянная)}$$

Векторы пересечения x имеют вид: $x=-2c\,[1\ \ 0\ \ -1\ \ 2]$. Пересечение одномерно, его базис $e=[1\ \ 0\ \ -1\ \ 2]$.

Лекция 17

ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ И ИХ МАТРИЧНОЕ ПРЕДСТАВЛЕНИЕ. ДЕЙСТВИЯ НАД ОПЕРАТОРАМИ И МАТРИЦАМИ

Пусть заданы линейные пространства $\mathscr{L}_n$ и M_m. Пусть $\boldsymbol{A}$ — отображение пространства $\mathscr{L}_n$ в пространство M_m. Образ вектора $x \in \mathscr{L}_n$ будем обозначать символом $\boldsymbol{A}x$ ($\boldsymbol{A}x \in M_m$).

Определение. Отображение $\boldsymbol{A}: \mathscr{L}_n \to M_m$ называется *линейным отображением* (иначе *морфизмом*), если выполнено $\boldsymbol{A}(x_1+x_2)=\boldsymbol{A}x_1+\boldsymbol{A}x_2$ и $\boldsymbol{A}(\lambda x)=\lambda \boldsymbol{A}x$ для любых векторов x_1, x_2, x из $\mathscr{L}_n$ и любого $\lambda \in \mathbf{K}$.

В дальнейшем все линейные отображения обозначаем большими прописными латинскими буквами.

Если $\boldsymbol{A}$ отображает векторы $\mathscr{L}_n$ *на все множество* векторов M_m, то отображение называется *эпиморфизмом.* Если же $\boldsymbol{A}$ задает отображение *на часть множества* векторов M_m, но отображение *взаимно-однозначное* ($x \leftrightarrow \boldsymbol{A}x$), то $\boldsymbol{A}$ называется *мономорфизмом.* Отображение $\boldsymbol{A}$, являющееся одновременно эпиморфизмом и мономорфизмом, называется *изоморфизмом.*

В случае, когда пространство M_m является n-мерным, как и $\mathscr{L}_n$ ($m=n$), линейное отображение $\boldsymbol{A}: \mathscr{L}_n \to M_n$ называется *эндоморфизмом* или *линейным оператором.*

Сейчас мы докажем теорему, на основании которой M_n можно отождествить с $\mathscr{L}_n$ и считать, что эндоморфизм (линейный оператор) задает отображение $\mathscr{L}_n$ в самого себя. Поэтому для эндоморфизма $\boldsymbol{A}$ часто употребляют символ $\boldsymbol{A}=\operatorname{End}\mathscr{L}_n$.

Теорема. *Любые два n-мерных пространства $\mathscr{L}_n$ и M_n (над полем* **K**) *изоморфны между собой.*

Действительно, пусть в $\mathscr{L}_n$ выбран базис $\{e_1, \dots, e_n\}$, а в M_n — базис $\{g_1, \dots, g_n\}$. Построим отображение $A: \mathscr{L}_n \to M_n$ по следующему правилу. Каждому $x = \sum_{k=1}^{n} \xi^k e_k$ сопоставим $y = \sum_{k=1}^{n} \xi^k g_k$, т. е. координаты образа $Ax = y$ в базисе $\{g_1, \dots, g_n\}$ те же самые, что и у вектора x в базисе $\{e_1, \dots, e_n\}$. Это отображение *линейное*, так как вектору

$$\alpha x_1 + \beta x_2 = \alpha \sum_{k=1}^{n} \xi_1^k e_k + \beta \sum_{k=1}^{n} \xi_2^k e_k = \sum_{k=1}^{n} (\alpha \xi_1^k + \beta \xi_2^k) e_k$$

соответствует, согласно правилу, вектор

$$\sum_{k=1}^{n} (\alpha \xi_1^k + \beta \xi_2^k) g_k = \alpha \sum_{k=1}^{n} \xi_1^k g_k + \beta \sum_{k=1}^{n} \xi_2^k g_k = \alpha A x_1 + \beta A x_2 .$$

Построенное отображение — мономорфизм, потому что при $x_1 \neq x_2$ имеем различные координатные матрицы-столбцы, построенные из ξ_1^k и ξ_2^k, а, следовательно, векторы Ax_1 и Ax_2 также различны между собой.

Построенное отображение в то же время и эпиморфизм, так как для любого $y = \sum_{k=1}^{n} \eta^k g_k$ существует прообраз $x = \sum_{k=1}^{n} \eta^k e_k$, для которого $Ax = y$. Теорема доказана.

Тем самым все n-мерные линейные пространства (если учитывать только линейную структуру) не различимы между собой. Поэтому символ M для другого линейного пространства можно не употреблять и писать $\mathscr{L}_n$, $\mathscr{L}_m$, $\mathscr{L}_p$ и т. д.

Пусть A — линейное отображение из $\mathscr{L}_n$ в $\mathscr{L}_m$. Из линейности A следует $A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2$ и вообще $A\left(\sum_{k=1}^{s} \alpha_k x_k\right) = \sum_{k=1}^{s} \alpha_k A x_k$. Пусть $\{e_1, \dots, e_n\}$ — базис $\mathscr{L}_n$, $\{g_1, \dots, g_m\}$ — базис $\mathscr{L}_m$. Векторам базиса e_k $(k = \overline{1, n})$ соответствуют образы $Ae_k \in \mathscr{L}_m$, которые мы можем разложить по векторам базиса $\{g_1, \dots, g_m\}$

$$\left.\begin{array}{l} Ae_1 = \alpha_1^1 g_1 + \alpha_1^2 g_2 + \ldots + \alpha_1^m g_m \\ Ae_2 = \alpha_2^1 g_1 + \alpha_2^2 g_2 + \ldots + \alpha_2^m g_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Ae_n = \alpha_n^1 g_1 + \alpha_n^2 g_2 + \ldots + \alpha_n^m g_m \end{array}\right\} \left(Ae_k = \sum_{i=1}^{m} \alpha_k^i g_i\right). \qquad (17.1)$$

Записывая координаты векторов $Ae_1, \ldots, Ae_n$ в виде столбцов, получим матрицу

$$A = \begin{bmatrix} \alpha_1^1 & \alpha_2^1 \ldots \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 \ldots \alpha_n^2 \\ \vdots & \vdots \quad \vdots \\ \alpha_1^m & \alpha_2^m \ldots \alpha_n^m \end{bmatrix}, \qquad (17.2)$$

которую кратко будем записывать в виде $A = [\alpha_k^i]$ (верхний индекс — номер строки, нижний индекс — номер столбца) и называть *матрицей линейного отображения* $A: \mathscr{L}_n \to \mathscr{L}_m$ в данных базисах $\{e_1, \ldots, e_n\}$, $\{g_1, \ldots, g_m\}$.

Если A — эндоморфизм (линейный оператор), то его матрица необходимо *квадратная*.

Если координаты вектора x заданы: $x = \sum_{k=1}^{n} \xi^k e_k$, то как найти координаты образа Ax? Используя линейность отображения A, получим

$$y = Ax = A\left(\sum_{k=1}^{n} \xi^k e_k\right) = \sum_{k=1}^{n} \xi^k Ae_k =$$
$$= \sum_{k=1}^{n} \xi^k \left(\sum_{i=1}^{m} \alpha_k^i g_i\right) = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} \alpha_k^i \xi^k\right) g_i.$$

Если обозначить координаты вектора y через η^i, то имеем

$$\eta^i = \sum_{k=1}^{n} \alpha_k^i \xi^k. \qquad (17.3)$$

Формула (17.3) служит для нахождения координат образа Ax, если известны матрица отображения и координаты вектора x.

Итак, *каждому линейному отображению* A *соответствует (в фиксированных базисах) прямоугольная матрица* $A = [\alpha_k^i]$ $(i = \overline{1, m};\ k = \overline{1, n})$.

Обратно, пусть задана некоторая матрица $A = [\alpha_k^i]$ $(i = \overline{1, m};\ k = \overline{1, n})$ и заданы базис $\{e_1, \ldots, e_n\}$ в $\mathscr{L}_n$ и

$\{g_1, \ldots, g_m\}$ в $\mathcal{L}_m$. Зададим отображение $\boldsymbol{A}: \mathcal{L}_n \to \mathcal{L}_m$ по следующему закону. Каждому вектору $x = \sum_{k=1}^{n} \xi^k e_k$ сопоставим вектор $y = \sum_{i=1}^{m} \eta^i g_i$, где числа η^i определяются с помощью матрицы $A = [\alpha_k^i]$ по формуле (17.3). Заданное отображение линейное, так как вектору $\alpha x_1 + \beta x_2 = \sum_{k=1}^{n} (\alpha \xi_1^k + \beta \xi_2^k) e_k$ по (17.3) соответствует вектор с координатами

$$\sum_{k=1}^{n} \alpha_k^i (\alpha \xi_1^k + \beta \xi_2^k) = \alpha \sum_{k=1}^{n} \alpha_k^i \xi_1^k + \beta \sum_{k=1}^{n} \alpha_k^i \xi_2^k = \alpha \eta_1^i + \beta \eta_2^i,$$

откуда следует, что $\boldsymbol{A}(\alpha x_1 + \beta x_2) = \alpha \boldsymbol{A} x_1 + \beta \boldsymbol{A} x_2$. Воспользовавшись формулой (17.3), подсчитаем координаты векторов $\boldsymbol{A} e_1, \ldots, \boldsymbol{A} e_n$.

У вектора e_1 первая координата равна 1, остальные равны нулю. Поэтому из (17.3) следует, что вектор $\boldsymbol{A} e_1$ имеет координаты $\alpha_1^1, \alpha_1^2, \ldots, \alpha_1^m$—первый столбец матрицы A. У вектора e_2 вторая координата равна 1, а остальные равны нулю. Поэтому из (17.3) следует, что вектор $\boldsymbol{A} e_2$ имеет координаты $\alpha_2^1, \alpha_2^2, \ldots, \alpha_2^m$ — второй столбец матрицы A. Аналогично получим, что остальные столбцы матрицы A есть координаты векторов $\boldsymbol{A} e_3, \ldots, \boldsymbol{A} e_n$. Таким образом, матрица построенного отображения $\boldsymbol{A}$ совпадает с заданной матрицей $\boldsymbol{A}$. Тем самым, *каждой матрице* (при фиксированных базисах $\{e_1, \ldots, e_n\}$, $\{g_1, \ldots, g_m\}$) *однозначно сопоставляется линейное отображение* $\boldsymbol{A}$. Можно заключить, что *между множеством всех линейных отображений (операторов) и множеством всех прямоугольных (квадратных) матриц существует взаимно-однозначное соответствие.*

Пример. Найти матрицу оператора, переводящего векторы $x_1 = [1, 0, -1]$, $x_2 = [0, 1, -2]$, $x_3 = [1, 1, 1]$ в векторы $y_1 = [0, 1, -1]$, $y_2 = [1, 0, 2]$, $y_3 = [0, 0, -1]$.

Решение. Используем (17.3) для определения 9 неизвестных α_k^i $(i, k = 1, 2, 3)$

$$\begin{cases} 0 = \alpha_1^1 - \alpha_3^1 \\ 1 = \alpha_1^2 - \alpha_3^2 \\ -1 = \alpha_1^3 - \alpha_3^3 \end{cases} \quad \begin{cases} 1 = \alpha_2^1 - 2\alpha_3^1 \\ 0 = \alpha_2^2 - 2\alpha_3^2 \\ 2 = \alpha_2^3 - 2\alpha_3^3 \end{cases} \quad \begin{cases} 0 = \alpha_1^1 + \alpha_2^1 + \alpha_3^1 \\ 0 = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \\ -1 = \alpha_1^3 + \alpha_2^3 + \alpha_3^3. \end{cases}$$

Данная система уравнений распадается на 3 подсистемы каждая с тремя неизвестными

$$\begin{cases} \alpha_1^1 + \alpha_2^1 + \alpha_3^1 = 0 \\ \alpha_2^1 - 2\alpha_3^1 = 1 \\ \alpha_1^1 - \alpha_3^1 = 0, \end{cases} \quad \begin{cases} \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 0 \\ \alpha_2^2 - 2\alpha_3^2 = 0 \\ \alpha_1^2 - \alpha_3^2 = 1, \end{cases} \quad \begin{cases} \alpha_1^3 + \alpha_2^3 + \alpha_3^3 = -1 \\ \alpha_2^3 - 2\alpha_3^3 = 2 \\ \alpha_1^3 - \alpha_3^3 = -1. \end{cases}$$

Решая их, находим вид матрицы оператора

$$A = \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ \frac{3}{4} & -\frac{1}{2} & -\frac{1}{4} \\ -\frac{3}{2} & 1 & -\frac{1}{2} \end{bmatrix}.$$

Среди множества отображений выделим 3 отображения, играющие важную роль в теории линейных операторов.

Определение. *Нулевым* морфизмом называется отображение O: $\mathscr{L}_n \to \mathscr{L}_m$, которое любой вектор переводит в нуль-вектор пространства $\mathscr{L}_m$, т. е. $Ox = 0$ для любого $x \in \mathscr{L}_n$.

Легко найти вид матрицы нулевого морфизма. Поскольку $Oe_k = 0$, то все $\alpha_k^i = 0$ $(i = \overline{1, m};\ k = \overline{1, n})$ и матрица имеет вид

$$0 = \begin{bmatrix} 0 & 0 \dots 0 \\ 0 & 0 \dots 0 \\ \vdots & \vdots \quad \vdots \\ 0 & 0 \dots 0 \end{bmatrix}. \tag{17.4}$$

Определение. Отображение A' называется *противоположным* по отношению к отображению A, если для любого $x \in \mathscr{L}_n$ выполнено $A'x = -Ax$.

Пусть матрица отображения A есть $A = [\alpha_k^i]$ $(i = \overline{1, m};\ k = \overline{1, n})$. Поскольку $A'e_k = -Ae_k = -\sum_{i=1}^{m} \alpha_k^i g_i = \sum_{i=1}^{m} (-\alpha_k^i) g_i$, то матрица противоположного оператора составлена из чисел α_k^i, умноженных на -1, т. е. $A' = [-\alpha_k^i]$.

Определение. Эндоморфизм (оператор) $\mathcal{E}$ называется *тождественным*, если он оставляет на месте все векторы пространства, т. е. $\mathcal{E}x = x$ для любого $x \in \mathcal{L}_n$.

Поскольку $\mathcal{E}e_k = e_k$, то $\alpha_k^i = \delta_k^i = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$ — *символ Кронекера* ($i, k = \overline{1, n}$). Матрица тождественного оператора имеет вид

$$E = \begin{bmatrix} 1 & 0 \dots 0 \\ 0 & 1 \dots 0 \\ \vdots & \vdots \quad \vdots \\ 0 & 0 \dots 1 \end{bmatrix} \tag{17.5}$$

и носит название *единичной матрицы*.

Перейдем к рассмотрению действий над отображениями и построим соответствующие действия для матриц.

Определение. Если для любого $x \in \mathcal{L}_n$ $\boldsymbol{A}x = \boldsymbol{B}x$, то отображения считаются *одинаковыми* (*равными*).

Поскольку $\boldsymbol{A}e_k = \boldsymbol{B}e_k$, то $\alpha_k^i = \beta_k^i$ ($i = \overline{1, m}$; $k = \overline{1, n}$). Поэтому две матрицы A и B считаются *совпадающими* (пишем $A = B$), если их элементы на соответствующих местах равны друг другу.

Определение. Суммой линейных отображений $\boldsymbol{A}$ и $\boldsymbol{B}$ называется такое отображение $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$, что для любого $x \in \mathcal{L}_n$ $\boldsymbol{C}x = \boldsymbol{A}x + \boldsymbol{B}x$. Покажем, что $\boldsymbol{C}$ — линейное отображение. Действительно, $\boldsymbol{C}(\alpha x_1 + \beta x_2) = \boldsymbol{A}(\alpha x_1 + \beta x_2) + \boldsymbol{B}(\alpha x_1 + \beta x_2) = \alpha \boldsymbol{A}x_1 + \beta \boldsymbol{A}x_2 + \alpha \boldsymbol{B}x_1 + \beta \boldsymbol{B}x_2 = \alpha(\boldsymbol{A}x_1 + \boldsymbol{B}x_1) + \beta(\boldsymbol{A}x_2 + \boldsymbol{B}x_2) = \alpha \boldsymbol{C}x_1 + \beta \boldsymbol{C}x_2$, т. е. при любых $x_1, x_2 \in \mathcal{L}_n$ и любых $\alpha, \beta \in \mathbf{K}$ $\boldsymbol{C}(\alpha x_1 + \beta x_2) = \alpha \boldsymbol{C}x_1 + \beta \boldsymbol{C}x_2$. Отображение $\boldsymbol{C}$ — линейное отображение.

Поскольку

$$\boldsymbol{C}e_k = \boldsymbol{A}e_k + \boldsymbol{B}e_k = \sum_{i=1}^{m} \alpha_k^i g_i + \sum_{i=1}^{m} \beta_k^i g_i = \sum_{i=1}^{m} (\alpha_k^i + \beta_k^i) g_i,$$

то, обозначая элементы матрицы отображения $\boldsymbol{C}$ через γ_k^i, мы получим

$$\gamma_k^i = \alpha_k^i + \beta_k^i \quad (i = \overline{1, m};\ k = \overline{1, n}). \tag{17.6}$$

Поэтому под *суммой матриц* $A = [\alpha_k^i]$ *и* $B = [\beta_k^i]$ *понимаем матрицу* $C = [\gamma_k^i]$, *элементы которой связаны соотношением* (17.6).

Легко проверить, что как для отображений, так и для матриц выполняются следующие четыре условия (записываем только для матриц):

$$(1)\ A+B=B+A; \quad (2)\ (A+B)+C=A+(B+C);$$
$$(3)\ A+O=A; \quad (4)\ A+A'=0. \qquad (17.7)$$

Определение. Если для любого $x \in \mathscr{L}_n$ выполнено $(\lambda A)x=\lambda(Ax)$, то отображение λA называется *произведением числа λ на отображение A* (введена *операция умножения на число отображения A*).

Поскольку $\quad (\lambda A)\,e_k=\lambda A e_k=\lambda \sum_{i=1}^{m} \alpha_k^i g_i=\sum_{i=1}^{m}(\lambda\alpha_k^i)\,g_i$,

то элементы матрицы оператора λA получаются путем умножения на λ всех элементов матрицы $A=[\alpha_k^i]$. Поэтому под произведением числа λ и матрицы A понимается матрица с элементами $[\lambda\alpha_k^i]$ $(i=\overline{1, m};\ k=\overline{1, n})$.

Легко убедиться, что выполнены условия

$$(1)\ 1\cdot A=A; \quad (2)\ (\lambda\mu)\,A=\lambda(\mu A);$$
$$(3)\ (\lambda+\mu)\,A=\lambda A+\mu A; \quad (4)\ \lambda(A+B)=\lambda A+\lambda B. \qquad (17.8)$$

Введенные операции и выполнение свойств (17.7) для суммы матриц и свойств (17.8) для операции умножения матрицы на число приводит нас к выводу (см. аксиомы группы (A) и группы (B) для линейных пространств): *совокупность всех матриц вида $m\times n$ представляет собой линейное векторное пространство, где роль векторов играют прямоугольные матрицы, имеющие m строк и n столбцов.*

Какова размерность введенного линейного пространства? Чтобы ответить на этот вопрос, рассмотрим mn матриц вида

$$e_{ik}=\begin{bmatrix} 0\ldots0\ldots0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ 0\ldots1\ldots0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ 0\ldots0\ldots0 \end{bmatrix}, \qquad (17.9)$$

где на пересечении i-й строки и k-го столбца стоит 1, а остальные элементы матрицы—нули. Эти матрицы линейно-независимы, так как, составив линейную комбинацию $\sum_{i=1}^{m}\sum_{k=1}^{n}\lambda_{ik}e_{ik}$ с произвольными коэффициентами λ_{ik} и приравняв ее нулевой матрице (нуль-вектору), из равенства матриц получим, что $\lambda_{ik}=0$ ($i=\overline{1,m}$; $k=\overline{1,n}$).

Кроме того, любая матрица $A=[\alpha_k^i]$ может быть представлена как линейная комбинация матриц e_{ik}: $A=\sum_{i=1}^{m}\sum_{k=1}^{n}\alpha_k^i e_{ik}$. Поэтому матрицы (17.9) играют роль базисных векторов в линейном пространстве матриц $m\times n$. Таких матриц mn. Поэтому размерность рассматриваемого линейного пространства равна mn.

Ясно, что все изложенное выше для отображений $A:\mathscr{L}_n\to\mathscr{L}_m$ и прямоугольных матриц имеет место для операторов (эндоморфизмов) и квадратных матриц $n\times n$. *Размерность линейного пространства операторов равна* n^2.

Лекция 18

КОМПОЗИЦИЯ ОТОБРАЖЕНИЙ И УМНОЖЕНИЕ МАТРИЦ. ОБРАТНЫЙ ОПЕРАТОР И ОБРАТНАЯ МАТРИЦА. ОБРАЗ И ЯДРО ЛИНЕЙНОГО ОТОБРАЖЕНИЯ

Пусть $\boldsymbol{A}$—линейное отображение из $\mathscr{L}_n$ в $\mathscr{L}_m$, $\boldsymbol{B}$—линейное отображение из $\mathscr{L}_m$ в $\mathscr{L}_p$, т. е. $\mathscr{L}_n\xrightarrow{A}\mathscr{L}_m\xrightarrow{B}\mathscr{L}_p$. Тогда возникает отображение $\boldsymbol{C}:\mathscr{L}_n\to\mathscr{L}_p$, которое называется *композицией* отображений $\boldsymbol{A}$ и $\boldsymbol{B}$ и обозначается символом $\boldsymbol{C}=\boldsymbol{B}\circ\boldsymbol{A}$ (при композиции последовательность действий пишется справа налево). Другими словами, для любого $x\in\mathscr{L}_n$ имеем $\boldsymbol{C}x=\boldsymbol{B}(\boldsymbol{A}x)$.

Покажем, что $\boldsymbol{C}$—линейный оператор. Этот факт легко проверяется, так как $\boldsymbol{C}(\alpha x_1+\beta x_2)=\boldsymbol{B}(\boldsymbol{A}(\alpha x_1+\beta x_2))=$ $=\boldsymbol{B}(\alpha\boldsymbol{A}x_1+\beta\boldsymbol{A}x_2)=\alpha\boldsymbol{B}(\boldsymbol{A}x_1)+\beta\boldsymbol{B}(\boldsymbol{A}x_2)=\alpha\boldsymbol{C}x_1+\beta\boldsymbol{C}x_2$.

Как записать композицию отображений в матричном представлении?

Пусть $\{e_1,\dots,e_n\}$, $\{g_1,\dots,g_m\}$, $\{h_1,\dots,h_p\}$ — базисы в пространствах $\mathscr{L}_n$, $\mathscr{L}_m$, $\mathscr{L}_p$, а соответствующие матрицы отображений — $A=[\alpha_k^i]$ $(i=\overline{1, m};\ k=\overline{1, n})$, $B=[\beta_i^s]$ $(s=\overline{1, p})$, $C=[\gamma_k^s]$ $(s=\overline{1, p};\ k=\overline{1, n})$. Можно написать следующую цепочку преобразований:

$$e_k \xrightarrow{A} \sum_{i=1}^{m} \alpha_k^i g_i \xrightarrow{B} \sum_{i=1}^{m} \alpha_k^i \left(\sum_{s=1}^{p} \beta_i^s h_s \right) = \sum_{s=1}^{p} \left(\sum_{i=1}^{m} \beta_i^s \alpha_k^i \right) h_s . \qquad (18.1)$$

С другой стороны имеем

$$Ce_k = \sum_{s=1}^{p} \gamma_k^s h_s . \qquad (18.2)$$

Сравнивая (18.1), (18.2) и учитывая линейную независимость векторов $h_1, \ldots, h_p$, получим

$$\gamma_k^s = \sum_{i=1}^{m} \beta_i^s \alpha_k^i , \qquad (18.3)$$

т. е. элемент матрицы C, стоящий на пересечении s-й строки и k-го столбца представляет собой *сумму произведений элементов s-й строки матрицы B с соответствующими элементами* (т. е. стоящими под теми же номерами, что и элементы s-й строки) *k-го столбца матрицы A.*

Построенную таким способом матрицу C называют *произведением матриц B, A* (пишут $C=BA$), а правило умножения матриц (18.3) называют правилом „строка на столбец“. Из этого правила вытекает, что не любые матрицы B и A могут быть перемножены, а только такие, когда число столбцов матрицы B совпадает с числом строк матрицы A. В результате получится матрица, имеющая число строк, равное числу строк матрицы B, а число ее столбцов совпадает с числом столбцов матрицы A.

Пример. Найти произведение AB, если

$$A=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 3 \\ -1 & 0 & -1 \end{bmatrix}, \quad B=\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & -1 & -1 & 1 \\ 1 & -1 & 0 & 0 \end{bmatrix}.$$

Пользуясь правилом „строка на столбец“, имеем

$$AB=\begin{bmatrix} 1\cdot 1+(-1)\cdot 0+0\cdot 1 & 2+1+0 & 3+1+0 & -1-1+0 \\ 0\cdot 1+2\cdot 0+3\cdot 1 & 0-2-3 & 0-2+0 & 0+2+0 \\ -1\cdot 1+0\cdot 0-1\cdot 1 & -2+0+1 & -3+0+0 & 1+0+0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 4 & -2 \\ 3 & -5 & -2 & 2 \\ -2 & -1 & -3 & 1 \end{bmatrix}.$$

На этом примере видно, что произведение AB существует, а произведение BA не существует.

Если речь идет об операторах $\boldsymbol{A}, \boldsymbol{B}$, то соответствующие матрицы квадратные и, следовательно, $C = BA$ — снова квадратная матрица.

В общем случае для квадратных матриц $AB \neq BA$, в чем легко убедиться на следующем примере:

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix},$$

$$BA = \begin{bmatrix} -1 & -2 \\ 0 & 5 \end{bmatrix}, \quad AB \neq BA.$$

Когда $AB \neq BA$, то этот факт отмечают словами: матрицы A и B *не коммутируют* между собой. Если же матрицы A и B таковы, что $AB = BA$, то они называются *коммутирующими*.

Для всякой квадратной матрицы A вводится понятие не отрицательной степени матрицы, если по определению считать, что

$$A^0 = E, \ A^1 = A, \ A^2 = A \cdot A, \ldots, A^{n+1} = A \cdot A^n, \ldots, \tag{18.4}$$

где E — единичная матрица (17.5). Используя (18.4), тотчас получим, что $A^p \cdot A^q = A^{p+q} = A^q \cdot A^p$, так как $EA = A = AE$ для любой матрицы A.

В дальнейшем нам понадобится факт из теории матриц, который мы приведем без доказательства.

Теорема. *Детерминант произведения квадратных матриц равен произведению детерминантов сомножителей*, т. е.

$$\det(AB) = \det A \cdot \det B. \tag{18.5}$$

Доказательство приведено, например, в книге Г. Е. Шилова (см. литературу в конце лекций).

Перейдем к более подробному рассмотрению эндоморфизмов пространства $\mathscr{L}_n$, называемых далее операторами.

Определение. Оператор $\boldsymbol{B}$ называется *обратным* по отношению к оператору $\boldsymbol{A}$, если выполнено $\boldsymbol{B} \circ \boldsymbol{A} = \mathcal{E}$, где $\mathcal{E}$ — тождественный оператор.

В матричном представлении данное условие имеет вид $BA = E$, где E — единичная матрица (17.5).

Поэтому можно дать аналогичное определение обратной матрицы по отношению к матрице A.

О п р е д е л е н и е. Матрица A^{-1} называется ***обратной*** к матрице A, если выполнено условие

$$A^{-1}A=E. \tag{18.6}$$

Какие из матриц A имеют обратные? Чтобы ответить на этот вопрос, воспользуемся соотношением (18.5). Из (18.6), учитывая, что $\det E=1$, следует $(\det A^{-1})\times\times(\det A)=1$. Следовательно, *детерминант матрицы A непременно должен быть отличен от нуля.* Если матрица A *вырожденная* ($\det A=0$), то понятие об обратной матрице для нее отсутствует. Этот факт выражают словами, что всякая вырожденная матрица A обратной матрицы не имеет.

Пусть матрица $A=[\alpha_k^i]$ — *невырожденная* ($\det A\neq 0$). Как найти элементы матрицы A^{-1}, зная элементы матрицы A? Обозначим элементы A^{-1} через χ_k^i, где χ_k^i — неизвестные пока числа. Запишем условие (18.6) через элементы матриц

$$\sum_{s=1}^{n}\chi_s^i\alpha_k^s=\delta_k^i=\begin{cases}1, & i=k\\ 0, & i\neq k.\end{cases} \tag{18.7}$$

Придадим индексу i значение 1, а индекс k заставим меняться от 1 до n. В результате получим систему n уравнений с n неизвестными χ_s^1:

$$\left.\begin{array}{l} i=1,\ k=1:\chi_1^1\alpha_1^1+\chi_2^1\alpha_1^2+\ldots+\chi_n^1\alpha_1^n=1\\ i=1,\ k=2:\chi_1^1\alpha_2^1+\chi_2^1\alpha_2^2+\ldots+\chi_n^1\alpha_2^n=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ i=1,\ k=n:\chi_1^1\alpha_n^1+\chi_2^1\alpha_n^2+\ldots+\chi_n^1\alpha_n^n=0\end{array}\right\}. \tag{18.8}$$

Поскольку $\det A\neq 0$, то применяем формулу Крамера

$$\chi_j^1=\frac{\begin{vmatrix}\alpha_1^1\ldots 1\ldots\alpha_1^n\\ \alpha_2^1\ldots 0\ldots\alpha_2^n\\ \ldots\ldots\\ \alpha_n^1\ldots 0\ldots\alpha_n^n\end{vmatrix}}{\det A}=\frac{A_1^j}{\det A}, \tag{18.9}$$

где A_1^j — алгебраическое дополнение элемента α_1^j в матрице A. Заметим, что элемент χ_j^1 стоит на пересечении первой строки и j-го столбца, а алгебраическое дополнение подсчитывается для элемента, стоящего на пересечении первого столбца j-й строки. Если сейчас индексу i придать значение, равное 2, индекс k заставить меняться от 1 до n и разрешить по формулам Крамера соответствующую систему уравнений, то получим, что $\chi_j^2 = \dfrac{A_2^j}{\det A}$. Аналогично имеем

$$\chi_j^3 = \frac{A_3^j}{\det A}, \ldots, \chi_j^n = \frac{A_n^j}{\det A}.$$

Таким образом, элементы A^{-1} определены однозначным образом $\chi_k^i = \dfrac{A_i^k}{\det A}$ и, следовательно, A^{-1} имеет вид

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} A_1^1 & A_1^2 \ldots A_1^n \\ A_2^1 & A_2^2 \ldots A_2^n \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ A_n^1 & A_n^2 \ldots A_n^n \end{bmatrix}, \tag{18.10}$$

где (еще раз напоминаем) A_k^j — алгебраические дополнения элементов α_k^j матрицы A.

Пользуясь формулой разложения определителя по алгебраическим дополнениям своей и чужой строки (столбца) (лекция 2)

$$\sum_{k=1}^{n} \alpha_k^i A_k^j = \sum_{k=1}^{n} \alpha_i^k A_j^k = \begin{cases} \det A, & i = j \\ 0, & i \neq j, \end{cases}$$

легко проверить выполнение следующих равенств:

$$\left.\begin{array}{l} \text{(а) } AA^{-1} = E; \quad \text{(б) } (AB)^{-1} = B^{-1}A^{-1}; \\ \text{(в) } (A^{-1})^{-1} = A. \end{array}\right\}. \tag{18.11}$$

На основании (18.6) и (18.11а) заключаем, что

$$A^{-1}A = AA^{-1} = E.$$

Определив отрицательную степень матрицы A в виде $A^{-n} = A^{-1}A^{-1} \ldots A^{-1}$, формулу $A^pA^q = A^{p+q}$ мы распространим тем самым на все множество целых чисел p и q.

Пример. Найти обратную матрицу для

$$A=\begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Решение. Вычисляя определитель и алгебраические дополнения, находим: $\det A=-3$, $A_1^1=-1$, $A_1^2=2$, $A_1^3=-3$, $A_2^1=0$, $A_2^2=0$, $A_2^3=-3$, $A_3^1=-1$, $A_3^2=-1$, $A_3^3=0$. Поэтому

$$A^{-1}=\begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & 1 \\ 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix}.$$

Вернемся к рассмотрению линейных отображений $\boldsymbol{A}: \mathscr{L}_n \to \mathscr{L}_m$ и установим, что представляет собой множество векторов $\boldsymbol{A}x$ из $\mathscr{L}_m$.

Определение. Совокупность всех векторов пространства $\mathscr{L}_m$ $y=\boldsymbol{A}x$, где x пробегает все множество векторов $\mathscr{L}_n$, называется *образом отображения* $\boldsymbol{A}$ и обозначается символом $\operatorname{Im}\boldsymbol{A}$.

Пусть $x=\sum_{k=1}^{n}\xi^k e_k$. Тогда

$$\boldsymbol{A}x=\boldsymbol{A}\left(\sum_{k=1}^{n}\xi^k e_k\right)=\sum_{k=1}^{n}\xi^k \boldsymbol{A}e_k.$$

Вспоминая определение линейной оболочки, заключаем: *образ линейного отображения* $\boldsymbol{A}$ *есть линейная оболочка, натянутая на образы базисных векторов* $\{\boldsymbol{A}e_1, \ldots, \boldsymbol{A}e_n\}$. Нам известно, что координаты векторов $\boldsymbol{A}e_k$ составляют столбцы матрицы $A=[\alpha_k^i]$. Поэтому *размерность* $\operatorname{Im}\boldsymbol{A}$ *совпадает с максимальным числом линейно-независимых столбцов матрицы* A, т. е. с рангом A.

Итак, $\dim\operatorname{Im}\boldsymbol{A}=$ ранг A.

Еще одной важной характеристикой отображения $\boldsymbol{A}$ является понятие ядра отображения.

Определение. *Ядром линейного отображения* $\boldsymbol{A}$ (символ $\operatorname{Ker}\boldsymbol{A}$) называется множество всех векторов из $\mathscr{L}_n$, которые $\boldsymbol{A}$ отображает в нуль-вектор пространства $\mathscr{L}_m$, т. е. совокупность $x \in \mathscr{L}_n$ таких, что $\boldsymbol{A}x = 0$.

Покажем, что Ker $\boldsymbol{A}$ есть линейное подпространство пространства $\mathscr{L}_n$. В самом деле, пусть $x_1, x_2 \in \operatorname{Ker} \boldsymbol{A}$, т. е. $\boldsymbol{A}x_1 = 0$, $\boldsymbol{A}x_2 = 0$. Видим, что $\boldsymbol{A}(\alpha x_1 + \beta x_2) = \alpha \boldsymbol{A}x_1 + \beta \boldsymbol{A}x_2 = \alpha \cdot 0 + \beta \cdot 0 = 0$, и, следовательно, вектор $(\alpha x_1 + \beta x_2) \in \operatorname{Ker} \boldsymbol{A}$ при любых $\alpha, \beta \in \mathbf{K}$.

Какова размерность ядра отображения? Чтобы ответить на этот вопрос, запишем в координатах условие $\boldsymbol{A}x = 0$

$$\boldsymbol{A}x = \boldsymbol{A}\left(\sum_{k=1}^{n} \xi^k e_k\right) = \sum_{k=1}^{n} \xi^k \boldsymbol{A}e_k = \sum_{k=1}^{n} \xi^k \left(\sum_{i=1}^{m} \alpha_k^i g_i\right) = 0.$$

Из-за линейной независимости векторов $g_1, \ldots, g_m$ имеем

$$\sum_{k=1}^{n} \alpha_k^i \xi^k = 0 \quad (i = \overline{1, m}) \tag{18.12}$$

—однородную систему линейных уравнений, матрица которой совпадает с матрицей A. Следовательно, вектор x принадлежит пространству решений системы (18.12), и *ядро отображений* $\boldsymbol{A}$ *совпадает с пространством решений однородной системы уравнений с матрицей* A. Нам известно, что размерность пространства решений равна $n-r$, где r—ранг матрицы A. Подводя итоги, можно заключить, что *ядро* (Ker $\boldsymbol{A}$) *есть подпространство пространства* $\mathscr{L}_n$ *размерности* $n - r$.

Теорема. *Если* $B = AT$ (*или если* $B = TA$), *где* T—*невырожденная матрица, то ранг* $B =$ *ранг* A.

Доказательство. Имеем, что $\boldsymbol{B}x = \boldsymbol{A}(\mathscr{T} x)$ для любого $x \in \mathscr{L}_n$. Но $\operatorname{Im} \mathscr{T} = \mathscr{L}_n$, поскольку ранг $T = n$ (T—невырожденная матрица). Поэтому $\dim \operatorname{Im} \boldsymbol{B} = \dim \operatorname{Im} \boldsymbol{A}$. Ранги матриц B и A совпадают.

Отметим, что систему уравнений (18.12) в матричном представлении можно записать в виде $A[x] = 0$, где $[x]$ — матрица-столбец с неизвестными коэффициентами. Если через $[b]$ обозначить матрицу-столбец из свободных членов $b^1, b^2, \ldots, b^m$ неоднородной системы, то неоднородная система линейных уравнений в матричной форме примет вид $A[x] = [b]$.

Поэтому нахождение общего решения однородной системы уравнений можно всегда интерпретировать как отыскание ядра отображения, задаваемого матрицей системы, а нахождение общего решения неоднородной системы уравнений — как отыскание множества всех векторов пространства $\mathscr{L}_n$, которые отображаются с помощью A в фиксированный вектор b пространства $\mathscr{L}_m$.

В частности, когда имеем систему $A[x]=[b]$ из n уравнений с n неизвестными и с $\det A \neq 0$, то формулы Крамера приобретают вид

$$[x]=A^{-1}[b]. \tag{18.13}$$

Для доказательства достаточно умножить равенство $A[x]=[b]$ слева на A^{-1} и учесть, что $A^{-1}A[x]=E[x]=[x]$.

Лекция 19

СОБСТВЕННЫЕ ЗНАЧЕНИЯ И СОБСТВЕННЫЕ ВЕКТОРЫ ОПЕРАТОРА. ПЕРЕХОД К НОВОМУ БАЗИСУ. ИНВАРИАНТЫ ОПЕРАТОРА

Очень важную роль в теории линейных операторов и в физических приложениях играют понятия собственных векторов и собственных значений оператора $\boldsymbol{A}=\operatorname{End}\mathscr{L}_n$.

Определение. Подпространство U называется *инвариантным* относительно оператора $\boldsymbol{A}$, если для любого $x\in U$ образ $\boldsymbol{A}x\in U$, т. е. $\boldsymbol{A}(U)\subseteq U$.

Нас будут интересовать одномерные инвариантные подпространства, называемые *собственными направлениями* оператора $\boldsymbol{A}$.

Определение. Вектор x $(x\neq 0)$ называется *собственным вектором* оператора $\boldsymbol{A}$, если $\boldsymbol{A}x=\lambda x$. Число λ называется *собственным значением* оператора.

Теорема 1. *Если x—собственный вектор оператора $\boldsymbol{A}$, то для любого $\alpha\in\mathbf{K}$ вектор αx так же является собственным вектором $\boldsymbol{A}$ с тем же собственным значением, что и вектор x.*

В самом деле, пусть $Ax = \lambda x$. Тогда $A(\alpha x) = \alpha Ax = \alpha(\lambda x) = \lambda(\alpha x)$.

Теорема 2. *Множество всех собственных векторов, отвечающих одному и тому же собственному значению, образует линейное инвариантное подпространство.*

Действительно, если $Ax_1 = \lambda x_1$, $Ax_2 = \lambda x_2$, то $A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2 = \lambda(\alpha x_1 + \beta x_2)$. Теорема доказана.

Теорема 3. *Собственные векторы $x_1, \dots, x_s$ с попарно различными собственными значениями $\lambda_1, \dots, \lambda_s$ ($\lambda_i \neq \lambda_j$; $i \neq j$) образуют линейно-независимую систему векторов.*

Для доказательства теоремы применим метод индукции. При $s = 1$ теорема выполнена, так как собственный вектор не нуль-вектор. Пусть теорема верна для любой системы из k собственных векторов. Докажем, что она верна и для системы из $(k+1)$ собственного вектора. Составим линейную комбинацию с произвольными коэффициентами

$$\alpha_1 x_1 + \dots + \alpha_k x_k + \alpha_{k+1} x_{k+1} = 0. \tag{19.1}$$

Действуем оператором A, учитывая, что $Ax_s = \lambda_s x_s$. Получаем

$$\alpha_1 \lambda_1 x_1 + \dots + \alpha_k \lambda_k x_k + \alpha_{k+1} \lambda_{k+1} x_{k+1} = 0. \tag{19.2}$$

Из собственных чисел $\lambda_1, \dots, \lambda_{k+1}$ хотя бы одно отлично от нуля, так как собственные значения попарно различны. За счет перенумерации всегда можно считать, что $\lambda_{k+1} \neq 0$. Умножив (19.1) на λ_{k+1} и вычитая полученное соотношение из (19.2), получим

$$\alpha_1(\lambda_1 - \lambda_{k+1}) x_1 + \alpha_2(\lambda_2 - \lambda_{k+1}) x_2 + \dots + \alpha_k(\lambda_k - \lambda_{k+1}) x_k = 0. \tag{19.3}$$

Поскольку для любых k собственных векторов теорема верна, то векторы $x_1, \dots, x_k$ линейно-независимы и, следовательно,

$$\alpha_1(\lambda_1 - \lambda_{k+1}) = 0, \quad \alpha_2(\lambda_2 - \lambda_{k+1}) = 0, \dots, \alpha_k(\lambda_k - \lambda_{k+1}) = 0.$$

Это означает, что $\alpha_1 = 0$, $\alpha_2 = 0, \dots, \alpha_k = 0$, так как $\lambda_1 \neq \lambda_{k+1}$, $\lambda_2 \neq \lambda_{k+1}, \dots, \lambda_k \neq \lambda_{k+1}$. В таком случае из (19.1) следует равенство $\alpha_{k+1} x_{k+1} = 0$. Но x_{k+1} как собственный вектор не является нуль-вектором. Поэтому коэффициент α_{k+1} также равен нулю. Теорема доказана.

Как практически находить собственные значения и собственные векторы оператора?

Пусть $\{e_1, \ldots, e_n\}$ — базис в $\mathcal{L}_n$, $A = [\alpha^i_k]$ — матрица оператора. Равенство $\boldsymbol{A}x = \lambda x$ можно записать в виде $(\boldsymbol{A} - \lambda\mathcal{E})\, x = 0$, где $\mathcal{E}$ — тождественный оператор $(\mathcal{E}x = x)$. В матричной форме приведенное операторное равенство выглядит следующим образом:

$$(A - \lambda E)\,[x] = 0 \text{ или } \sum_{k=1}^{n} (\alpha^i_k - \lambda\delta^i_k)\,\xi^k = 0.$$

Подробно расписывая последнее равенство, получим систему n линейных уравнений

$$\begin{cases} (\alpha^1_1 - \lambda)\,\xi^1 + \alpha^1_2\,\xi^2 + \ldots + \alpha^1_n\,\xi^n = 0 \\ \alpha^2_1\,\xi^1 + (\alpha^2_2 - \lambda)\,\xi^2 + \ldots + \alpha^2_n\,\xi^n = 0 \\ \dots\dots\dots\dots\dots\dots \\ \alpha^n_1\,\xi^1 + \alpha^n_2\,\xi^2 + \ldots + (\alpha^n_n - \lambda)\,\xi^n = 0. \end{cases} \tag{19.4}$$

Из теории систем линейных уравнений следует, что данная система имеет не нулевое решение тогда и только тогда, когда определитель матрицы системы равен нулю (ранг $(A - \lambda E) < n$).

Приравнивая $\det(A - \lambda E)$ к нулю, мы получим *характеристическое уравнение*

$$\begin{vmatrix} \alpha^1_1 - \lambda & \alpha^1_2 & \ldots & \alpha^1_n \\ \alpha^2_1 & \alpha^2_2 - \lambda & \ldots & \alpha^2_n \\ \alpha^n_1 & \alpha^n_2 & \ldots & \alpha^n_n - \lambda \end{vmatrix} = (-1)^n \lambda^n + I_1 \lambda^{n-1} + \ldots + I_n = 0. \tag{19.5}$$

Полином n-й степени, стоящий в левой части равенства (19.5), называется *характеристическим полиномом* матрицы A. *Корни этого полинома суть собственные значения оператора.*

Итак, *для того, чтобы оператор* $\boldsymbol{A}$ *имел собственные векторы, необходимо и достаточно, чтобы характеристический полином его матрицы A был равен нулю, причем соответствующие собственные значения оператора совпадают с корнями характеристического полинома.*

Если мы рассматриваем комплексные линейные пространства ($\mathbf{K} \equiv \mathbf{C}$), то согласно основной теореме алгебры (о наличии корней полинома) имеем n корней (включая кратные)

$$\underbrace{\lambda_1, \ldots, \lambda_1}_{r_1}; \quad \underbrace{\lambda_2, \ldots, \lambda_2}_{r_2}; \ldots ; \underbrace{\lambda_s, \ldots, \lambda_s}_{r_s}. \tag{19.6}$$

Тем самым мы найдем s попарно различных собственных значений оператора. Подставляя λ_k в систему (19.4) вместо параметра λ и составляя нормальную фундаментальную систему решений, мы находим базис подпространства собственных векторов, отвечающих данному собственному значению λ_k оператора $\mathcal{A}$.

Если мы рассматриваем вещественные линейные пространства ($\mathbf{K} \equiv \mathbf{R}$), то ситуация складывается несколько сложнее, чем в предыдущем случае. Когда вещественные корни у характеристического полинома отсутствуют (корни только комплексные), то мы говорим, что вещественных собственных векторов оператор не имеет. Пусть $\underbrace{\lambda_1, \ldots, \lambda_1}_{p_1}; \ldots ; \underbrace{\lambda_m, \ldots, \lambda_m}_{p_m}$ — вещественные корни полинома, а $\underbrace{\alpha_1 \pm i\beta_1, \ldots, \alpha_1 \pm i\beta_1}_{q_1};$ $\ldots ; \underbrace{\alpha_t \pm i\beta_t, \ldots, \alpha_t \pm i\beta_t}_{q_t}$ — комплексные и комплексно-сопряженные корни полинома (вспомним, что у полинома с вещественными коэффициентами наряду с комплексным корнем $\alpha + i\beta$ непременно присутствует корень $\alpha - i\beta$), причем $p_1 + \ldots + p_m + 2q_1 + \ldots + 2q_t = n$. В вещественном линейном пространстве m вещественным собственным значениям $\lambda_1, \ldots, \lambda_m$ будут отвечать вещественные собственные векторы (процедура нахождения базиса оболочки собственных векторов, отвечающих фиксированному вещественному собственному значению, полностью совпадает с указанной выше для комплексных пространств, только координаты ξ^k — вещественные числа), а комплексным собственным значениям будут отвечать комплексные собственные векторы (их координаты ξ^k—комплексные числа).

Пример. Найти собственные значения и собственные векторы оператора, заданного матрицей

$$\begin{bmatrix} -1 & 0 & 1 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & -1 \end{bmatrix}.$$

Составляем характеристическое уравнение

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

Имеем 2-кратный вещественный корень $\lambda_1 = -1$ и комплексные корни $\mu_1 = i$, $\mu_2 = -i$. Подставляем вместо λ в систему вида (19.4) $\lambda_1 = -1$. Имеем

$$\begin{cases} 0\cdot\xi^1 + 0\cdot\xi^2 + \xi^3 + \xi^4 = 0 \\ 0\cdot\xi^1 + 0\cdot\xi^2 + 0\cdot\xi^3 - \xi^4 = 0 \\ 0\cdot\xi^1 + 0\cdot\xi^2 + 2\xi^3 + \xi^4 = 0 \\ 0\cdot\xi^1 + 0\cdot\xi^2 - 2\xi^3 + 0\cdot\xi^4 = 0, \end{cases}$$

откуда следует, что $\xi^3 = \xi^4 = 0$, а $\xi^1 = c^1$, $\xi^2 = c^2$ (c^1, c^2 — произвольные вещественные параметры). Строим нормальную фундаментальную систему решений, полагая сначала $c^1 = 1$, $c^2 = 0$, затем $c^1 = 0$, $c^2 = 1$. Получим два базисных собственных вектора

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

для оболочки собственных векторов, отвечающих собственному значению $\lambda_1 = -1$. Любой другой собственный вектор x, отвечающий $\lambda_1 = -1$, есть линейная комбинация векторов x_1 и x_2

$$x = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Подставляя вместо λ в систему вида (19.4) комплексный корень $\mu_1 = i$, получим следующую систему уравнений:

$$\begin{cases} (-1-i)\xi^1 + 0\cdot\xi^2 + \xi^3 + \xi^4 = 0 \\ 0\cdot\xi^1 + (-1-i)\xi^2 + 0\cdot\xi^3 - \xi^4 = 0 \\ 0\cdot\xi^1 + 0\cdot\xi^2 + (1-i)\xi^3 + \xi^4 = 0 \\ 0\cdot\xi^1 + 0\cdot\xi^2 - 2\xi^3 + (-1-i)\xi^4 = 0, \end{cases}$$

откуда находим, что

$$\xi^1 = \frac{1}{2}(1+i)c, \quad \xi^2 = \frac{1}{2}(1-i)^2 c, \ \xi^3 = c, \ \xi^4 = (i-1)c,$$

где c — произвольный комплексный параметр. Фиксируя значение c, полагая, например, $c=2$, получим базисный комплексный собственный вектор для одномерного инвариантного подпространства, отвечающего собственному значению $\mu_1 = i$

$$y_1 = \begin{bmatrix} 1+i \\ -2i \\ 2 \\ -2+2i \end{bmatrix}.$$

Если провести аналогичные рассмотрения с сопряженным корнем $\mu_2 = -i$, то при соответствующем выборе $\xi^3 = \bar{c}$ базисный собственный вектор, отвечающий собственному значению $\mu_2 = -i$, будет комплексно сопряженным к вектору y_1

$$y_2 = \begin{bmatrix} 1-i \\ 2i \\ 2 \\ -2-2i \end{bmatrix}.$$

Что можно сказать о размерности инвариантного пространства собственных векторов, отвечающих кратному собственному значению?

Теорема 4. *Размерность подпространства собственных векторов, отвечающих кратному собственному значению, не превосходит кратности этого значения.*

Пусть U—подпространство собственных векторов, отвечающих собственному значению λ_1 кратности r_1, имеет размерность m ($\dim U = m$). Пусть $\{f_1, \ldots, f_m\}$ — базис в U. Поскольку $f_1, \ldots, f_m$ — собственные векторы, отвечающие собственному значению λ_1, то

$$\boldsymbol{A}f_1 = \lambda_1 f_1, \ \boldsymbol{A}f_2 = \lambda_1 f_2, \ldots, \boldsymbol{A}f_m = \lambda_1 f_m. \quad (19.7)$$

Дополним базис U до базиса $\mathcal{L}_n$ векторами $e_{m+1}, \dots, e_n$. Образы этих векторов при действии оператора A есть линейные комбинации векторов базиса $\{f_1, \dots, f_m, e_{m+1}, \dots, e_n\}$

$$\left\{\begin{aligned} Ae_{m+1} &= \sum_{j=1}^{m} \alpha_{m+1}^{j} f_j + \sum_{s=m+1}^{n} \alpha_{m+1}^{s} e_s \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ Ae_n &= \sum_{j=1}^{m} \alpha_n^j f_j + \sum_{s=m+1}^{n} \alpha_n^s e_s. \end{aligned}\right. \tag{19.8}$$

Матрица A в базисе $\{f_1, \dots, f_m, e_{m+1}, \dots, e_n\}$, согласно (19.7), (19.8), имеет вид

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & \vdots & \alpha_{m+1}^1 & \dots & \alpha_n^1 \\ 0 & \lambda_1 & \dots & 0 & \vdots & \alpha_{m+1}^2 & \dots & \alpha_n^2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \\ 0 & 0 & \dots & \lambda_1 & \vdots & \alpha_{m+1}^m & \dots & \alpha_n^m \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & \vdots & \alpha_{m+1}^{m+1} & \dots & \alpha_n^{m+1} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \\ 0 & 0 & \dots & 0 & \vdots & \alpha_{m+1}^n & \dots & \alpha_n^n \end{bmatrix}. \tag{19.9}$$

Составляя характеристический полином $\det(A - \lambda E)$, получим

$$\det(A - \lambda E) = (\lambda_1 - \lambda)^m Q(\lambda), \tag{19.10}$$

где $Q(\lambda)$ — полином степени $n - m$.

Из (19.10) следует, очевидно, что кратность r_1 корня $\lambda = \lambda_1$ не меньше m ($\lambda = \lambda_1$ может быть и корнем полинома $Q(\lambda)$). Итак, $m \leqslant r_1$. Теорема доказана.

В том, что размерность подпространства собственных векторов бывает меньше кратности соответствующего корня, легко убедиться на следующем примере. Пусть

$$A = \begin{bmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

— матрица оператора. Ее характеристический полином имеет вид

$$\begin{vmatrix} -\lambda & 0 & 1 & -1 \\ -1 & -\lambda & 1 & -1 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix} = \lambda^3(\lambda-1).$$

Первый корень $\lambda_1 = 0$ имеет кратность 3, второй корень $\lambda_2 = 1$ однократен. В случае первого собственного значения получаем

$$\xi^3 - \xi^4 = 0, \quad -\xi^1 + \xi^3 - \xi^4 = 0, \quad \xi^4 = 0,$$

откуда следует, что $\xi^1 = \xi^3 = \xi^4 = 0$, $\xi^2 = c$ (c — произвольный вещественный параметр). Таким образом, подпространство собственных векторов, отвечающих собственному 3-кратному значению $\lambda_1 = 0$, имеет размерность, равную единице. Его произвольный вектор-x имеет вид

$$x = \alpha \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

В n-мерном линейном пространстве $\mathscr{L}$ любая система из линейно-независимых векторов служит базисом пространства, т. е. базисы в $\mathscr{L}_n$ могут быть выбраны различным образом. Возникают следующие вопросы: (1) как задать переход от одного базиса к другому? (2) как при этом меняются координаты вектора x? (3) как меняется матрица оператора при переходе к новому базису? (4) какие величины, связанные с матрицей оператора, остаются неизменными или, как говорят, являются инвариантами преобразования базисов? Ответим на эти вопросы.

Пусть $\{e_1, \dots, e_n\}$ — базис в $\mathscr{L}_n$ и $\{e_{1'}, \dots, e_{n'}\}$ — новый базис в $\mathscr{L}_n$. Разложим векторы $e_{i'}$ $(i = \overline{1, n})$ нового базиса по векторам базиса $\{e_1, \dots, e_n\}$

$$e_{i'} = \sum_{k=1}^{n} \tau_{i'}^{k} e_k \quad (i = \overline{1, n}). \tag{19.11}$$

Координаты $e_{i'}$ запишем, как обычно, в виде столбцов и получим матрицу

$$T=\begin{bmatrix}\tau_{1'}^1 & \tau_{2'}^1 \dots \tau_{n'}^1\\ \tau_{1'}^2 & \tau_{2'}^2 \dots \tau_{n'}^2\\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\\ \tau_{1'}^n & \tau_{2'}^n \dots \tau_{n'}^n\end{bmatrix}, \qquad (19.12)$$

которая называется *матрицей перехода* от нештрихованного базиса к штрихованному. Ясно, что $\det T \neq 0$, так как векторы $e_{l'}$ как векторы базиса являются линейно-независимыми и ранг матрицы T полный. Итак, *матрица перехода есть невырожденная матрица.*

Соотношения (19.11) могут быть записаны в матричной форме, если ввести в рассмотрение матрицы-строки $[e]' = [e_{1'}, e_{2'}, \dots, e_{n'}]$, $[e] = [e_1, e_2, \dots, e_n]$, составленные из векторов штрихованного и нештрихованного базисов. Тогда (19.11) эквивалентно следующему соотношению:

$$[e]' = [e]\,T. \qquad (19.13)$$

Поскольку T — невырожденная матрица, то всегда существует T^{-1}. Умножая (19.13) на T^{-1} справа и пользуясь равенствами $TT^{-1} = E$ и $[e]\,E = [e]$, получим

$$[e] = [e]'\,T^{-1}. \qquad (19.14)$$

Итак, *переход от штрихованного базиса к нештрихованному осуществляется с помощью обратной матрицы.*

Для любого $x \in \mathscr{L}_n$ мы имеем $x = \sum\limits_{k=1}^{n} \xi^k e_k$ в базисе $\{e_1, \dots, e_n\}$ и $x = \sum\limits_{l=1}^{n} \xi^{l'} e_{l'}$ в базисе $\{e_{1'}, \dots, e_{n'}\}$. Подставив $e_{l'}$ из (19.11), получим, что $x = \sum\limits_{l=1}^{n} \xi^{l'} \sum\limits_{k=1}^{n} \tau_{l'}^k e_k = \sum\limits_{k=1}^{n} \left(\sum\limits_{l=1}^{n} \tau_{l'}^k \xi^{l'} \right) e_k$. Поэтому

$$\sum_{k=1}^{n} \xi^k e_k = \sum_{k=1}^{n} \left(\sum_{l=1}^{n} \tau_{l'}^k \xi^{l'} \right) e_k .$$

В силу линейной независимости векторов базиса $\{e_1, \dots, e_n\}$ вытекает равенство

$$\xi^k = \sum_{l=1}^{n} \tau^k_{l'} \xi^{l'}, \qquad (19.15)$$

которое в матричной форме имеет вид

$$[x] = T[x]', \qquad (19.16)$$

где под $[x]$ и $[x]'$ подразумеваются матрицы-столбцы, составленные из координат ξ^k и $\xi^{k'}$ вектора x. Умножая (19.16) слева на T^{-1}, мы получим, как „новые" координаты вектора x выражаются через „старые":

$$[x]' = T^{-1}[x]. \qquad (19.17)$$

Таким образом, *формулы* (19.16), (19.17) *задают законы, связывающие между собой координаты вектора x в штрихованном и нештрихованном базисах.*

Пусть $A = [\alpha^i_k]$ — матрица оператора в базисе $\{e_1, \dots, e_n\}$. Через $A' = [\alpha^{i'}_{k'}]$ обозначим матрицу того же оператора в базисе $\{e_{1'}, \dots, e_{n'}\}$.

Заметив, что $Ae_k = \sum_{i=1}^{n} \alpha^i_k e_i$ и $Ae_{k'} = \sum_{m=1}^{n} \alpha^{m'}_{k'} e_{m'}$, подставив согласно (19.11) $e_{k'} = \sum_{s=1}^{n} \tau^s_{k'} e_s$, $e_{m'} = \sum_{i=1}^{n} \tau^i_{m'} e_i$ в левую и правую части второго равенства, получим

$$\sum_{s=1}^{n} \tau^s_{k'} Ae_s = \sum_{m=1}^{n} \alpha^{m'}_{k'} \sum_{i=1}^{n} \tau^i_{m'} e_i$$

или

$$\sum_{s=1}^{n} \tau^s_{k'} \sum_{i=1}^{n} \alpha^i_s e_i = \sum_{m=1}^{n} \alpha^{m'}_{k'} \sum_{i=1}^{n} \tau^i_{m'} e_i .$$

В силу линейной независимости векторов базиса $\{e_1, \dots, e_n\}$ имеем равенство

$$\sum_{s=1}^{n} \alpha^i_s \tau^s_{k'} = \sum_{m=1}^{n} \tau^i_{m'} \alpha^{m'}_{k'}, \qquad (19.18)$$

которое в матричной форме принимает вид

$$AT = TA', \qquad (19.19)$$

откуда следует, что

$$\text{(а) } A' = T^{-1}AT; \qquad \text{(б) } A = TA'T^{-1}. \qquad (19.20)$$

Формулы (19.20) *задают законы, согласно которым связаны матрицы оператора в штрихованном и нештрихованном базисах.*

Перейдем сейчас к очень важному для физических приложений вопросу об инвариантах линейных операторов. Прежде всего отметим, что *ранг A есть инвариант преобразования базисов*, так как ранг $A' =$ ранг A из-за умножения A на невырожденные матрицы.

Теорема 5. *Собственные значения оператора инвариантны относительно преобразования базисов.*

В самом деле, пусть λ_0 — собственное значение оператора и x — собственный вектор, отвечающий данному собственному значению. Тогда имеет место следующее матричное соотношение:

$$A[x] = \lambda_0 [x], \tag{19.21}$$

где A — матрица оператора в базисе $\{e_1, \ldots, e_n\}$. Пусть T—матрица перехода к штрихованному базису. Умножив (19.21) на T^{-1} слева и вставив между A и $[x]$ множитель $E = TT^{-1}$, получим

$$T^{-1} ATT^{-1}[x] = \lambda_0 T^{-1}[x]. \tag{19.22}$$

С учетом (19.17) и (19.20a) равенство (19.22) имеет вид

$$A'[x]' = \lambda_0 [x]'.$$

Собственное значение собственного вектора x в штрихованной системе координат осталось неизменным. Теорема доказана.

Отметим также, что матрица тождественного преобразования при переходе к новому базису не меняет своего вида и остается единичной, так как $E' = T^{-1}ET = T^{-1}T = E$.

Рассмотрим характеристическую матрицу $(A - \lambda E)$ для оператора $A - \lambda \mathcal{E}$. Согласно (19.20a) имеем: $(A - \lambda E)' = T^{-1}(A - \lambda E)\,T = T^{-1}AT - \lambda T^{-1}ET = A' - \lambda E$. Но

$$\det (A - \lambda E)' = \det T^{-1} \cdot \det (A - \lambda E) \cdot \det T = \\ = \frac{1}{\det T} \cdot \det (A - \lambda E) \cdot \det T = \det (A - \lambda E).$$

В силу равенства $(A - \lambda E)' = A' - \lambda E$ получим следующее равенство:

$$\det(A - \lambda E) = \det(A' - \lambda E), \tag{19.23}$$

выполняющееся при любом значении параметра λ, т. е. *характеристический полином оператора есть инвариант преобразования базисов.*

В (19.23) слева и справа стоят полиномы n-й степени

$$(-1)^n \lambda^n + I_1 \lambda^{n-1} + \ldots + I_{n-1} \lambda + I_n = (-1)^n \lambda^n + \\ + I_1' \lambda^{n-1} + \ldots + I_{n-1}' \lambda + I_n'.$$

Известно, что полиномы тождественно равны друг другу, если равны их коэффициенты при соответствующих степенях λ. Итак, *коэффициенты характеристического полинома суть инварианты относительно преобразования базисов.*

Каким образом они связаны с элементами матрицы оператора? Раскрывая определитель n-го порядка $\det(A - \lambda E)$ и собирая члены при соответствующих степенях λ, мы получим, что

$$I_1 = (-1)^{n-1} (\alpha_1^1 + \alpha_2^2 + \ldots + \alpha_n^n) = (-1)^{n-1} \operatorname{Sp} A, \tag{19.24}$$

где символом $\operatorname{Sp} A$ (либо $\operatorname{tr} A$) обозначают $\sum\limits_{i=1}^{n} \alpha_i^i$ и называют *следом* матрицы A (*шпуром* A).

Аналогично I_2 есть *сумма всех диагональных миноров второго порядка*, взятая со множителем $(-1)^{n-2}$, I_3 — сумма всех диагональных миноров третьего порядка, взятая со множителем $(-1)^{n-3}$ и т. д. Наконец, $I_n = \det A$, поскольку минор n-го порядка совпадает с определителем матрицы. Тот факт, что определитель матрицы оператора есть инвариант преобразования базисов, может быть получен и из (19.20а), поскольку $\det A' = \det T^{-1} \det A \det T = \det A$.

Лекция 20

ЛИНЕЙНЫЕ ФОРМЫ И СОПРЯЖЕННОЕ ЛИНЕЙНОЕ ПРОСТРАНСТВО. БИЛИНЕЙНЫЕ ФОРМЫ. ПРЕОБРАЗОВАНИЕ МАТРИЦЫ БИЛИНЕЙНОЙ ФОРМЫ ПРИ ИЗМЕНЕНИИ БАЗИСА И ЕЕ ИНВАРИАНТЫ. СИММЕТРИЧНЫЕ И АНТИСИММЕТРИЧНЫЕ БИЛИНЕЙНЫЕ ФОРМЫ

На числовое поле **K**, под которым мы подразумеваем **C** или **R**, если иметь в виду лишь операции сложения чисел и умножения их на числа из того же поля **K**, можно смотреть как на одномерное линейное пространство K_1 (комплексное в случае **C** и вещественное в случае **R**), векторами которого являются числа из множества **K**.

Определение. Линейное отображение $\varphi : \mathscr{L}_n \to K_1$ называется *линейной формой*, заданной на векторах $\mathscr{L}_n$, т. е. для любого $x \in \mathscr{L}_n$ сопоставлено число $\varphi(x) \in K_1$, причем $\varphi(\alpha x_1 + \beta x_2) = \alpha\varphi(x_1) + \beta\varphi(x_2)$ для любых $\alpha, \beta \in \mathbf{K}$.

Например, если в трехмерном евклидовом пространстве зафиксировать вектор $\vec{c}$ и рассмотреть скалярное произведение $(\vec{c}\,\vec{x})$, где $\vec{x}$ пробегает все множество векторов пространства, то мы получим линейную форму $\varphi(\vec{x}) = (\vec{c}\,\vec{x})$, так как $(\vec{c}(\alpha\vec{x}_1 + \beta\vec{x}_2)) = \alpha(\vec{c}\,\vec{x}_1) + \beta(\vec{c}\,\vec{x}_2) = \alpha\varphi(\vec{x}_1) + \beta\varphi(\vec{x}_2)$.

Пусть $\{e_1, \dots, e_n\}$ — базис в $\mathscr{L}_n$. Найдем матрицу отображения φ, вычислив φ на векторах базиса. Имеем $\varphi(e_k) = \varphi_k$ $(k = \overline{1, n})$. Матрица $[\varphi]$ имеет одну строку, поскольку K_1 одномерно. Поэтому

$$[\varphi] = [\varphi_1\varphi_2 \cdots \varphi_n]. \tag{20.1}$$

Элементы этой матрицы носят название *коэффициентов линейной формы*. Зная матрицу $[\varphi]$, легко найти значение формы на любом векторе $x \in \mathscr{L}_n$. В самом деле, если $x = \sum_{k=1}^{n} \xi^k e_k$, то

$$\varphi(x) = \varphi\left(\sum_{k=1}^{n} \xi^k e_k\right) = \sum_{k=1}^{n} \xi^k \varphi(e_k) = \sum_{k=1}^{n} \varphi_k \xi^k.$$

В матричной форме это выражение имеет простой вид

$$\varphi(x) = [\varphi][x], \tag{20.2}$$

где под $[x]$ подразумевается матрица-столбец из координат вектора x.

Мы знаем, что относительно операции сложения и умножения на число совокупность всех линейных форм (как линейных отображений) образует свое линейное пространство размерности $1 \cdot n = n$.

Определение. Линейное n-мерное пространство, векторами которого являются линейные формы, называется сопряженным линейным пространством (по отношению к пространству $\mathcal{L}_n$) и обозначается символом $\mathcal{L}_n^*$.

Базис в $\mathcal{L}_n^*$ будем обозначать той же буквой e, но, ставя индексы наверху $\{e^1, e^2, \ldots, e^n\}$, в отличие от базиса $\{e_1, \ldots, e_n\}$ в $\mathcal{L}_n$.

Всякая линейная форма φ как вектор из $\mathcal{L}_n^*$ раскладывается по векторам базиса $\{e^1, \ldots, e^n\}$: $\varphi = \sum_{k=1}^{n} \theta_k e^k$, где θ_k — координаты линейной формы φ в базисе $\{e^1, \ldots, e^n\}$.

Удобно выбрать базис в $\mathcal{L}_n^*$ таким образом, чтобы линейные формы e^k на векторах базиса $\mathcal{L}_n$ $\{e_1, \ldots, e_n\}$ принимали наиболее простые значения. А именно

$$e^k(e_i) = \delta_i^k = \begin{cases} 1, & i = k \\ 0, & i \neq k. \end{cases} \tag{20.3}$$

Такой базис в $\mathcal{L}_n^*$ называется *каноническим* или *взаимным* к базису $\{e_1, \ldots, e_n\}$ в $\mathcal{L}_n$.

Почему целесообразно выбирать в $\mathcal{L}_n^*$ канонический (взаимный) базис? Чтобы ответить на поставленный вопрос, находим

$$\varphi(x) = \sum_{k=1}^{n} \theta_k e^k(x) = \sum_{k=1}^{n} \theta_k e^k\left(\sum_{l=1}^{n} \xi^l e_l\right) = \sum_{l,\, k=1}^{n} \theta_k \xi^l e^k(e_l) =$$

$$= \sum_{l,\, k=1}^{n} \theta_k \xi^l \delta_l^k = \sum_{k=1}^{n} \theta_k \xi^k,$$

где θ_k—координаты формы. С другой стороны, ранее нами получено выражение $\varphi(x) = \sum_{k=1}^{n} \varphi_k \xi^k$, где φ_k—коэффициенты линейной формы. Итак, если базис $\{e^1, \ldots, e^n\}$ канонический, то $\sum_{k=1}^{n} \theta_k \xi^k = \sum_{k=1}^{n} \varphi_k \xi^k$ для *любого* вектора x. Это означает, что $\theta_k = \varphi_k$ $(k = \overline{1, n})$ — *во взаимных базисах* $\{e_1, \ldots, e_n\}$, $\{e^1, \ldots, e^n\}$ $(e^k(e_i) = \delta_i^k)$ *координаты линейной формы и коэффициенты линейной формы совпадают*. В дальнейшем мы будем предполагать, что базис $\{e_1, \ldots, e_n\}$ в $\mathscr{L}_n$ и базис $\{e^1, \ldots, e^n\}$ в $\mathscr{L}_n^*$ взаимные.

Как меняются коэффициенты линейной формы, если от базиса $\{e_1, \ldots, e_n\}$ в $\mathscr{L}_n$ перейти к базису $\{e_{1'}, \ldots, e_{n'}\}$ с помощью матрицы перехода T? Пусть $e_{k'} = \sum_{i=1}^{n} \tau_{k'}^i e_i$. Тогда $\varphi_{k'} = \varphi(e_{k'}) = \varphi\left(\sum_{i=1}^{n} \tau_{k'}^i e_i\right) = \sum_{i=1}^{n} \tau_{k'}^i \varphi(e_i) = \sum_{i=1}^{n} \varphi_i \tau_{k'}^i$. В матричной форме, если через $[\varphi]'$ обозначить матрицу-строку линейной формы в штрихованном базисе, последнее соотношение может быть записано в виде

$$[\varphi]' = [\varphi]\, T. \qquad (20.4)$$

Умножая справа на T^{-1}, получим выражение для матрицы-строки $[\varphi]$ в нештрихованном базисе

$$[\varphi] = [\varphi]'\, T^{-1}. \qquad (20.5)$$

Мы видим, что *закон преобразования коэффициентов линейной формы совпадает по форме с законами преобразования базисов в пространстве* $\mathscr{L}_n$.

Перейдем к рассмотрению линейных скалярных функций, зависящих не от одного векторного аргумента, как это было в случае линейных форм, а от двух векторных аргументов. С этой целью будем рассматривать отображение (функцию) $B: L \times L \to \mathbf{K}$, где $L \times L$ — прямое произведение множеств векторов пространства $\mathscr{L}_n$.

Определение. Числовая функция $B(x, y)$ называется *билинейной формой*, если для любых $x, y, z \in \mathscr{L}_n$ и любого $\lambda \in \mathbf{K}$ выполнено: (1) $B(x + z, y) = B(x, y) + B(z, y)$; (2) $B(\lambda x, y) = \lambda B(x, y)$; (3) $B(x, y + z) = B(x, y) + B(x, z)$; (4) $B(x, \lambda y) = \lambda B(x, y)$, т. е. числовая функция линейна как по первому, так и по второму аргументу.

Из условий (1)—(4) следует

$$B\left(\sum_{i=1}^{p}\lambda_i x_i,\ \sum_{k=1}^{q}\mu_k y_k\right)=\sum_{i=1}^{p}\ \sum_{k=1}^{q}\lambda_i\mu_k B(x_i,\ y_k)$$

и поэтому, если $\{e_1, \ldots, e_n\}$ — базис в $\mathscr{L}_n$, то

$$B(x,\ y)=B\left(\sum_{i=1}^{n}\xi^i e_i,\ \sum_{k=1}^{n}\eta^k e_k\right)=\sum_{i,\ k=1}^{n}\xi^i\eta^k B(e_i,\ e_k).$$

Обозначим $B(e_i,\ e_k)=\beta_{ik}$, и тогда

$$B(x,\ y)=\sum_{i,\ k=1}^{n}\beta_{ik}\xi^i\eta^k, \tag{20.6}$$

где числа β_{ik} называются *коэффициентами билинейной формы*. Из них мы можем построить матрицу (первый индекс номер строки, второй — номер столбца)

$$B=\begin{bmatrix}\beta_{11} & \beta_{12} \cdots \beta_{1n}\\ \beta_{21} & \beta_{22} \cdots \beta_{2n}\\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\ \beta_{n1} & \beta_{n2} \cdots \beta_{nn}\end{bmatrix}, \tag{20.7}$$

которая носит название *матрицы билинейной формы относительно данного базиса.*

Как преобразуется матрица билинейной формы, если мы перейдем к штрихованному базису? Какие можно указать инварианты для билинейных форм?

Пусть $e_{k'}=\sum_{i=1}^{n}\tau_{k'}^{i}e_i$ — векторы нового базиса. Тогда

$$\beta_{i'k'}=B(e_{i'},\ e_{k'})=B\left(\sum_{p=1}^{n}\tau_{i'}^{p}e_p,\ \sum_{q=1}^{n}\tau_{k'}^{q}e_q\right)=\sum_{p,\ q=1}^{n}\tau_{i'}^{p}\tau_{k'}^{q}\beta_{pq}.$$

Чтобы записать полученное равенство в матричной форме, запишем последнюю двойную сумму в раздельном виде

$$\beta_{i'k'}=\sum_{p=1}^{n}\tau_{i'}^{p}\left(\sum_{q=1}^{n}\beta_{pq}\tau_{k'}^{q}\right). \tag{20.8}$$

Внутренняя сумма дает нам элементы $\gamma_{pk'}$ матрицы BT.

Из равенства $\beta_{i'k'}=\sum_{p=1}^{n}\tau_{i'}^{p}\gamma_{pk'}$ следует, что перемножаются соответствующие элементы столбцов матрицы T

и BT и складываются. Правило умножения матриц (строка на столбец) нарушено. Чтобы правило умножения матриц не нарушалось, рассмотрим транспонированную T^{tr}, полученную из T заменой строк на столбцы и столбцов на строки. Тогда $\sum_{p=1}^{n} \tau_{i'}^{p} \gamma_{pk'}$ будет давать элементы матрицы, полученной от перемножения T^{tr} и BT. Таким образом, равенство (20.8) в матричном виде имеет вид

$$B' = T^{tr} BT. \tag{20.9}$$

Поскольку матрица B слева и справа умножается на невырожденные матрицы, то заключаем: *ранг матрицы билинейной формы есть инвариант преобразования базисов*. В силу того, что $\det T^{tr} = \det T$ (первое свойство определителей), имеем

$$\det B' = (\det T)^2 \det B, \tag{20.10}$$

откуда заключаем, что *знак детерминанта матрицы билинейной формы есть инвариант преобразования базисов.*

В приложениях чаще всего находят себя симметричные и антисимметричные (кососимметричные) билинейные формы.

Прежде чем переходить к их рассмотрению, дадим определение симметрической и антисимметрической матрицы.

Определение 1. Матрица B называется *симметрической*, если при транспонировании она не меняется, т. е. $B^{tr} = B$.

Определение 2. Матрица B называется *антисимметрической* (*кососимметрической*), если при транспонировании меняет знак на противоположный, т. е. $B^{tr} = -B$.

Если элементы матрицы B обозначить через β_{ik}, то условие симметричности означает следующее равенство между элементами матрицы:

$$\beta_{ik} = \beta_{ki} \quad (i,\ k = \overline{1, n}), \tag{20.11}$$

а условие антисимметричности означает, что

$$\beta_{ik} = -\beta_{ki} \quad (i, k = \overline{1, n}). \tag{20.12}$$

Условия (20.12) при $i=k$ дают, что $\beta_{ii}=-\beta_{ii}$ и все элементы по главной диагонали матрицы нулевые, а при $i \neq k$ элементы совпадают по модулю, но разнятся знаком.

Переходим к рассмотрению симметричных и антисимметричных билинейных форм.

Определение 3. Билинейная форма $B(x, y)$ называется *симметричной*, если для любых $x, y \in \mathscr{L}_n$ выполнено $B(x, y) = B(y, x)$.

Пусть в $\mathscr{L}_n$ задан базис $\{e_1, \ldots, e_n\}$. По определению 3 имеем, что $B(e_i, e_k) = B(e_k, e_i)$. Это означает, что коэффициенты билинейной формы β_{ik} удовлетворяют условию $\beta_{ik} = \beta_{ki}$, т. е. *матрица симметричной билинейной формы необходимо симметрическая*. Обратно, пусть у билинейной формы $B(x, y)$ матрица симметрическая, т. е. $\beta_{ik} = \beta_{ki}$. Тогда

$$B(x, y) = \sum_{i,\, k=1}^{n} \beta_{ik} \xi^i \eta^k = \sum_{k,\, i=1}^{n} \beta_{ki} \eta^k \xi^i = B(y, x)$$

для любых x и y из $\mathscr{L}_n$. *Билинейная форма с симметрической матрицей непременно симметричная билинейная форма.*

Определение 4. Билинейная форма $B(x, y)$ называется *антисимметричной*, если для любых $x, y \in \mathscr{L}_n: B(x, y) = -B(y, x)$.

Если в $\mathscr{L}_n$ задан базис $\{e_1, \ldots, e_n\}$, то согласно определению 4 $B(e_i, e_k) = -B(e_k, e_i)$ и, следовательно, $\beta_{ik} = -\beta_{ki}$. Итак, *матрица антисимметричной билинейной формы с необходимостью является антисимметрической матрицей*. Обратно, если у билинейной формы $B(x, y)$ матрица B антисимметрическая, то форма $B(x, y)$ является антисимметричной, так как

$$B(x, y) = \sum_{i,\, k=1}^{n} \beta_{ik} \xi^i \eta^k = \sum_{k,\, i=1}^{n} (-\beta_{ki}) \eta^k \xi^i = -B(y, x).$$

Поскольку определения симметричности и антисимметричности не связаны с каким-либо базисом и носят инвариантный характер, то условия симметричности и антисимметричности матриц симметричной и антисимметричной билинейных форм носят также инвариантный характер и не зависят от выбора базиса в $\mathscr{L}_n$.

Вернемся к рассмотрению общих билинейных форм $B(x, y)$. Построим билинейные формы $B_1(x, y)$, $B_2(x, y)$ по правилам

$$\left.\begin{aligned} &(а)\ B_1(x, y) \overset{\text{def}}{=} \frac{1}{2}(B(x, y)+B(y, x)) \\ &(б)\ B_2(x, y) \overset{\text{def}}{=} \frac{1}{2}(B(x, y)-B(y, x)). \end{aligned}\right\}. \tag{20.13}$$

Совершенно очевидно, что $B_1(x, y)$ — симметричная билинейная форма, так как $B_1(y, x)=\frac{1}{2}(B(y, x)+ + B(x, y))=B_1(x, y)$, а $B_2(x, y)$ — антисимметричная билинейная форма, поскольку $B_2(y, x)=\frac{1}{2}(B(y, x)- - B(x, y))=-B_2(x, y)$.

Если сейчас сложить оба равенства из (20.13), то получим

$$B(x, y)=B_1(x, y)+B_2(x, y). \tag{20.14}$$

Итак, *любая билинейная форма может быть всегда представлена как сумма некоторой симметричной и некоторой антисимметричной билинейных форм.*

Пример. Задана билинейная форма $B(x, y)= =2\xi^1\eta^1-\xi^1\eta^2-2\xi^1\eta^3+3\xi^2\eta^1-2\xi^2\eta^2-5\xi^2\eta^3+\xi^3\eta^1+\xi^3\eta^2+ +\xi^3\eta^3$. Представить ее в виде суммы симметричной и антисимметричной билинейных форм.

Решение. Составляем матрицу билинейной формы, а затем по правилу (20.13) построим матрицы форм $B_1(x, y)$ и $B_2(x, y)$. Итак,

$$B=\begin{bmatrix} 2 & -1 & -2 \\ 3 & -2 & -5 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\left.\begin{aligned} &\overset{(1)}{\beta}_{11}=2,\ \overset{(1)}{\beta}_{12}=\frac{1}{2}(-1+3)=1, \\ &\overset{(1)}{\beta}_{13}=\frac{1}{2}(-2+1)=-\frac{1}{2},\ \overset{(1)}{\beta}_{22}=-2, \\ &\overset{(1)}{\beta}_{23}=\frac{1}{2}(-5+1)=-2,\ \overset{(1)}{\beta}_{33}=1,\ \overset{(2)}{\beta}_{11}=\overset{(2)}{\beta}_{22}= \\ &=\overset{(2)}{\beta}_{33}=0,\ \overset{(2)}{\beta}_{12}=\frac{1}{2}(-1-3)=-2, \\ &\overset{(2)}{\beta}_{13}=\frac{1}{2}(-2-1)=-\frac{3}{2},\ \overset{(1)}{\beta}_{23}=\frac{1}{2}(-5-1)=-3 \end{aligned}\right\}.$$

Поэтому матрицы симметричной и антисимметричной форм имеют вид:

$$B_1=\begin{bmatrix} 2 & 1 & -\frac{1}{2} \\ 1 & -2 & -2 \\ -\frac{1}{2} & -2 & 1 \end{bmatrix}, \quad B_2=\begin{bmatrix} 0 & -2 & -\frac{3}{2} \\ 2 & 0 & -3 \\ \frac{3}{2} & 3 & 0 \end{bmatrix}.$$

(сами билинейные формы выписывать не будем).

Проверка решения проста. Сложив B_1 и B_2, мы получаем B.

Лекция 21

КВАДРАТИЧНЫЕ ФОРМЫ. ПРИВЕДЕНИЕ КВАДРАТИЧНОЙ ФОРМЫ К КАНОНИЧЕСКОМУ ВИДУ МЕТОДОМ ЛАГРАНЖА. ТЕОРЕМА ИНЕРЦИИ КВАДРАТИЧНЫХ ФОРМ. ПОЛОЖИТЕЛЬНО ОПРЕДЕЛЕННЫЕ КВАДРАТИЧНЫЕ ФОРМЫ

Определение. Квадратичной формой $B(x, x)$ называется числовая функция, полученная из билинейной формы $B(x, y)$ путем замены аргумента y на аргумент x.

Для установления вида квадратичной формы воспользуемся соотношениями (20.14) и заменим в них аргумент y на x. В этом случае $B_2(x, x)\equiv 0$ (см. (20.13б)) и $B(x, x)=B_1(x, x)$.

Итак, *всякая квадратичная форма, построенная на основе билинейной формы* $B(x, y)$, *совпадает с квадратичной формой, построенной на основе симметричной части* $B_1(x, y)$ *общей билинейной формы* $B(x, y)$.

В силу установленного факта матрица всякой квадратичной формы (понимаемая как матрица соответствующей симметричной билинейной формы) есть симметрическая матрица

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} \dots \beta_{1n} \\ \beta_{12} & \beta_{22} \dots \beta_{2n} \\ \cdot \cdot \cdot & \cdot \cdot \cdot \cdot \cdot \\ \beta_{1n} & \beta_{2n} \dots \beta_{nn} \end{bmatrix},$$

а сама квадратичная форма в некотором базисе имеет вид

$$B(x, x) = \beta_{11}(\xi^1)^2 + 2\beta_{12}\xi^1\xi^2 + \dots + 2\beta_{1n}\xi^1\xi^n +$$
$$+ \beta_{22}(\xi^2)^2 + 2\beta_{23}\xi^2\xi^3 + \dots + 2\beta_{n-1n}\xi^{n-1}\xi^n + \beta_{nn}(\xi^n)^2. \quad (21.1)$$

Поскольку матрица квадратичной формы совпадает с матрицей симметричной билинейной формы, то, как и для любых билинейных форм, имеем, что (1) *ранг матрицы квадратичной формы* (называемый просто *рангом квадратичной формы*) *есть инвариант преобразования базисов*, (2) *знак детерминанта матрицы квадратичной формы сохраняется при преобразованиях базисов*, (3) *при переходе от одного базиса к другому закон преобразования матрицы квадратичной формы имеет вид* $B' = T^{tr}BT$.

Отметим, что если ранг квадратичной формы полный (ранг $B = n$), то квадратичная форма называется *невырожденной*. Их рассмотрением мы займемся ниже, а сейчас зададимся вопросом, нельзя ли перейти в $\mathscr{L}_n$ к такому базису, в котором квадратичная форма представляет сумму квадратов новых координат, умноженных на некоторые коэффициенты?

Если такое возможно, то базис называется *каноническим* и матрица квадратичной формы в нем будет иметь диагональный вид.

Теорема 1. *В пространстве $\mathscr{L}_n$ всегда существует канонический базис $\{e_{1'}, \dots, e_{n'}\}$, в котором квадратичная форма примет вид*

$$B(x, x) = \alpha_1(\xi^{1'})^2 + \alpha_2(\xi^{2'})^2 + \dots + \alpha_n(\xi^{n'})^2, \quad (21.2)$$

где $\alpha_1, \dots, \alpha_n$ — фиксированные числа (некоторые из них, может быть, и равные нулю).

Для доказательства применим *метод Лагранжа*. Предварительно отметим, что если $\beta_{11} = 0$ и какое-либо из чисел $\beta_{1s} \neq 0$, то, проведя замену координат по правилу

$$\xi^1 = \eta^1 + \eta^s, \ \xi^2 = \eta^2, \dots, \xi^{s-1} = \eta^{s-1}, \ \xi^s = \eta^1 - \eta^s,$$
$$\xi^{s+1} = \eta^{s+1}, \dots, \xi^n = \eta^n,$$

для слагаемого $2\beta_{1s}\xi^1\xi^s$ в новом базисе получим

$$2\beta_{1s}(\eta^1+\eta^s)(\eta^1-\eta^s)=2\beta_{1s}(\eta^1)^2-2\beta_{1s}(\eta^s)^2,$$

т. е. в новом базисе $\beta_{1'1'}=2\beta_{1s}\neq 0$. Поэтому, не ограничивая общности рассуждений, применяем метод Лагранжа к квадратичной форме, где $\beta_{11}\neq 0$

$$B(x, x)=\beta_{11}(\xi^1)^2+2\beta_{12}\xi^1\xi^2+\ldots+2\beta_{1n}\xi^1\xi^n+\sum_{i, k=2}^{n}\beta_{ik}\xi^i\xi^k. \tag{21.3}$$

Запишем (21.3) в следующем виде:

$$B(x, x)=\frac{1}{\beta_{11}}(\beta_{11}\xi^1+\beta_{12}\xi^2+\ldots+\beta_{1n}\xi^n)^2-\frac{1}{\beta_{11}}(\beta_{12}^2(\xi^2)^2+$$
$$+\ldots+\beta_{1n}^2(\xi^n)^2+2\beta_{12}\beta_{13}\xi^2\xi^3+\ldots+2\beta_{1n-1}\beta_{1n}\xi^{n-1}\xi^n)+$$
$$+\sum_{i, k=2}^{n}\beta_{ik}\xi^i\xi^k=\frac{1}{\beta_{11}}(\beta_{11}\xi^1+\beta_{12}\xi^2+\ldots+\beta_{1n}\xi^n)^2+\widetilde{B}(x, x),$$

где $\widetilde{B}(x, x)$ — квадратичная форма, в которую входят лишь переменные $\xi^2, \ldots, \xi^n$.

Делаем замену координат

$$\eta^1=\beta_{11}\xi^1+\beta_{12}\xi^2+\ldots+\beta_{1n}\xi^n,\ \eta^2=\xi^2, \ldots, \eta^n=\xi^n, \tag{21.4}$$

после которой $B(x, x)=\frac{1}{\beta_{11}}(\eta^1)^2+\widetilde{B}(x, x)$, где $\widetilde{B}(x, x)=$ $=\sum_{i, k=2}^{n}\widetilde{\beta}_{ik}\eta^i\eta^k$. Предложенный алгоритм применяем к $\widetilde{B}(x, x)$ и после конечного числа шагов приходим к (21.2). Чтобы найти матрицу перехода от начальных координат $\xi^1, \ldots, \xi^n$ к каноническим $\xi^{1'}, \ldots, \xi^{n'}$, достаточно перемножать матрицы промежуточных переходов (вспомним, что если $[\eta]=T_1^{-1}[x]$ и $[\zeta]=T_2^{-1}[\eta]$, то $[\zeta]=T_2^{-1}T_1^{-1}[x]$ и т. д.).

Пример. Методом Лагранжа привести к каноническому виду квадратичную форму $B(x, x)=x_1^2+$ $+4x_1x_2+x_2^2+3x_3x_4$.

Решение. Первый шаг: $B(x, x)=(x_1+2x_2)^2-$ $-3x_2^2+3x_3x_4$. После замены координат $y_1=x_1+2x_2$, $y_2=x_2$, $y_3=x_3-x_4$, $y_4=x_3+x_4$ (эквивалентно относительно старых координат $x_1 = y_1 - 2y_2$, $x_2 = y_2$, $x_3 = \frac{1}{2}(y_3 + y_4)$, $x_4 = \frac{1}{2}(y_4 - y_3)$) имеем $B(x, x) = y_1^2 - 3y_2^2 - \frac{3}{4}y_3^2 + \frac{3}{4}y_4^2$. Матрица перехода T от первоначального базиса к каноническому (помним, что $[x] = T[y]$) имеет вид

$$T = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Все, что излагалось до сих пор относительно квадратичных форм, имеет место как для вещественных, так и для комплексных квадратичных форм. Дальнейшее изучение квадратичных форм мы продолжаем только для *вещественных* линейных пространств и для квадратичной формы $B(x, x)$ коэффициенты β_{ik} и координаты ξ^k с необходимостью вещественные числа.

Пусть квадратичная форма приведена к каноническому виду (21.2), где $\alpha_1, \ldots, \alpha_n$ — вещественные числа. Поскольку ранг квадратичной формы инвариант преобразования базисов, то среди чисел $\alpha_1, \ldots, \alpha_n$ имеется ровно r чисел, отличных от нуля, где $r =$ ранг B, а остальные равны нулю. Если среди этих r чисел s чисел положительны, а остальные $(r - s)$ отрицательны, то, проведя соответствующую перенумерацию переменных, приведем $B(x, x)$ к виду

$$B(x, x) = \lambda_1(\eta^1)^2 + \ldots + \lambda_s(\eta^s)^2 - \lambda_{s+1}(\eta^{s+1})^2 - \\ - \ldots - \lambda_r(\eta^r)^2, \qquad (21.5)$$

гве $\lambda_1, \ldots, \lambda_r$ — положительные числа.

Ясно, что после преобразования координат

$$\eta^{k'} = \sqrt{\lambda_k}\,\eta^k, \quad \eta^{t'} = \eta^t \quad (k = \overline{1, r}; \ t = \overline{r+1, n})$$

квадратичная форма примет наиболее простой вид

$$B(x, x) = (\eta^{1'})^2 + \ldots + (\eta^{s'})^2 - (\eta^{s+1'})^2 - \ldots - (\eta^{r'})^2,$$

который и называется *каноническим для вещественных квадратичных форм*, рассматриваемых в пространстве $\mathcal{L}_n$, при этом число s — *положительным индексом инерции*, $(r-s)$—*отрицательным индексом инерции*, разность $s-(r-s)$ при полном ранге $r=n$— *сигнатурой квадратичной формы*.

Легко заметить, что ни канонический базис, ни канонический вид квадратичной формы не определены однозначно, так как числа $\alpha_1, \ldots, \alpha_n$ в (21.2) после различных линейных преобразований, сохраняющих диагональную форму, принимают различные значения. Ясно, что число r (как ранг матрицы квадратичной формы) при изменении канонического базиса остается неизменным. А что можно сказать о числах s, $r-s$?

Теорема 2 (*инерции квадратичных форм*). *Положительный индекс инерции s и отрицательный индекс инерции $r-s$ являются инвариантами квадратичных форм.*

Доказательство. Пусть квадратичная форма в базисе $\{e_1, \ldots, e_n\}$ имеет вид $B(x, x) = \sum\limits_{i, k=1}^{n} \beta_{ik} \xi^i \xi^k$.

Пусть в каноническом базисе $\{e_{1'}, \ldots, e_{n'}\}$

$$B(x, x) = (\eta^1)^2 + \ldots + (\eta^s)^2 - (\eta^{s+1})^2 - \ldots - (\eta^r)^2, \quad (21.6)$$

а в каноническом базисе $\{e_{1''}, \ldots, e_{n''})$

$$B(x, x) = (\zeta^1)^2 + \ldots + (\zeta^p)^2 - (\zeta^{p+1})^2 - \ldots - (\zeta^r)^2. \quad (21.7)$$

Пусть соответствующие преобразования координат следующие:

$$\text{(а)}\ \eta^k = \sum_{i=1}^{n} \theta_i^k \xi^i; \quad \text{(б)}\ \zeta^k = \sum_{i=1}^{n} \sigma_i^k \xi^i. \quad (21.8)$$

Вычитая из (21.6) соотношения (21.7), для любого $x \in \mathcal{L}_n$ получим

$$(\eta^1)^2 + \ldots + (\eta^s)^2 + (\zeta^{p+1})^2 + \ldots + (\zeta^r)^2 = (\zeta^1)^2 + \ldots + (\zeta^p)^2 + \\ + (\eta^{s+1})^2 + \ldots + (\eta^r)^2. \quad (21.9)$$

Предположим, что $s < p$ и рассмотрим вектор x, координаты которого в базисе $\{e_{1'}, \ldots, e_{n'}\}$ удовлетворяют условию $\eta^1 = 0, \eta^2 = 0, \ldots, \eta^s = 0$, а в базисе $\{e_{1''}, \ldots, e_{n''}\}$ условию $\zeta^{p+1} = 0, \ldots, \zeta^n = 0$. Такой вектор x существует, и он не нулевой, так как наложенные условия для координат вектора x в базисе $\{e_1, \ldots, e_n\}$ эквивалентны однородной системе уравнений

$$\begin{cases} \sum_{i=1}^{n} \theta_i^k \xi^i = 0 \quad (k=\overline{1, s}) \\ \sum_{i=1}^{n} \sigma_i^m \xi^i = 0 \quad (m=\overline{p+1, n}). \end{cases} \tag{21.10}$$

Число уравнений в (21.10), равное $s+(n-p)=n-(p-s)$, меньше n из-за нашего предположения $s<p$, и поэтому система всегда имеет нетривиальное решение. Выбрав такой вектор x (для этого достаточно взять любое частное решение системы (21.10)), мы получим из (21.9) для его канонических координат соотношение

$$(\zeta^1)^2+\ldots+(\zeta^p)^2+(\eta^{s+1})^2+\ldots+(\eta^r)^2=0,$$

откуда следует, что $\zeta^1=0, \ldots, \zeta^p=0, \eta^{s+1}=0, \ldots, \eta^r=0$, т. е. вектор x таков, что в базисе $\{e_{1''}, \ldots, e_{n''}\}$ его координаты $\zeta^1=0, \ldots, \zeta^p=0$, $\zeta^{p+1}=0, \ldots, \zeta^n=0$. Получается, что вектор x нулевой. Пришли к противоречию. Остается предположить, что $s \geqslant p$. Однако в силу симметрии чисел s и p предположение $p<s$ также приводит к противоречию. Следовательно, $s=p$. Теорема доказана.

Сейчас более подробно рассмотрим положительно определенные квадратичные формы.

Определение. Квадратичная форма $B(x, x)$ называется *положительно определенной*, если для любого $x \neq 0$ $B(x, x)>0$.

Теорема 3. *Для того, чтобы квадратичная форма была положительно определенной, необходимо и достаточно*, чтобы положительный *индекс инерции был равен n.*

Необходимость. Дано, что $B(x, x)>0$ для любого $x \neq 0$. Предположим, что $s<n$. Приведем $B(x, x)$ к каноническому виду

$$B(x, x)=(\xi^{1'})^2+\ldots+(\xi^{s'})^2-(\xi^{s+1'})^2-\ldots-(\xi^{r'})^2. \tag{21.11}$$

Если $r<n$, то, взяв вектор x_0, у которого координаты в каноническом базисе имеют вид $\xi^{1'}=0, \ldots, \xi^{n-1'}=0$, $\xi^{n'}=1$, из (21.11) получим $B(x_0, x_0)=0$. Имеем противоречие. Итак, $r=n$, т. е. *всякая положительно определенная квадратичная форма является невырожденной*. Пусть $s<r=n$. Рассмотрим тот же век-

тор x_0, у которого координаты в каноническом базисе имеют вид $\xi^{1'}=0, \ldots, \xi^{n-1'}=0$, $\xi^{n'}=1$. Из (21.11), где $r=n$, получим, что $B(x_0, x_0) = -1 < 0$. Противоречие. Необходимость доказана.

Достаточность. Дано, что $s=n$. Это означает, что в каноническом базисе $B(x, x)$ имеет вид

$$B(x, x) = (\xi^{1'})^2 + (\xi^{2'})^2 + \ldots + (\xi^{n'})^2, \tag{21.12}$$

откуда следует, что для любого $x \neq 0$ $B(x, x) > 0$. Имеем положительно определенную квадратичную форму.

Отметим, что соответствующая положительно определенной квадратичной форме симметричная билинейная форма в каноническом базисе имеет вид

$$B(x, y) = \xi^{1'}\eta^{1'} + \xi^{2'}\eta^{2'} + \ldots + \xi^{n'}\eta^{n'}. \tag{21.13}$$

Ее матрица является единичной.

Если квадратичная форма записана в произвольном (не каноническом) базисе, то как узнать, является ли она положительно определенной или не является?

Теорема 4 (*Сильвестр*). *Квадратичная форма* $B(x, x) = \sum_{i,k=1}^{n} \beta_{ik}\xi^i\xi^k$ *положительно определённая тогда и только тогда*, когда *главные миноры матрицы* B строго положительны:

$$\beta_{11} > 0, \quad \begin{vmatrix} \beta_{11} & \beta_{12} \\ \beta_{12} & \beta_{22} \end{vmatrix} > 0, \ldots, \begin{vmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1n} \\ \beta_{12} & \beta_{22} & \ldots & \beta_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \beta_{1n} & \beta_{2n} & \ldots & \beta_{nn} \end{vmatrix} > 0. \tag{21.14}$$

Для доказательства представим квадратичную форму в виде

$$B(x, x) = \widetilde{B}(x, x) + 2\left(\sum_{i=1}^{n}\beta_{in}\xi^i\right)\xi^n + \beta_{nn}(\xi^n)^2, \tag{21.15}$$

где $\widetilde{B}(x, x)$ — квадратичная форма, содержащая переменные $\xi^1, \ldots, \xi^{n-1}$. Главные миноры матрицы $\widetilde{B}$ совпадают с минорами матрицы B, если исключить из рассмотрения последний, совпадающий с $\det B$. Далее будем применять метод индукции. При $n=1$ теорема верна (этот факт очевиден). Предположим, что теорема верна для квадратичной формы от $(n-1)$ переменных. Необходимо доказать теорему для случая n переменных.

Необходимость. Дано, что $B(x, x)$ — положительно определенная квадратичная форма. Требуется доказать неравенства (21.14). Из соотношения (21.15) тотчас следует, что $\widetilde{B}(x, x)$ также положительно определенная квадратичная форма, так как иначе существует такой вектор x_0, у которого первые $n-1$ координат отличны от нуля, а $\xi^n = 0$, что $B(x_0, x_0) = \widetilde{B}(x_0, x_0) \leqslant 0$. Следовательно, в силу индуктивного рассмотрения все главные миноры в (21.14), кроме последнего, строго положительны. Докажем, что и последний минор, равный $\det B$, строго положителен. Поскольку $B(x, x)$ положительно определенная квадратичная форма, то в каноническом базисе ее матрица $B' = E$ (см. (21.12)) и $\det B' = 1$. Но мы знаем, что $\det B' = (\det T)^2 \det B$. Поэтому $\det B = \frac{1}{(\det T)^2} > 0$. Необходимость доказана.

Достаточность. Дано, что выполнены условия (21.14). Доказать, что $B(x, x)$ положительно определенная квадратичная форма. Поскольку выполнены условия (21.14), то согласно индуктивному подходу квадратичная форма $\widetilde{B}(x, x)$ положительно определенная. Не меняя последней переменной $(\xi^n = \eta^n)$ за счет невырожденных линейных преобразований, связывающих канонические координаты $\eta^1, \ldots, \eta^{n-1}$ и координаты $\xi^1, \ldots, \xi^{n-1}$, приведем $\widetilde{B}(x, x)$ к каноническому виду. В результате получим

$$B(x, x) = \sum_{i=1}^{n-1} (\eta^i)^2 + 2\left(\sum_{i=1}^{n-1} \gamma_{in} \eta^i\right) \eta^n + \beta_{nn} (\eta^n)^2 =$$

$$= \sum_{i=1}^{n-1} (\eta^i + \gamma_{in} \eta^n)^2 - \left(\sum_{i=1}^{n-1} \gamma_{in}^2\right) (\eta^n)^2 + \beta_{nn} (\eta^n)^2.$$

Положив $\xi^{i'} = \eta^i + \gamma_{in} \eta^n \ (i = \overline{1, n-1})$,

$$\xi^{n'} = \eta^n, \quad \beta_{nn} - \sum_{i=1}^{n-1} \gamma_{in}^2 = \alpha,$$

имеем

$$B(x, x) = \sum_{i=1}^{n-1} (\xi^{i'})^2 + \alpha (\xi^{n'})^2. \tag{21.16}$$

Из (21.16) следует, что $\det B' = \alpha$. На основании известной нам формулы $\det B' = (\det T)^2 \det B$ и того факта, что $\det B > 0$ (условия (21.14) выполнены), выводим $\alpha > 0$. Следовательно, $B(x, x)$ по (21.16) строго положительна для любого $x \neq 0$. Достаточность также доказана.

В заключение лекции отметим, что если бы мы рассматривали квадратичные формы в комплексных линейных пространствах, то не получили бы таких дополнительных свойств квадратичных форм, как в вещественном случае, потому что за счет комплексных невырожденных линейных преобразований из (21.2) следует единственный канонический вид квадратичной формы

$$B(x, x) = (\eta^1)^2 + (\eta^2)^2 + \ldots + (\eta^r)^2, \tag{21.17}$$

где $r < n$ для вырожденных квадратичных форм и $r = n$ для невырожденных квадратичных форм.

Свойства, установленные выше для вещественных квадратичных форм, позволяют нам перейти в следующих лекциях к рассмотрению вещественных евклидовых пространств.

Лекция 22

ЛИНЕЙНЫЕ ВЕКТОРНО-ТОЧЕЧНЫЕ (АФФИННЫЕ) ПРОСТРАНСТВА. ПРЕОБРАЗОВАНИЕ АФФИННОЙ СИСТЕМЫ КООРДИНАТ. СОБСТВЕННО ЕВКЛИДОВЫ И ПСЕВДОЕВКЛИДОВЫ ПРОСТРАНСТВА. ПРЕОБРАЗОВАНИЕ ПРЯМОУГОЛЬНОЙ СИСТЕМЫ КООРДИНАТ

Нашей дальнейшей целью является изучение линейных операторов в евклидовом пространстве, и мы могли бы ограничиться рассмотрением только векторных евклидовых пространств, но физические приложения требуют рассмотрения векторно-точечных собственно евклидовых и псевдоевклидовых пространств. Поэтому мы начинаем с определения аффинных n-мерных пространств, на основе которых строится понятие евклидовых пространств.

Пусть $\mathcal{L}_n = (L, \mathbf{K})$ — n-мерное линейное векторное пространство, а M — некоторое множество, элементы которого назовем точками.

Определение. Пара (L, M) называется n-мерным *аффинным* пространством, если выполнено: (I) для L имеют место 8 аксиом линейного векторного пространства (т. е. $(L, \mathbf{K})$ — n-мерное линейное пространство); (II) для множества M имеют место аксиомы: (II_1) любой упорядоченной паре точек (A, B) соответствует один и только один вектор из $\mathcal{L}_n$, обозначаемый символом $\overrightarrow{AB}$; (II_2) для любой точки $A \in M$ и любого вектора $x \in \mathcal{L}_n$ существует такая единственная точка B, что $\overrightarrow{AB} = x$; (II_3) если $\overrightarrow{AB} = \overrightarrow{CD}$, то $\overrightarrow{AC} = \overrightarrow{BD}$ (аксиома параллелограмма).

Аффинное n-мерное пространство является векторно-точечным и обозначается символом A_n. Для A_n с применением аксиом линейного пространства и аксиом группы (II) легко получить (доказательства просты и мы их опускаем), что (1) векторы $\overrightarrow{AA}$ и $\overrightarrow{BB}$ для любых точек равны между собой и представляют нуль-вектор, (2) вектор $\overrightarrow{BA}$ есть противоположный вектор по отношению к $\overrightarrow{AB}$.

Если поле $\mathbf{K} \equiv \mathbf{C}$, то имеем дело с комплексным аффинным пространством, если $\mathbf{K} \equiv \mathbf{R}$, то речь идет о вещественных аффинных пространствах.

На протяжении последующих лекций (до рассмотрения унитарных пространств) мы будем иметь дело только с вещественными аффинными и евклидовыми пространствами.

Определение. Совокупность точки O и системы линейно-независимых векторов (базиса $\{e_1, \ldots, e_n\}$): $\{O, e_1, \ldots, e_n\}$ в A_n называется *аффинной системой координат*, а координаты радиус-вектора $\overrightarrow{OM}$ (для любой точки M) называются *аффинными координатами* точки M в заданной системе координат.

Поэтому, если $\overrightarrow{OM}=\sum_{k=1}^{n}\xi^k e_k$, то говорим, что точка M имеет аффинные координаты $(\xi^1; \xi^2; \ldots; \xi^n)$.

Если начало координат O осталось прежним, а $e_{k'}=\sum_{i=1}^{n}\tau_{k'}^{i} e_i$ — векторы нового базиса, то мы получим новую аффинную систему координат с прежним началом $\{O, e_{1'}, \ldots, e_{n'}\}$. В новой системе координат радиус-вектор $x=\overrightarrow{OM}$ имеет новые координаты (т. е. и точка M имеет новые координаты), которые подсчитываются (см. лекцию 18) по правилу

$$\text{(а) } [x]'=T^{-1}[x]; \quad \text{(б) } [x]=T[x]'. \tag{22.1}$$

В координатной записи преобразование (22.1) аффинной системы координат с сохранением начала координат имеет вид

$$\text{(а) } \xi^{k'}=\sum \tau_i^{k'}\xi^i; \quad \text{(б) } \xi^i=\sum_{k=1}^{n}\tau_{k'}^{i}\xi^{k'}, \tag{22.2}$$

где $\tau_i^{k'}$ — элементы обратной матрицы T^{-1}.

Более общее преобразование аффинной системы координат затрагивает и перенос начала координат. Пусть $\{O, e_1, \ldots, e_n\}$ и $\{O', e_{1'}, \ldots, e_{n'}\}$ — аффинные системы координат. Поскольку $\overrightarrow{OM}=\overrightarrow{OO'}+\overrightarrow{O'M}$, то $\overrightarrow{OM}=\sum_{k=1}^{n}\alpha^k e_k+\sum_{i=1}^{n}\xi^{i'} e_{i'}$, где $\alpha^1, \ldots, \alpha^n$ — координаты нового начала координат O' относительно нештрихованной системы координат. Поэтому имеем следующее равенство: $\sum_{k=1}^{n}\xi^k e_k=\sum_{i=1}^{n}\xi^{i'}\sum_{k=1}^{n}\tau_{i'}^{k} e_k+\sum_{k=1}^{n}\alpha^k e_k$, откуда следует

$$\xi^k=\sum_{i=1}^{n}\tau_{i'}^{k}\xi^{i'}+\alpha^k. \tag{22.3}$$

Соотношения (22.3) в матричной форме приобретают вид

$$[x]=T[x]'+[\alpha], \tag{22.4}$$

где $[\alpha]$ — матрица-столбец, построенная из координат начала O'.

Умножив слева на T^{-1} (22.4), получим

$$[x]' = T^{-1}[x] - T^{-1}[\alpha] \tag{22.5}$$

или в координатной записи

$$\xi^{k'} = \sum_{i=1}^{n} \tau_i^{k'} \xi^i + \alpha^{k'}, \tag{22.6}$$

где за $\alpha^{k'}$ обозначены числа $\left(-\sum_{i=1}^{n} \tau_i^{k'} \alpha^i\right)$.

Формулы (22.3) и (22.6) полностью исчерпывают вопрос об общем преобразовании аффинной системы координат.

Определение. Аффинное пространство A_n называется *евклидовым*, если в A_n зафиксирована *симметричная невырожденная билинейная форма* $g(x, y)$, называемая *метрической* билинейной формой (или просто *метрикой*).

Из этого определения следует, что если в некотором базисе $\{e_1, \ldots, e_n\}$ $g(x, y) = \sum_{i, k=1}^{n} g_{ik} \xi^i \eta^k$, то (1) $g_{ik} = g_{ki}$ (симметричность), (2) $\det[g_{ik}] \neq 0$ (невырожденность).

Метрическую билинейную форму принято называть *скалярным произведением векторов* x и y и обозначать просто (x, y), опуская ради простоты букву g — символ метрической билинейной формы.

Зная свойства симметричных билинейных форм, установленные нами ранее, запишем их для скалярного произведения векторов

$$\text{(а)}\ (x, y) = (y, x); \quad \text{(б)}\ (x + z, y) = (x, y) + (z, y);$$
$$\text{(в)}\ (\lambda x, y) = \lambda (x, y). \tag{22.7}$$

В аффинной системе координат скалярное произведение (x, y) подсчитывается, согласно определению билинейной формы, следующим образом:

$$(x, y) = \sum_{i, k=1}^{n} g_{ik} \xi^i \eta^k, \tag{22.8}$$

где $g_{ik} = (e_i, e_k)$ — скалярные произведения векторов базиса.

Как известно, всякая симметричная билинейная форма однозначно определяет квадратичную форму.

Метрической билинейной форме однозначно соответствует невырожденная квадратичная форма

$$(x, x) = x^2 = \sum_{i, k=1}^{n} g_{ik} \xi^i \xi^k, \tag{22.9}$$

называемая квадратом длины вектора x.

Длина вектора обозначается символом $|x|$ и, следовательно, в аффинной системе координат вычисляется по формуле

$$|x| = \sqrt{\sum_{i, k=1}^{n} g_{ik} \xi^i \xi^k}. \tag{22.10}$$

Определение. Евклидово пространство называется *собственно евклидовым* и обозначается символом E_n, если невырожденная квадратичная форма $x^2 = \sum_{i, k=1}^{n} g_{ik} \xi^i \xi^k$ является *положительно определенной*, и называется *псевдоевклидовым*, если невырожденная квадратичная форма не является положительно определенной.

В силу данного определения и определения для положительно определенных квадратичных форм следует, что для собственно евклидовых пространств к свойствам (22.7а, б, в) скалярного произведения прибавляется еще одно

$$x^2 = (x, x) > 0 \tag{22.11}$$

для любого $x \neq 0$ и $(x, x) = 0$ тогда и только тогда, когда x является нуль-вектором.

Для псевдоевклидовых пространств условие (22.11) выполняется только для части векторов и (x, x) может обращаться в нуль и для не нулевых векторов (в физике они называются *светоподобными* или *изотропными*).

Классическая физика базируется на рассмотрении 3-мерных собственных евклидовых пространств (они называются в физике как и в школьной геометрии просто евклидовыми пространствами), а релятивистская физика базируется на рассмотрении определенного класса 4-мерных псевдоевклидовых пространств. Этот класс псевдоевклидовых пространств носит название *пространств Минковского* и будет нами определен ниже.

В предыдущей лекции нами показано, что всегда существует такой базис $\{i_1, i_2, \ldots, i_n\}$ (называемый каноническим), в котором положительно определенная для собственно евклидовых пространств невырожденная квадратичная форма x^2 имеет вид

$$x^2 = (\xi^1)^2 + (\xi^2)^2 + \ldots + (\xi^n)^2. \tag{22.12}$$

Для псевдоевклидовых пространств x^2 может быть приведена к следующей канонической форме:

$$x^2 = (\xi^1)^2 + \ldots + (\xi^s)^2 - (\xi^{s+1})^2 - \ldots - (\xi^n)^2. \tag{22.13}$$

Это означает, что в базисе $\{i_1, \ldots, i_n\}$ для собственн евклидовых пространств коэффициенты метрической формы $g_{ik} = \delta_{ik} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$ и матрица $[g_{ik}]$ имеет вид

$$g = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \cdot & & 0 \\ & 0 & & \cdot & \\ & & & & 1 \end{bmatrix}, \tag{22.14}$$

а для псевдоевклидовых

$$g = \begin{bmatrix} \left.\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}\right\} s & & 0 \\ & & \\ 0 & & \left.\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}\right\} n-s \end{bmatrix}. \tag{22.15}$$

Разность между числом положительных слагаемых и числом отрицательных слагаемых в (22.13) называется сигнатурой псевдоевклидова пространства. Все псевдоевклидовы пространства разбиваются на классы в зависимости от сигнатуры пространства. Иногда сигнатуру обозначают символически следующим образом: $(\underbrace{+\ldots+}_{s}\ \underbrace{-\ldots-}_{n-s})$.

Определение. *Пространством Минковского* называется 4-мерное псевдоевклидово пространство сигнатуры $(+\,-\,-\,-)$, которая носит название *сигнатуры Лоренца.*

Часто в псевдоевклидовых пространствах вместо квадратичной формы x^2 рассматривается квадратичная форма $-x^2$, и в этом случае пространство Минковского имеет сигнатуру вида $(- + + +)$. Вследствие этого в учебниках по физике можно встретиться как с одной, так и с другой формой сигнатуры.

К псевдоевклидовым пространствам в этом курсе мы обращаться далее не будем, вернемся к рассмотрению собственно евклидовых пространств, которые будем называть евклидовыми (как это сделано во всех учебниках по линейной алгебре), и дадим определение ортогональных векторов (используемое также в случае псевдоевклидовых пространств).

Определение. Векторы x, y называются *ортогональными* между собой, если их скалярное произведение равно нулю: $(x, y) = 0$.

Обратившись к каноническому базису $\{i_1, \ldots, i_n\}$ и зная, что коэффициенты метрической билинейной формы суть скалярные произведения векторов базиса, из (22.14) заключаем

$$(i_k, i_m) = \delta_{km} = \begin{cases} 1, & k = m \\ 0, & k \neq m, \end{cases} \tag{22.16}$$

т. е. векторы канонического базиса взаимно ортогональны друг к другу и имеют длины, равные единице. Каждый такой базис называется *ортонормированным* (или *орторепером*). Итак, согласно (22.14) *в евклидовом пространстве E_n всегда существует ортонормированный базис, в котором матрица метрической формы совпадает с единичной матрицей.*

Определение. Система координат $\{O, i_1, \ldots, i_n\}$ в E_n называется *прямоугольной системой координат.*

В прямоугольной системе координат все формулы, связанные с вычислением скалярного произведения (метрических величин), сильно упрощаются, и в этом смысле прямоугольная система координат является предпочтительной. Например, скалярное произведение векторов и длина вектора имеют вид

$$\text{(а)}\ (x, y) = \sum_{k=1}^{n} \xi^k \eta^k; \quad \text{(б)}\ |x| = \sqrt{\sum_{k=1}^{n} (\xi^k)^2}. \tag{22.17}$$

Для того, чтобы ввести понятие угла между векторами в E_n, докажем предварительно следующее утверждение

Теорема. *В пространстве E_n для любых векторов x, y имеет место неравенство Коши-Буняковского*

$$|(x, y)| \leqslant |x||y|. \qquad (22.18)$$

В самом деле, для любого $\lambda \in \mathbf{R}$ и любых x, y согласно (22.11) выполнено: $(\lambda x - y, \lambda x - y) = \lambda^2 (x, x) - 2\lambda (x, y) + (y, y) \geqslant 0$,. т. е. квадратный трехчлен относительно λ не может иметь различных вещественных корней. Поэтому его дискриминант $D = (x, y)^2 - (x, x)(y, y) \leqslant 0$, откуда следует неравенство (22.18).

Неравенство Коши-Буняковского позволяет ввести понятие угла между векторами.

Определение. Углом θ между не нулевыми векторами x, y мы будем называть угол (в пределах от 0 до π), косинус которого равен $\dfrac{(x, y)}{|x||y|}$. В аффинной системе координат он подсчитывается по формуле

$$\cos\theta = \frac{\sum\limits_{i, k=1}^{n} g_{ik}\xi^i\eta^k}{\sqrt{\sum\limits_{i, k=1}^{n} g_{ik}\xi^i\xi^k}\sqrt{\sum\limits_{i, k=1}^{n} g_{ik}\eta^i\eta^k}}. \qquad (22.19)$$

В прямоугольной системе координат (22.19) принимает более простой вид

$$\cos\theta = \frac{\sum\limits_{k=1}^{n} \xi^k\eta^k}{\sqrt{\sum\limits_{k=1}^{n} (\xi^k)^2}\sqrt{\sum\limits_{k=1}^{n} (\eta^k)^2}}. \qquad (22.20)$$

Поставим вопрос о переходе в E_n от одной прямоугольной системы координат к другой прямоугольной системе координат. Предварительно дадим определение ортогональной матрицы.

Определение. Матрица $A = [\alpha_k^i]$ называется *ортогональной*, если она удовлетворяет условию

$$A^{tr}A = E, \qquad (22.21)$$

которое в подробной записи имеет вид

$$\sum_{p=1}^{n} \alpha_k^p \alpha_m^p = \delta_{km} = \begin{cases} 1, & k = m \\ 0, & k \neq m. \end{cases} \tag{22.22}$$

Из (22.22) заключаем, что *у всякой ортогональной матрицы сумма* произведений элементов столбца на самого себя равна 1, а сумма произведений элементов разных столбцов равна нулю. Это правило легко позволяет отличить ортогональную матрицу от неортогональной. Его можно было бы взять за новое определение ортогональной матрицы, поскольку оно означает выполнение (22.22).

Из (22.21) следует, что $(\det A^{tr})\cdot(\det A) = \det E$ или $(\det A)^2 = 1$, и, следовательно, если A — ортогональная матрица, то

$$\det A = \pm 1. \tag{22.23}$$

Это означает, что ортогональная матрица обязательно невырожденная, и если (22.21) умножить справа на A^{-1}, то получим эквивалентное (22.21) условие

$$A^{tr} = A^{-1}. \tag{22.24}$$

Условие (22.24) можно было бы взять еще за одно определение ортогональной матрицы.

Пусть $\{O, i_1, \ldots, i_n\}$ и $\{O, i_{1'}, \ldots, i_{n'}\}$ — две прямоугольные системы координат в E_n с общим началом, т. е. совершается переход от одного ортонормированного базиса к другому. Обозначим матрицу перехода от нештрихованного базиса к штрихованному через $Q = [q_{k'}^m]$, так что $i_{k'} = \sum\limits_{m=1}^{n} q_{k'}^m i_m$. Ввиду того, что $i_{1'}, i_{2'}, \ldots, i_{n'}$ — единичные взаимно ортогональные векторы, имеем

$$\delta_{ks} = (i_{k'}, i_{s'}) = \Big(\sum_{m=1}^{n} q_{k'}^m i_m, \ \sum_{l=1}^{n} q_{s'}^l i_l\Big) =$$

$$= \sum_{m,\, l=1}^{n} q_{k'}^m q_{s'}^l (i_m, i_l) = \sum_{m,\, l=1}^{n} q_{k'}^m q_{s'}^l \delta_{ml} = \sum_{m=1}^{n} q_{k'}^m q_{s'}^m .$$

Итак, элементы матрицы перехода связаны соотношением

$$\sum_{p=1}^{n} q_{k'}^p q_{s'}^p = \delta_{ks},$$

которое совпадает с (22.22) и показывает, что *матрица перехода от одного ортонормированного базиса к другому необходимо является ортогональной матрицей.*

Обратно, если задана некоторая ортогональная матрица $A=[\alpha^i_k]$, то, как известно, она невырожденная и с ее помощью можно осуществить переход от одного базиса к другому. Пусть $\{i_1, \dots, i_n\}$ — ортонормированный базис в E_n. Построим векторы

$$e_{k'}=\sum_{m=1}^{n}\alpha^m_k i_m .$$

Они линейно независимые ($\det A \neq 0$), и мы можем построить базис $\{e_{1'}, \dots, e_{n'}\}$. Вычислим скалярное произведение векторов нового базиса. Имеем

$$(e_{k'}, e_{m'})=\left(\sum_{p=1}^{n}\alpha^p_k i_p, \sum_{q=1}^{n}\alpha^q_m i_q\right)=\sum_{p, q=1}^{n}\alpha^p_k\alpha^q_m\delta_{pq}=\sum_{p=1}^{n}\alpha^p_k\alpha^p_m .$$

Поскольку A — ортогональная матрица, то в силу (22.22) $(e_{k'}, e_{m'})=\delta_{km}$, и, следовательно, базис $\{e_{1'}, \dots, e_{n'}\}$ является также ортонормированным базисом.

Таким образом, установлено *взаимно-однозначное соответствие между ортогональными матрицами и вращениями базиса (репера) в пространстве* E_n. Если детерминант ортогональной матрицы перехода равен $+1$, то вращения называются *собственными* (ориентация репера сохраняется. Например, в E_3 правая тройка переходит в правую), если же детерминант ортогональной матрицы перехода равен -1, то вращения называются *несобственными* (ориентация репера меняется. Например, в E_3 правая тройка переходит в левую).

Примером собственных вращений в трехмерном евклидовом пространстве E_3 являются вращения, задаваемые углами Эйлера, изученные нами в лекции 11.

Применяя сейчас формулы преобразования координат (22.3), (22.6) в произвольном случае, мы можем написать формулы преобразования от одной прямоугольной системы координат к другой прямоугольной системе координат, когда перенесено и начало координат

$$
\left.\begin{aligned}
&(\text{а})\ \ \xi^k = \sum_{m=1}^{n} q^k_{m'}\,\xi^{m'} + \alpha^k \\
&(\text{б})\ \ \xi^{k'} = \sum_{m=1}^{n} q^{k'}_m\,\xi^m + \alpha^{k'}
\end{aligned}\right\} \qquad (22.25)
$$

или в матричной форме

$$
\begin{aligned}
&(\text{а})\ \ [x] = Q\,[x]' + [\alpha] \\
&(\text{б})\ \ [x]' = Q^{-1}\,[x] + [\alpha]'
\end{aligned} \qquad (22.26)
$$

где $Q = [q^k_{m'}]$—ортогональная матрица и, следовательно, в (22.26б) Q^{-1} может быть заменено на Q^{tr} согласно (22.24).

Лекция 23

ИЗОМЕТРИЧЕСКИЙ ОПЕРАТОР. САМОСОПРЯЖЕННЫЙ (СИММЕТРИЧЕСКИЙ) ОПЕРАТОР И ЕГО МАТРИЦА. СВЯЗЬ СИММЕТРИЧЕСКОЙ БИЛИНЕЙНОЙ ФОРМЫ С СООТВЕТСТВУЮЩИМ ЕЙ САМОСОПРЯЖЕННЫМ ОПЕРАТОРОМ

Переходим к изучению свойств некоторых операторов в E_n, играющих важную роль в приложениях (в теоретической механике и квантовой механике).

Определение. Оператор A в E_n называется *изометрическим* (*изометричным*), если он сохраняет скалярное произведение векторов, т. е. $(Ax, Ay) = (x, y)$ для любых $x, y \in E_n$.

Другими словами, изометричный оператор сохраняет метрику пространства и соответственно все метрические свойства пространства E_n.

Пусть $A = [\alpha^i_k]$—матрица изометричного оператора в произвольном базисе $\{e_1, \ldots, e_n\}$. Согласно определению имеем $(Ae_k, Ae_m) = (e_k, e_m)$. Рассмотрим подробнее левую часть этого равенства и учтем, что $(e_k, e_m) = g_{km}$.

$$
(Ae_k, Ae_m) = \left(\sum_{i=1}^{n} \alpha^i_k e_i,\ \sum_{p=1}^{n} \alpha^p_m e_p\right) =
$$

$$= \sum_{i,\,p=1}^{n} \alpha_k^i \alpha_m^p (e_i\,,\, e_p) = \sum_{i,\,p=1}^{n} \alpha_k^i g_{ip} \alpha_m^p\,,$$

т. е. элементы матрицы изометричного оператора и коэффициенты метрической формы связаны соотношением

$$\sum_{i,\,p=1}^{n} \alpha_k^i g_{ip} \alpha_m^p = g_{km} \tag{23.1}$$

или в матричном представлении

$$A^{tr} g A = g, \tag{23.2}$$

где через $g = [g_{km}]$ обозначена матрица метрической билинейной формы.

Если в E_n выбрана прямоугольная система координат и $\{i_1\,, \ldots, i_n\}$ — ортонормированный базис, то в нем $g = E$ и (23.2) переходит в соотношение $A^{tr} A = E$. Это значит, что *у всякого изометричного оператора его матрица невырожденная и в ортонормированном базисе есть ортогональная матрица.*

Обратно, пусть в ортонормированном базисе матрица A некоторого оператора $\boldsymbol{A}$ является ортогональной. Тогда $\boldsymbol{A}$ — изометричный оператор.

В самом деле,

$$(\boldsymbol{A}x,\ \boldsymbol{A}y) = \left(\boldsymbol{A}\left(\sum_{k=1}^{n} \xi^k i_k\right),\ \boldsymbol{A}\left(\sum_{m=1}^{n} \eta^m i_m\right)\right) =$$

$$= \sum_{k,\,m=1}^{n} \xi^k \eta^m (\boldsymbol{A}i_k\,,\ \boldsymbol{A}i_m) = \sum_{k,\,m=1}^{n} \xi^k \eta^m \left(\sum_{p=1}^{n} \alpha_k^p i_p\,,\ \sum_{q=1}^{n} \alpha_m^q i_q\right) =$$

$$= \sum_{k,\,m=1}^{n} \xi^k \eta^m \left(\sum_{p,\,q=1}^{n} \alpha_k^p \alpha_m^q \delta_{pq}\right) = \sum_{k,\,m=1}^{n} \xi^k \eta^m \delta_{km} = (x,\ y).$$

Определение. Оператор $\boldsymbol{B}$ называется *сопряженным* к оператору $\boldsymbol{A}$ относительно операции скалярного произведения, если $(\boldsymbol{B}x,\ y) = (x,\ \boldsymbol{A}y)$ для любых $x,\ y \in E_n$.

Пусть $A = [\alpha_k^i]$ — заданная матрица оператора $\boldsymbol{A}$ в произвольном базисе $\{e_1\,, \ldots, e_n\}$. Какова матрица оператора $\boldsymbol{B}$ в этом базисе? Пусть $B = [\beta_k^i]$, где β_k^i — пока неизвестные числа. Из определения следует: $(\boldsymbol{B}e_k\,,\ e_m) = (e_k\,,\ \boldsymbol{A}e_m)$ или в подробной записи

$$\left(\sum_{p=1}^{n} \beta_k^p e_p,\, e_m\right) = \left(e_k,\, \sum_{q=1}^{n} \alpha_m^q e_q\right)$$

т. е.

$$\sum_{p=1}^{n} \beta_k^p g_{pm} = \sum_{q=1}^{n} g_{kq} \alpha_m^q. \tag{23.3}$$

Последнее равенство в матричном виде записывается следующим образом (учтем, что слева стоят элементы с индексами k, m, а справа с индексами m, k):

$$gB = (gA)^{tr}. \tag{23.4}$$

Но при транспонировании произведения матриц необходимо перемножать транспонированные матрицы в обратном порядке, т. е. $(gA)^{tr} = A^{tr} g^{tr}$. Поскольку для метрической билинейной формы $g^{tr} = g$ (форма — симметричная), то (23.4) эквивалентно соотношению

$$B = g^{-1} A^{tr} g. \tag{23.5}$$

Условие (23.5) и есть то условие, которому удовлетворяет матрица сопряженного оператора в произвольном базисе.

Если базис ортонормированный, то $g = E$, $g^{-1} = E$ и из (23.5) следует, что $B = A^{tr}$. Итак, *если в ортонормированном базисе A — матрица некоторого оператора $\boldsymbol{A}$, то матрица A^{tr} определяет сопряженный ему оператор*. Сопряженный к $\boldsymbol{A}$ оператор обычно обозначают символом $\boldsymbol{A}^c$ (или $\boldsymbol{A}^+$), и, следовательно, в ортонормированном базисе его матрица $A^c = A^{tr}$.

Определение. Оператор $\boldsymbol{A}$ называется *самосопряженным* (*симметрическим*), если $\boldsymbol{A} = \boldsymbol{A}^c$, т. е. $(\boldsymbol{A}x, y) = (x, \boldsymbol{A}y)$ для любых $x, y \in E_n$.

Согласно (23.5) матрица A самосопряженного оператора в произвольном базисе удовлетворяет условию

$$A = g^{-1} A^{tr} g, \tag{23.6}$$

а в ортонормированном базисе, когда $g = E$, имеем

$$A = A^{tr} \Rightarrow \alpha_k^l = \alpha_l^k. \tag{23.7}$$

Итак, *у всякого самосопряженного оператора в ортонормированном базисе матрица необходимо* *симметрическая* (отсюда и еще одно название у самосопряженного оператора — симметрический оператор).

Наоборот, пусть в ортонормированном базисе матрица некоторого оператора A симметрическая, т. е. $\alpha^m_k = \alpha^k_m$. Покажем, что A — самосопряженный оператор.

В самом деле,

$$Ax = \left(\sum_{k=1}^{n} \alpha^1_k \xi^k\right) i_1 + \ldots + \left(\sum_{k=1}^{n} \alpha^n_k \xi^k\right) i_n,$$

$$Ay = \left(\sum_{m=1}^{n} \alpha^1_m \eta^m\right) i_1 + \left(\sum_{m=1}^{n} \alpha^2_m \eta^m\right) i_2 + \ldots + \left(\sum_{m=1}^{n} \alpha^n_m \eta^m\right) i_n.$$

Тогда

$$(Ax, y) = \sum_{m=1}^{n}\left(\sum_{k=1}^{n} \alpha^m_k \xi^k\right)\eta^m = \sum_{k=1}^{n}\left(\sum_{m=1}^{n} \alpha^k_m \eta^m\right)\xi^k = (x, Ay).$$

Тем самым *между множеством всех симметричных матриц и множеством самосопряженных операторов существует взаимно-однозначное соответствие.*

Но каждая симметрическая матрица однозначно определяет симметричную билинейную форму и соответствующую квадратичную форму. Поэтому установим *связь между самосопряженными операторами и симметричными билинейными формами* (эквивалентно *квадратичными формами*).

Пусть A — самосопряженный оператор, $A = [\alpha^i_k]$ — его матрица в ортонормированном базисе. Введем обозначение $\widetilde{\alpha}_{km} = \alpha^k_m$ и построим симметричную билинейную форму $\widetilde{A}(x, y) = \sum_{k, m=1}^{n} \widetilde{\alpha}_{km} \xi^k \eta^m$ ($\widetilde{\alpha}_{km} = \widetilde{\alpha}_{mk}$ из-за симметричности матрицы $A = [\alpha^k_m]$). Что она собой представляет? Для ответа на вопрос подсчитаем скалярное произведение (x, Ay). Имеем

$$(x, Ay) = \sum_{k=1}^{n} \xi^k \left(\sum_{m=1}^{n} \alpha^k_m \eta^m\right) = \sum_{k, m=1}^{n} \widetilde{\alpha}_{km} \xi^k \eta^m = \widetilde{A}(x, y).$$

Таким образом, мы доказали, что *каждому самосопряженному оператору однозначно сопоставляется* *симметричная билинейная форма, которая представляет собой скалярное произведение вектора x и образа вектора y при самосопряженном эндоморфизме* $\boldsymbol{A}$.

Отметим, что из-за самосопряженности $(x, \boldsymbol{A}y) = (\boldsymbol{A}x, y)$, и поэтому симметричную билинейную форму $\widetilde{A}(x, y)$ можно задать и таким скалярным произведением $(\boldsymbol{A}x, y) = \widetilde{A}(x, y)$.

Докажем сейчас обратное, что *любой симметрической билинейной форме можно однозначно сопоставить самосопряженный оператор.*

Действительно, пусть $B(x, y) = \sum\limits_{k, m=1}^{n} \beta_{km} \xi^k \eta^m$ — симметричная билинейная форма ($\beta_{km} = \beta_{mk}$), заданная в ортонормированном базисе. Введем в рассмотрение элементы матрицы $A = [\alpha_m^k]$, определив их равенством $\alpha_m^k = \beta_{km}$. Из-за симметричности билинейной формы получим, что $A = A^{tr}$. Но в таком случае матрица A определяет по свойству, доказанному нами выше, самосопряженный оператор.

Итак, *между множеством самосопряженных операторов и симметрическими билинейными формами (квадратичными формами) существует взаимно-однозначное соответствие.*

Полученный результат имеет очень важное значение при решении вопроса о приведении ортогональными преобразованиями к каноническому виду квадратичных форм (или что то же самое — симметричных билинейных форм).

В самом деле, пусть Q — матрица перехода от одного ортонормированного базиса к другому ортонормированному базису, т. е. Q — ортогональная матрица, и пусть $A = [\alpha_{km}]$ — матрица квадратичной формы ($\alpha_{km} = \alpha_{mk}$). Тогда согласно общей теории билинейных форм матрица A преобразуется по закону $A' = Q^{tr}AQ$. Но в ортонормированном базисе матрица A определяет и самосопряженный оператор $\boldsymbol{A}$. Согласно общей теории линейных операторов закон преобразования для матриц операторов таков: $A' = Q^{-1}AQ$. Но для ортогональных матриц $Q^{-1} = Q^{tr}$, и, следовательно, *закон преобразования матрицы самосопряженного* *оператора при переходе к новому ортонормированному базису совпадает с законом преобразования матрицы квадратичной формы.* Из этого факта тотчас следует вывод: *если удастся ортогональными преобразованиями привести к диагональному виду матрицу самосопряженного оператора, то тем самым будет указан канонический вид квадратичной формы*, ибо матрица квадратичной формы будет в новом базисе иметь тот же диагональный вид, что и матрица оператора.

Нам известно, что наиболее простой вид матрица оператора имеет в базисе, где векторы базиса являются собственными векторами оператора. Поэтому необходимо выяснить все свойства собственных значений и собственных векторов самосопряженного оператора.

Теорема 1. *Все собственные значения самосопряженного оператора вещественные числа.*

В самом деле, пусть λ_0 — корень характеристического уравнения $\det(\alpha_k^i - \lambda\delta_k^i) = 0$, где $[\alpha_k^i]$ — матрица самосопряженного оператора в ортонормированном базисе (для доказательства теоремы достаточно ограничиться рассмотрением ортонормированного базиса, так как собственные значения оператора—инварианты преобразования базиса). Из системы алгебраических уравнений

$$\sum_{q=1}^{n} \alpha_q^p \xi^q = \lambda_0 \xi^p \quad (p = \overline{1, n}) \tag{23.8}$$

определяются координаты ξ^q собственных векторов, отвечающих собственному значению λ_0. При этом мы не знаем пока, вещественные или комплексные числа $\lambda_0, \xi^1, \ldots, \xi^n$. Будем предполагать, что имеет место общий случай — эти числа комплексные, потому что вещественные числа есть частный случай комплексных. Умножим равенство (23.8) на комплексно-сопряженные числа $\bar{\xi}^p$ и просуммируем по p от 1 до n. Имеем

$$\sum_{p,\, q=1}^{n} \alpha_q^p \bar{\xi}^p \xi^q = \lambda_0 \sum_{p=1}^{n} \xi^p \bar{\xi}^p. \tag{23.9}$$

Выражение $\sum_{p=1}^{n} \xi^p \bar{\xi}^p$ — вещественное число, так как каждое слагаемое $\xi^p \bar{\xi}^p$ — вещественное число. Если мы покажем, что и левая часть равенства (23.9) вещественна, то λ_0 — вещественное число и теорема будет доказана.

В левой части равенства (23.9) слагаемые, для которых $p=q$ являются вещественными числами. Те члены, где $p \neq q$, будем рассматривать попарно, объединяя, например, члены при $p=s$, $q=t$ и при $p=t$, $q=s$ (s, t — фиксированные индексы). Тогда такая пара слагаемых представляет собой $(\alpha_t^s \xi^t \bar{\xi}^s + \alpha_s^t \xi^s \bar{\xi}^t)$. Но в силу самосопряженности оператора $\alpha_s^t = \alpha_t^s$, и, следовательно, рассматриваемая пара слагаемых равна $\alpha_t^s (\xi^t \bar{\xi}^s + \xi^s \bar{\xi}^t) = \alpha_t^s (\xi^t \bar{\xi}^s + \overline{\bar{\xi}^s \xi^t})$. В скобках стоит сумма комплексного числа и сопряженного ему комплексного числа, т. е. скобка есть вещественное число. Итак, левая часть равенства (23.9) — вещественное число. Теорема доказана.

Теорема 2. *Собственные векторы самосопряженного оператора, отвечающие различным собственным значениям, ортогональны между собой.*

Действительно, пусть $Ax = \lambda x$, $Ay = \mu y$ $(\lambda \neq \mu)$. Тогда $(Ax, y) = \lambda (x, y)$ и $(x, Ay) = \mu \cdot (x, y)$. Но оператор A самосопряженный, и, следовательно, $(Ax, y) = (x, Ay)$. Это означает, что $\lambda \cdot (x, y) = \mu \cdot (x, y)$ или $(\lambda - \mu) \cdot (x, y) = 0$, откуда следует, что $(x, y) = 0$. Собственные векторы ортогональны между собой.

Доказанные теоремы позволяют нам перейти к вопросу о каноническом виде матрицы самосопряженного оператора и соответственно каноническом виде квадратичной формы, полученном с помощью ортогональных преобразований. К решению поставленного вопроса мы перейдем в следующей лекции.

Лекция 24

ОРТОГОНАЛИЗАЦИЯ СИСТЕМЫ ЛИНЕЙНО-НЕЗАВИСИМЫХ ВЕКТОРОВ. ОСНОВНАЯ ТЕОРЕМА О ДИАГОНАЛЬНОМ ВИДЕ МАТРИЦЫ САМОСОПРЯЖЕННОГО ОПЕРАТОРА И КАНОНИЧЕСКОМ ВИДЕ КВАДРАТИЧНОЙ ФОРМЫ

Взаимно-однозначное соответствие между симметричными билинейными формами и самосопряженными операторами в E_n, установленное в предыдущей лекции, позволяет решить один из наиболее важных для физических приложений вопросов о приведении ортогональными преобразованиями к главным осям симметрических матриц, описывающих физические характеристики той или иной физической системы (в механике и электродинамике сплошных сред).

Предварительно решим задачу о построении ортонормированного базиса для линейной оболочки в E_n, натянутой на систему векторов $\{f_1, f_2, \dots, f_k\}$ $(k \leqslant n)$, причем, не ограничивая общности рассуждений, будем считать эти векторы линейно-независимыми (в противном случае из системы векторов выделяем максимальное число линейно-независимых и процесс ортогонализации будем осуществлять только для них).

Рассмотрим вектор f_1, пронормируем его и обозначим $g_1 = \frac{f_1}{|f_1|}$. Рассмотрим далее вектор $f_2' = \alpha_1 g_1 + f_2$ с неизвестным коэффициентом α_1 и выберем его так, чтобы f_2' был ортогонален к g_1, т. е. $(g_1, f_2') = 0$. Последнее равенство означает, что $\alpha_1 \cdot 1 + (g_1, f_2) = 0$, откуда следует

$$\alpha_1 = -(g_1, f_2).$$

Пронормируем вектор f_2' и обозначим $g_2 = \frac{f_2'}{|f_2'|}$. Затем строим вектор $f_3' = \beta_1 g_1 + \beta_2 g_2 + f_3$ с неизвестными коэффициентами β_1, β_2, которые определим из условий ортогональности f_3' к g_1 и к g_2: $(g_1, f_3') = \beta_1 \cdot 1 + (g_1, f_3) = 0$ и $(g_2, f_3') = \beta_2 \cdot 1 + (g_2, f_3) = 0$, т. е.

$$\beta_1 = -(g_1, f_3), \ \beta_2 = -(g_2, f_3).$$

Пронормируем вектор f_3' и обозначим $g_3 = \frac{f_3'}{|f_3'|}$. Продолжая процесс, мы построим систему единичных взаимно ортогональных векторов $\{g_1, g_2, \dots, g_k\}$, которую можно взять за базис рассматриваемой оболочки, так как система построенных векторов будет линейно-независима. Последний факт нуждается в доказательстве. Действительно, составим линейную комбинацию векторов $g_1, \ldots, g_k$ с произвольными коэффициентами $\lambda_1, \ldots, \lambda_k$ и приравняем ее нулю

$$\lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_k g_k = 0. \qquad (24.1)$$

Умножая (24.1) скалярно на вектор g_1 и учитывая ортогональность к g_1 векторов $g_2, \ldots, g_k$, получим $\lambda_1 \cdot 1 = 0$. Аналогично умножая (24.1) скалярно на $g_2, \ldots, g_k$, получим $\lambda_2 = 0$, $\lambda_3 = 0, \ldots, \lambda_k = 0$, что и доказывает линейную независимость построенных векторов.

Если нам необходимо перейти от ортонормированного базиса $\{i_1, \ldots, i_n\}$, относительно которого были заданы векторы $f_1, f_2, \ldots, f_k$, к другому ортонормированному базису в E_n, первыми векторами которого служили бы векторы $g_1, \ldots, g_k$, то поступаем следующим образом.

Пусть $g_1 = \sum_{m=1}^{n} \xi_1^m i_m, \ldots, g_k = \sum_{m=1}^{n} \xi_k^m i_m$. Ищем множество всех таких векторов $x = \sum_{m=1}^{n} \xi^m i_m$, которые были бы ортогональны одновременно к $g_1, \ldots, g_k$. Для этого строим нормальную систему решений для системы линейных уравнений

$$\left.\begin{aligned} (g_1, x) &= \sum_{m=1}^{n} \xi_1^m \xi^m = 0 \\ (g_2, x) &= \sum_{m=1}^{n} \xi_2^m \xi^m = 0 \\ &\ldots\ldots\ldots\ldots \\ (g_k, x) &= \sum_{m=1}^{n} \xi_k^m \xi^m = 0 \end{aligned}\right\}. \qquad (24.2)$$

Тогда совокупность векторов x будет образовывать линейную оболочку, натянутую на векторы нормальной фундаментальной системы решений. Поскольку ранг матрицы системы уравнений (24.2) равен k (векторы $g_1, \ldots, g_k$ линейно-независимы), то имеем $(n-k)$ векторов нормальной фундаментальной системы решений. Для них применим процесс ортогонализации и обозначим полученные единичные взаимноортогональные векторы через $h_1, h_2, \ldots, h_{n-k}$. Система векторов $\{g_1, \ldots, g_k, h_1, \ldots, h_{n-k}\}$ и будет служить новым ортонормированным базисом в E_n. Ортогональная матрица перехода от базиса $\{i_1, \ldots, i_n\}$ к базису $\{g_1, \ldots, g_k, h_1, \ldots, h_{n-k}\}$ строится очень просто. Ее столбцами служат матрицы-столбцы из координат векторов $\{g_1, \ldots, g_k, h_1, \ldots, h_{n-k}\}$.

Пример. В E_5 задана оболочка, натянутая на линейно-независимые векторы

$$f_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad f_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad f_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}.$$

Построить ортонормированный базис $\{g_1, g_2, g_3\}$ оболочки и перейти к новому ортонормированному базису в E_5, первыми тремя векторами которого служили бы g_1, g_2, g_3.

Решение. Пронормируем f_1. Имеем $|f_1| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$. Тогда $g_1 = \frac{1}{\sqrt{2}}(i_1 - i_2)$. Находим, что $\alpha_1 = -(g_1, f_2) = \frac{1}{\sqrt{2}}$, $f_2' = \frac{1}{\sqrt{2}} g_1 + f_2 = \frac{1}{2} i_1 - \frac{1}{2} i_2 + i_2 + 2i_3 = \frac{1}{2} i_1 + \frac{1}{2} i_2 + 2i_3$. Поэтому $g_2 = \frac{f_2'}{|f_2'|} = \frac{1}{3\sqrt{2}}(i_1 + i_2 + 4i_3)$. Далее находим, что

$$\beta_1 = -(g_1, f_3) = 0, \ \beta_2 = -(g_2, f_3) = -\frac{4}{3\sqrt{2}},$$

$$f_3' = -\frac{4}{3\sqrt{2}} g_2 + f_3 = -\frac{1}{9}(2i_1 + 2i_2 - i_3) - i_4.$$

Поэтому $g_3 = -\frac{2}{3\sqrt{10}}\left(i_1 + i_2 - \frac{1}{2} i_3 + \frac{9}{2} i_4\right)$.

Решаем систему линейных уравнений (24.2)

$$\left.\begin{array}{l} \frac{1}{\sqrt{2}}\xi^1 - \frac{1}{\sqrt{2}}\xi^2 = 0 \\ \frac{1}{3\sqrt{2}}\xi^1 + \frac{1}{3\sqrt{2}}\xi^2 + \frac{4}{3\sqrt{2}}\xi^3 = 0 \\ -\frac{2}{3\sqrt{10}}\xi^1 - \frac{2}{3\sqrt{10}}\xi^2 + \frac{1}{3\sqrt{10}}\xi^3 - \frac{9}{3\sqrt{10}}\xi^4 = 0 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \left\{\begin{array}{l} \xi^1 = -2c_1 \\ \xi^2 = -2c_1 \\ \xi^3 = \xi^4 = c_1 \\ \xi^5 = c_2 . \end{array}\right.$$

Векторы нормальной фундаментальной системы решений имеют вид $x_1 = -2i_1 - 2i_2 + i_3 + i_4$, $x_2 = i_5$. Поэтому $h_1 = \frac{1}{\sqrt{10}}(-2i_1 - 2i_2 + i_3 + i_4)$, $h_2 = i_5$.

Ортогональная матрица перехода от базиса $\{i_1, \dots, i_5\}$ к базису $\{g_1, g_2, g_3, h_1, h_2\}$ имеет вид

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3\sqrt{10}} & -\frac{2}{\sqrt{10}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3\sqrt{10}} & -\frac{2}{\sqrt{10}} & 0 \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3\sqrt{10}} & \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 & -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \qquad (24.3)$$

Перейдем к формулировке и доказательству основной теоремы.

Теорема 1. *У каждого самосопряженного оператора в E_n существует n единичных взаимноортогональных собственных векторов. Матрица оператора относительно ортонормированного базиса, построенного на указанных собственных векторах оператора, имеет диагональный вид*

$$\begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & & \\ & & & \ddots & & \\ & & & & \lambda_s & \\ & & & & & \ddots \\ & & & & & & \lambda_s \end{bmatrix}, \qquad (24.4)$$

где по диагонали стоят собственные значения оператора, повторенные столько раз, какова их кратность в характеристическом уравнении.

Доказательство. Пусть λ_1 — собственное значение самосопряженного оператора A, заданного в некотором ортонормированном базисе $\{i_1, \ldots, i_n\}$ своей симметрической матрицей $A = [\alpha_q^p]$. Рассмотрим собственный вектор f_1 (для этого достаточно взять первый вектор из нормальной фундаментальной системы решений системы однородных уравнений

$$\sum_{q=1}^{n} (\alpha_q^p - \lambda_1 \delta_q^p)\, \xi^q = 0)$$

и пронормируем его. Получим единичный собственный вектор $g_1 = \frac{f_1}{|f_1|}$, для которого $Ag_1 = \lambda_1 g_1$.

Рассмотрим далее множество всех векторов $x = \sum_{m=1}^{n} \xi^m i_m$, ортогональных к $g_1 = \sum_{m=1}^{n} \xi_1^m i_m$. Для этого необходимо построить векторы нормальной фундаментальной системы решений для системы уравнений, состоящей лишь из одного уравнения

$$(g_1, x) = \xi_1^1 \xi^1 + \xi_1^2 \xi^2 + \ldots + \xi_1^n \xi^n = 0.$$

Таких векторов будет $n - 1$. Применим к ним процесс ортогонализации и получим $(n - 1)$ единичных взаимно-ортогональных векторов $\{h_2, h_3, \ldots, h_n\}$, каждый из которых также ортогонален и к вектору g_1. Очевидно, что искомые векторы x принадлежат к линейной оболочке, натянутой на векторы $\{h\}$, а сама она представляет $(n - 1)$-мерное векторное пространство E_{n-1}, ортогональное к g_1 (т. е. любой вектор из E_{n-1} ортогонален к g_1). Покажем, что векторы из E_{n-1} под действием самосопряженного оператора A переходят в векторы того же E_{n-1} (другими словами E_{n-1} является инвариантным подпространством для $\boldsymbol{A}$). В самом деле, используя самосопряженность $\boldsymbol{A}$, имеем $(\boldsymbol{A}x, g_1) = (x, \boldsymbol{A}g_1) = (x, \lambda_1 g_1) = \lambda_1 (x, g_1) = 0$, т. е. $\boldsymbol{A}x \in E_{n-1}$. От ортонормированного базиса $\{i_1, \ldots, i_n\}$ перейдем к ортонормированному базису $\{g_1, h_2, \ldots, h_n\}$. Поскольку

$\boldsymbol{A}g_1 = \lambda_1 g_1$, $\boldsymbol{A}h_s = \sum\limits_{r=2}^{n} \beta_s^r h_r$ $(s = \overline{2, n})$, то матрица самосопряженного оператора в новом базисе имеет вид

$$\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \beta_2^2 & \ldots & \beta_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \beta_2^n & \ldots & \beta_n^n \end{bmatrix}. \tag{24.5}$$

В силу вышесказанного матрица $B = [\beta_s^r]$ $(r, s = \overline{2, n})$ есть матрица самосопряженного оператора $\boldsymbol{A}$, действующего на векторах пространства E_{n-1}. Затем весь рассмотренный нами процесс повторяем уже для E_{n-1} и матрицы B. А именно, находим корни характеристического уравнения $\det(B - \lambda E) = 0$, и если, например, λ_1 корень уравнения $\det(A - \lambda E) = 0$ кратности $r_1 > 1$, то λ_1 непременно будет корнем уравнения $\det(B - \lambda E) = 0$, поскольку характеристический полином оператора (вместе с ним и собственные значения) есть инвариант преобразования базисов, а характеристический полином матрицы (24.5) имеет вид $\chi(\lambda) = (\lambda_1 - \lambda)\det(B - \lambda E)$.

Из системы уравнений $\sum\limits_{s=2}^{n} (\beta_s^r - \lambda_1 \delta_s^r)\, \xi^s = 0$ (если λ_1 — однократный корень уравнения $\det(A - \lambda E) = 0$, то в системе вместо λ_1 стоит другой корень уравнения $\det(B - \lambda E) = 0$) определяем собственный вектор f_2, нормируем его и получим единичный собственный вектор g_2 $(\boldsymbol{A}g_2 = \lambda_1 g_2)$, ортогональный к единичному собственному вектору g_1. Как и выше, в E_{n-1} строим подпространство E_{n-2}, ортогональное к g_2 (автоматически и к g_1), с ортонормированным базисом $\{h_3', \ldots, h_n'\}$. Перейдя от базиса $\{g_1, h_2, \ldots, h_n\}$ к базису $\{g_1, g_2, h_3', \ldots, h_n'\}$, мы получим следующий вид матрицы самосопряженного оператора в новом базисе:

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_1 & 0 & \ldots & 0 \\ 0 & 0 & \gamma_3^3 & \ldots & \gamma_n^3 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \gamma_3^n & \ldots & \gamma_n^n \end{bmatrix}. \qquad (24.6)$$

Заметим, что если λ_1 — однократное собственное значение, то в матрице (24.6) на пересечении 2-й строки и 2-го столбца будет стоять не λ_1, а другой корень характеристического уравнения $\det(A - \lambda E) = 0$.

Продолжая предложенный алгоритм построения ортогональных друг другу единичных собственных векторов оператора, через $n-2$ шага мы придем к ортонормированному базису $\{g_1, g_2, \ldots, g_n\}$, состоящему из собственных векторов оператора A, в котором матрица оператора примет вид (24.4).

Важнейшим следствием доказанной теоремы является тот факт, что *r-кратному собственному значению самосопряженного оператора соответствует r-мерное инвариантное подпространство собственных векторов оператора, отвечающих данному собственному значению*, в чем мы убедились, применяя конструктивный метод доказательства основной теоремы.

Это следствие дает практический способ построения того ортонормированного базиса, в котором матрица оператора имеет диагональный вид. Пусть $f_1, \ldots, f_{r_1}$ — линейно-независимые собственные векторы, отвечающие собственному значению λ_1 кратности r_1 (для их нахождения достаточно построить векторы нормальной фундаментальной системы решений для однородной системы уравнений $\sum_{q=1}^{n} (\alpha_q^p - \lambda_1 \delta_q^p) \xi^q = 0$. Согласно следствию будет r_1 таких векторов). Далее совершим их ортогонализацию и получим r_1 единичных взаимноортогональных собственных векторов $\{g_1, \ldots, g_{r_1}\}$. Поскольку различным собственным значениям отвечают ортогональные собственные векторы, то достаточно проделать такую же процедуру с собственным значением λ_2 кратности r_2 и т. д. В итоге мы имеем n единичных взаимно ортогональных собственных векторов $\{g_1, \ldots, g_{r_1}; h_1, \ldots, h_{r_2}; \ldots; k_1, \ldots, k_{r_s}\}$

$(r_1+r_2+\ldots+r_s=n)$, служащих базисом, в котором матрица оператора имеет диагональный вид (24.4).

Координаты собственных векторов $\{g_1, \ldots, k_{r_s}\}$, записанные в виде столбцов, будут составлять ортогональную матрицу перехода от первоначального ортонормированного базиса к каноническому ортонормированному базису.

Но заданной в ортонормированном базисе матрице самосопряженного оператора однозначно сопоставляется симметрическая билинейная форма (квадратичная форма), причем матрицы оператора и формы при переходе в другой ортонормированный базис преобразуются по одному и тому же закону. Поэтому теорему 1 можно записать в следующей эквивалентной форме (она-то и играет в приложениях главную роль).

Теорема 2. *Любую квадратичную форму* $B(x, x) = \sum_{i,k=1}^{n} \alpha_{ik}\xi^i\xi^k$ $(\alpha_{ik}=\alpha_{ki})$ *ортогональными преобразованиями всегда можно привести к следующему каноническому виду:*

$$B(x, x) = \lambda_1(\eta^1)^2+\ldots+\lambda_1(\eta^{r_1})^2+\ldots+\lambda_s(\eta^n)^2, \qquad (24.7)$$

где $\lambda_1, \ldots, \lambda_s$ — *корни характеристического уравнения* $\chi(\lambda)=\det(\alpha_{ik}-\lambda\delta_{ik})$, *встречающиеся столько раз, какова их кратность.*

Пример. Привести ортогональными преобразованиями к каноническому виду квадратичную форму: $2x_1x_2+2x_1x_3-2x_1x_4-2x_2x_3+2x_2x_4+2x_3x_4$ и найти матрицу перехода от данного базиса к каноническому.

Решение. Найдем характеристический полином матрицы квадратичной формы

$$\det(A-\lambda E)=\begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix} = (\lambda-1)^3(\lambda+3). \qquad (24.8)$$

Из (24.8) следует, что A имеет собственное значение $\lambda_1=1$ кратности 3 и однократное собственное значение $\lambda_2=-3$. Согласно теореме 2 по (24.7) имеем следующий канонический вид квадратичной формы:

$$A(x, x)=y_1^2+y_2^2+y_3^2-3y_4^2. \qquad (24.9)$$

Найдем ортогональное преобразование, осуществляющее приведение квадратичной формы к каноническому виду. С этой целью нам необходимо указать четыре единичных взаимно ортогональных собственных вектора матрицы A.

При $\lambda_1 = 1$ система уравнений $\sum_{q=1}^{4}(\alpha_{pq} - \lambda_1 \delta_{pq})\xi^q = 0$, соответствующая данному примеру, выглядит так:

$$\left.\begin{aligned} -\xi^1 + \xi^2 + \xi^3 - \xi^4 = 0 \\ \xi^1 - \xi^2 - \xi^3 + \xi^4 = 0 \\ \xi^1 - \xi^2 - \xi^3 + \xi^4 = 0 \\ -\xi^1 + \xi^2 + \xi^3 - \xi^4 = 0 \end{aligned}\right\}. \qquad (24.10)$$

Ранг матрицы системы равен 1 (уравнения 2, 3, 4 совпадают с первым). Поэтому

$$\xi^1 = c^2 + c^3 - c^4,\ \xi^2 = c^2,\ \xi^3 = c^3,\ \xi^4 = c^4.$$

Строим нормальную фундаментальную систему решений. Имеем

$$f_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},\quad f_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix},\quad f_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Применяя процесс ортогонализации и нормировки, придем к следующим единичным взаимно ортогональным векторам:

$$g_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix},\quad g_2 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \\ 0 \end{bmatrix},\quad g_3 = \begin{bmatrix} -\frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{\sqrt{3}}{2} \end{bmatrix}.$$

При $\lambda_2 = -3$ система уравнений

$$\sum_{q=1}^{4}(\alpha_{pq} - \lambda_2 \delta_{pq})\xi^q = 0$$

имеет ранг, равный трем, и нормальная фундаментальная система решений состоит из одного вектора

$$f_4 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

Нормируя собственный вектор f_4, получим единичный собственный вектор, ортогональный к g_1, g_2, g_3, так как $\lambda_1 \neq \lambda_2$

$$g_4 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

Таким образом, ортогональная матрица перехода от старого базиса к каноническому имеет следующий вид:

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Лекция 25

ПРИВЕДЕНИЕ В E_n ОБЩЕГО УРАВНЕНИЯ ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА К КАНОНИЧЕСКОМУ ВИДУ. НЕВЫРОЖДЕННЫЕ ЦЕНТРАЛЬНЫЕ И НЕЦЕНТРАЛЬНЫЕ ПОВЕРХНОСТИ. ЦИЛИНДРЫ

В этой лекции мы вернемся к рассмотрению векторно-точечных пространств E_n. В предыдущих двух лекциях мы ограничивались в основном рассмотрением векторных евклидовых пространств, поскольку теория операторов затрагивает лишь векторную структуру пространства.

Определение. Поверхностью второго порядка в E_n называется геометрическое место точек M, прямоугольные координаты $\xi^1; \ldots; \xi^n$ которых удовлетворяют уравнению вида

$$\sum_{i,k=1}^{n} \alpha_{ik}\xi^i\xi^k + 2\sum_{i=1}^{n} \varphi_i\xi^i + \gamma = 0, \tag{25.1}$$

где $A(x, x) = \sum_{i,k=1}^{n} \alpha_{ik}\xi^i\xi^k$ — квадратичная форма от радиус-вектора x, $\varphi(x) = \sum_{i=1}^{n} \varphi_i\xi^i$ — линейная форма от x, γ — постоянная, ранг $[\alpha_{ik}] \neq 0$.

Нашей задачей является выбор в E_n новой прямоугольной декартовой системы координат, в которой поверхность второго порядка описывалась бы наиболее простым уравнением, называемым *каноническим*.

По каноническому уравнению легко изучать свойства этих поверхностей.

Первый этап. Совершим ортогональное преобразование базиса, оставляя начало координат на месте, и перейдем к базису, в котором квадратичная форма $A(x, x)$ принимает канонический вид. Пусть ранг формы равен r и $\lambda_1, \ldots, \lambda_r$ — собственные значения матрицы $A = [\alpha_{ik}]$, отличные от нуля (среди них совпадающие также занумерованы). В результате получим новое уравнение поверхности

$$\lambda_1 (\eta^1)^2 + \dots + \lambda_r (\eta^r)^2 + 2 \sum_{l=1}^{n} \beta_l \eta^l + \gamma = 0, \tag{25.2}$$

которое может быть записано в следующем виде:

$$\sum_{s=1}^{r} \lambda_s \left(\eta^s + \frac{\beta_s}{\lambda_s}\right)^2 + 2 \sum_{p=r+1}^{n} \beta_p \eta^p + \gamma - \sum_{s=1}^{r} \frac{\beta_s^2}{\lambda_s} = 0. \tag{25.3}$$

Второй этап. Совершим перенос начала координат в точку $O'\left(-\frac{\beta_1}{\lambda_1}; \dots; -\frac{\beta_r}{\lambda_r}; 0; \dots 0\right)$

$$\eta^{s'} = \eta^s + \frac{\beta_s}{\lambda_s} \; (s = \overline{1, r}), \; \eta^{p'} = \eta^p \; (p = \overline{r+1, n}). \tag{25.4}$$

После этого уравнение поверхности примет следующий вид:

$$\lambda_1 (\eta^{1'})^2 + \dots + \lambda_r (\eta^{r'})^2 + 2\beta_{r+1} \eta^{r+1'} + \dots + 2\beta_n \eta^{n'} + \tilde{\gamma} = 0. \tag{25.5}$$

Определение. Поверхность называется *центральной*, если в ее уравнении (25.5) коэффициенты линейной формы $\beta_{r+1}, \dots, \beta_n$ равны нулю, причем — *невырожденной центральной*, когда $r = n$ и *вырожденной центральной*, когда $r < n$. Если же хотя бы одно из чисел β_p $(p = \overline{r+1, n})$ отлично от нуля, то поверхность называется *нецентральной*.

Такое определение связано с тем, что точка O' является единственным центром симметрии для поверхностей с уравнением

$$\lambda_1 (\eta^{1'})^2 + \dots + \lambda_r (\eta^{r'})^2 + \tilde{\gamma} = 0, \tag{25.6}$$

поскольку отражения $\eta^{k'} \to -\eta^{k'}$ не меняют вида данного уравнения.

Для нецентральных поверхностей продолжим процесс выбора новой прямоугольной системы координат с целью дальнейшего упрощения вида уравнения нецентральной поверхности (вместе с этим объяснится и тот факт, почему поверхность названа нецентральной).

Третий этап. (Для нецентральных поверхностей). Начало координат O' оставляем на месте и совершим следующее преобразование (новые координаты обозначим буквой ζ):

$$\left.\begin{aligned}
&\zeta^s = \eta^{s'} \quad (s = \overline{1, r}) \\
&\zeta^{r+1} = \alpha\,(\beta_{r+1}\,\eta^{r+1'} + \ldots + \beta_n\,\eta^{n'}) \\
&\zeta^{r+2} = \sum_{p=r+1}^{n} q_p^{r+2}\,\eta^{p'} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\zeta^{n} = \sum_{p=r+1}^{n} q_p^{n}\,\eta^{p'}
\end{aligned}\right\} \tag{25.7}$$

$$\left(\alpha = -\frac{1}{\sqrt{\beta_{r+1}^2 + \ldots + \beta_n^2}}\right),$$

подобрав числа $q_p^{r+2}, \ldots, q_p^n$ $(p = \overline{r+1, n})$ так, чтобы матрица преобразования была ортогональной. Покажем, что это всегда может быть осуществлено.

Действительно, матрицы-столбцы $[\zeta]$ из новых координат и $[\eta]'$ из старых координат связаны с матрицей перехода T соотношением $[\zeta] = T^{-1}[\eta]'$. Но для ортогональных преобразований $T^{-1} = T^{tr}$, и поэтому для (25.7) матрица T^{tr} имеет вид

$$T^{tr} = \begin{bmatrix}
1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \ldots & 0 & \alpha\beta_{r+1} & \alpha\beta_{r+2} & \ldots & \alpha\beta_n \\
0 & 0 & \ldots & 0 & q_{r+1}^{r+2} & q_{r+2}^{r+2} & \ldots & q_n^{r+2} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \ldots & 0 & q_{r+1}^{n} & q_{r+2}^{n} & \ldots & q_n^{n}
\end{bmatrix}, \tag{25.8}$$

где $q_p^{r+2}, \ldots, q_p^n$ $(p = \overline{r+1, n})$ — неизвестные пока элементы матрицы.

Очевидно, для того, чтобы матрица T была ортогональной, необходимо в E_{n-r} найти $(n - r - 1)$ единичных взаимно-ортогональных векторов, ортогональных к вектору x_0 с координатами $\alpha\beta_{r+1}, \ldots, \alpha\beta_n$, и их координаты подставить вместо $q_{r+1}^{r+2}, \dots, q_n^{r+2}, \dots, q_{r+1}^n, \dots$ $\dots, q_n^n$. Для этого строим векторы нормальной фундаментальной системы решений для системы уравнений, состоящей из одного уравнения

$$(x_0, x) = \sum_{p=r+1}^{n} (\alpha\beta_p) \xi^p = 0.$$

Таких векторов будет $(n-r-1)$ штук. Затем для них проводим процесс ортогонализации и нормировки. Тем самым неизвестные числа $q_p^{r+2}, \dots, q_p^n$ $(p = \overline{r+1, n})$ будут определены, и ортогональное преобразование (25.7) принимает конкретный вид. После преобразования (25.7) уравнение (25.5) поверхности перейдет в следующее:

$$\lambda_1 (\zeta^1)^2 + \dots + \lambda_r (\zeta^r)^2 + \frac{2}{\alpha} \zeta^{r+1} + \tilde{\gamma} = 0. \qquad (25.9)$$

Четвертый этап состоит в переносе начала координат

$$\xi^{1'} = \zeta^1, \dots, \xi^{r'} = \zeta^r, \ \xi^{r+1'} = \zeta^{r+1} + \frac{\alpha\tilde{\gamma}}{2},$$
$$\xi^{r+2'} = \zeta^{r+2}, \dots, \xi^{n'} = \zeta^n,$$

после которого каноническое уравнение нецентральной поверхности примет окончательный вид

$$\lambda_1 (\xi^{1'})^2 + \dots + \lambda_r (\xi^{r'})^2 = 2\delta\xi^{r+1'}, \qquad (25.10)$$

где $\delta = \sqrt{\beta_{r+1}^2 + \dots + \beta_n^2}$.

Если в (25.10) $r+1 = n$, то нецентральная поверхность называется *невырожденной*, если же $r+1 < n$ — то вырожденной.

Переходим к конкретной классификации невырожденных центральных и нецентральных поверхностей второго порядка, учитывающей знаки собственных значений.

Невырожденные центральные поверхности.

Каноническое уравнение имеет вид (25.6), где $r = n$. В зависимости от того, равен нулю свободный член $\tilde{\gamma}$ или отличен от нуля, все невырожденные центральные поверхности разбиваются на 2 типа. Рассмотрим каждый в отдельности.

(I) $\tilde{\gamma} \neq 0$. Поделим левую часть уравнения (25.6) на $\tilde{\gamma}$ и переобозначим переменные так, чтобы сначала шли те, у которых коэффициенты положительные, а затем те, у которых коэффициенты отрицательные. В результате получим для первого типа следующий вид уравнения центральной поверхности:

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_n^2}{a_n^2} = 1. \qquad (25.11)$$

Имеем n различных классов невырожденных центральных поверхностей первого типа при $k = 1, 2, \ldots$, (при $k = 0$ имеем мнимую поверхность 2-го порядка).

Ограничимся рассмотрением частных случаев $n = 2$ и $n = 3$. При $n = 2$ (на плоскости) имеем 2 центральные кривые второго порядка

$$k = 1: \ \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1 \quad \text{(гипербола)}$$

$$k = 2: \ \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1 \quad \text{(эллипс).}$$

При $n = 3$ (в пространстве) имеем 3 центральные поверхности 2-го порядка

$$k = 1: \ \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1 \quad \text{(двуполостной гиперболоид)}$$

$$k = 2: \ \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1 \quad \text{(однополостной гиперболоид)}$$

$$k = 3: \ \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1 \quad \text{(эллипсоид).}$$

Вид приведенных кривых и поверхностей подробно изучен нами в аналитической геометрии, и на этих вопросах мы не останавливаемся.

(II) $\tilde{\gamma} = 0$. Нумеруя переменные так, чтобы сначала шли переменные с положительными коэффициентами, а затем с отрицательными, мы получим уравнения для центральных невырожденных поверхностей второго типа

$$\frac{x_1^2}{a_1^2}+\ldots+\frac{x_k^2}{a_k^2}-\frac{x_{k+1}^2}{a_{k+1}^2}-\ldots-\frac{x_n^2}{a_n^2}=0. \qquad (25.12)$$

Такие поверхности относятся к *коническим*, так как наряду с точкой $M(x_1;\ldots;x_n)$ этому уравнению удовлетворяют точки с координатами $tx_1, tx_2, \ldots, tx_n$ при любом $t\in\mathbf{R}$, т. е. точки, лежащие на прямой, проходящей через M и начало координат. Сколько различных классов конических поверхностей описывает уравнение (25.12)? Чтобы ответить на этот вопрос, заметим, что k достаточно менять лишь в пределах от 0 до $\frac{n}{2}$, так как при $k>\frac{n}{2}$, умножая уравнение (25.12) на -1, придем к уравнению, в котором положительных слагаемых меньше $\frac{n}{2}$. Кроме того, при $k=0$ имеем мнимую поверхность с одной вещественной точкой $(0, 0, \ldots, 0)$. Итак, $0<k\leqslant\frac{n}{2}$, при этом мы должны учитывать, что k — целое число.

Опять-таки ограничимся рассмотрением частных случаев $n=2$ и $n=3$. Согласно нашим оценкам для k имеем лишь по одному классу в каждом случае.

При $n=2$: $\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}=0$ или $\left(\frac{x_1}{a_1}-\frac{x_2}{a_2}\right)\left(\frac{x_1}{a_1}+\frac{x_2}{a_2}\right)=0$ (пара пересекающихся прямых).

При $n=3$: $\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}=0$ (эллиптический конус с центральной осью x_1).

Невырожденные нецентральные поверхности.

Как и выше, поделив на δ и перенумеровав переменные в (25.10), получим уравнение вида

$$\frac{x_1^2}{a_1^2}+\ldots+\frac{x_k^2}{a_k^2}-\frac{x_{k+1}^2}{a_{k+1}^2}-\ldots-\frac{x_{n-1}^2}{a_{n-1}^2}=2x_n. \qquad (25.13)$$

Число различных классов поверхностей зависит от числа значений принимаемых индексом k. Очевидно, что $0 \leqslant k \leqslant \frac{n-1}{2}$, так как если $k > \frac{n-1}{2}$, то, умножая (25.13) на -1 и обозначая $-x_n$ через x'_n, мы придем опять же к уравнению вида (25.13), в котором положительных слагаемых (квадратичных) меньше $\frac{n-1}{2}$.

Обратимся к конкретным случаям $n=2$ и $n=3$. При $n=2$ имеем 1 класс нецентральных кривых 2-го порядка:

$$k=0:\ -\frac{x_1^2}{a_1^2}=2x_2 \text{ (парабола).}$$

При $n=3$ ($k=0$, $k=1$) имеем 2 класса нецентральных поверхностей второго порядка

$$k=0:\ -\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}=2x_3 \quad \text{(эллиптический параболоид)}$$

$$k=1:\ \frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}=2x_3 \quad \text{(гиперболический параболоид).}$$

Форма параболы и формы указанных поверхностей подробно изучались нами в аналитической геометрии, и на их построении мы не останавливаемся.

Сейчас остановимся на рассмотрении *вырожденных центральных и нецентральных поверхностей 2-го порядка.*

Поскольку поверхность вырожденная, то в канонических уравнениях (25.11), (25.12), (25.13) отсутствует хотя бы одна переменная, например, переменная x_1. В этом случае все сечения рассматриваемой поверхности гиперплоскостями $x_1 = \text{const}$ представляют собой одну и ту же поверхность в пространстве E_{n-1}. Это означает, что каждая вырожденная поверхность образуется за счет поступательного движения поверхности 2-го порядка, заданной в E_{n-1}, вдоль оси x_1, перпендикулярной к E_{n-1}. Следовательно, исследуемая поверхность необходимо содержит прямолинейные образующие параллельные оси x_1. Поэтому вырожденные поверхности второго порядка часто называют *цилиндрическими.*

Перейдем к конкретным частным случаям $n=2$ и $n=3$.

При $n=2$ согласно (25.11) и (25.12) все центральные вырожденные кривые описываются уравнением

$$\frac{x_2^2}{a_2^2}=\varepsilon, \tag{25.14}$$

где $\varepsilon=-1, 1, 0$.

При $\varepsilon=-1$ вещественных геометрических образов не существует. Когда $\varepsilon=1$, имеем *пару параллельных прямых*. При $\varepsilon=0$ — *пару слившихся прямых*.

Согласно (25.13), когда $n=2$, нецентральных вырожденных кривых второго порядка не имеется.

Чтобы построить все цилиндрические поверхности в E_3, необходимо подвергнуть поступательному движению вдоль оси x_3 все кривые 2-го порядка, лежащие в плоскости x_1Ox_2.

Учитывая приведенную выше классификацию всех невырожденных и вырожденных кривых 2-го порядка, получим следующие вещественные цилиндрические поверхности второго порядка:

(1) $\dfrac{x_1^2}{a_1^2}+\dfrac{x_2^2}{a_2^2}=1$ (эллиптический цилиндр)

(2) $\dfrac{x_1^2}{a_1^2}-\dfrac{x_2^2}{a_2^2}=1$ (гиперболический цилиндр)

(3) $\dfrac{x_1^2}{a_1^2}=2x_2$ (параболический цилиндр)

(4) $\dfrac{x_1^2}{a_1^2}-\dfrac{x_2^2}{a_2^2}=0$ (пара пересекающихся плоскостей)

(5) $\dfrac{x_1^2}{a_1^2}=1$ (пара параллельных плоскостей)

(6) $\dfrac{x_1^2}{a_1^2}=0$ (пара слившихся плоскостей).

В заключение проведем подробный разбор примера, связанного с приведением к каноническому виду общего уравнения кривой второго порядка на плоскости.

Пример. Требуется привести к каноническому виду уравнение кривой 2-го порядка $3x^2 + 10xy + 3y^2 - 2x - 14y - 13 = 0$ и указать преобразование координат.

Решение. Найдем собственные значения матрицы квадратичной формы

$$\det(A - \lambda E) = \begin{vmatrix} 3-\lambda & 5 \\ 5 & 3-\lambda \end{vmatrix} = \lambda^2 - 6\lambda - 16 = 0,$$

$$\lambda_1 = 8, \ \lambda_2 = -2.$$

Соответствующие им собственные векторы из-за $\lambda_1 \neq \lambda_2$ ортогональны между собой, и их остается только пронормировать. Имеем

$$g_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad g_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Следовательно, после преобразования координат, индуцированного переходом к базису с векторами g_1, g_2

$$\begin{cases} x = \frac{1}{\sqrt{2}} x' - \frac{1}{\sqrt{2}} y' \\ y = \frac{1}{\sqrt{2}} x' + \frac{1}{\sqrt{2}} y', \end{cases}$$

мы придем к уравнению $8(x')^2 - 2(y')^2 - 8\sqrt{2}x' - 6\sqrt{2}y' - 13 = 0$. Его можно преобразовать к следующему виду:

$$8\left(x' - \frac{1}{\sqrt{2}}\right)^2 - 2\left(y' + \frac{3}{\sqrt{2}}\right)^2 - 13 - 4 + 9 = 0.$$

После переноса начала координат: $x' = X + \frac{1}{\sqrt{2}}$, $y' = Y - \frac{3}{\sqrt{2}}$, получим следующее уравнение:

$$8X^2 - 2Y^2 = 8,$$

которое в канонической записи примет вид

$$X^2 - \frac{Y^2}{4} = 1. \tag{25.15}$$

Преобразование, связывающее координаты x, y и X, Y, легко устанавливается

$$\begin{cases} x = X\cos\frac{\pi}{4} - Y\sin\frac{\pi}{4} + 2 \\ y = X\sin\frac{\pi}{4} + Y\cos\frac{\pi}{4} - 1. \end{cases}$$

Таким образом, в новой системе координат, которая получена из старой поворотом на угол $\frac{\pi}{4}$ (против часовой стрелки) и переносом начала координат в точку O' $(2; -1)$, мы имеем гиперболу (25.15) с вещественной осью $O'X$ и асимптотами $Y = \pm 2X$.

На этом мы закончим рассмотрение вещественных евклидовых пространств и перейдем к изучению специальных комплексных линейных пространств, свойства которых очень близки к свойствам вещественных собственно евклидовых пространств. Такие пространства называются унитарными. Теория операторов в указанных пространствах находит широкое приложение в квантовой механике.

Лекция 26

ЭРМИТОВЫ И ЭРМИТОВО-КВАДРАТИЧНЫЕ ФОРМЫ В КОМПЛЕКСНОМ ЛИНЕЙНОМ ПРОСТРАНСТВЕ. УНИТАРНОЕ n-МЕРНОЕ ПРОСТРАНСТВО. УНИТАРНЫЕ ОПЕРАТОРЫ. ЭРМИТОВЫ ОПЕРАТОРЫ И ИХ СВОЙСТВА

На протяжении данной лекции мы будем рассматривать комплексные линейные пространства ($\mathbf{K} \equiv \mathbf{C}$). Нашей целью является введение таких комплексных линейных пространств, которые были бы близки по своим свойствам к собственно евклидовым вещественным пространствам. Евклидовы комплексные линейные пространства для этой цели не годятся, так как для них скалярное произведение (x, x) может обращаться в нуль не только для нулевых векторов, в то время как для собственно евклидовых вещественных пространств $(x, x) > 0$ для любого $x \neq 0$ и $(x, x) = 0$ тогда и только тогда, когда $x=0$. Чтобы подойти к решению поставленной задачи, в линейном комплексном пространстве мы будем помимо билинейных форм рассматривать эрмитовы формы.

Определение. Числовая функция $H(x, y)$ называется *эрмитовой формой*, если для любых $x, y, z \in \mathscr{L}$ и любого $\alpha \in \mathbf{C}$ выполнено

$$\left.\begin{aligned} &(1)\ H(x+z, y) = H(x, y) + H(z, y) \\ &(2)\ H(\alpha x, y) = \alpha H(x, y) \\ &(3)\ H(x, y+z) = H(x, y) + H(x, z) \\ &(4)\ H(x, \alpha y) = \bar{\alpha} H(x, y) \end{aligned}\right\}, \tag{26.1}$$

где $\bar{\alpha}$ — комплексно-сопряженное число по отношению к α.

Очевидно, что первые 3 условия из (26.1) совпадают с условиями для билинейных форм, а четвертое условие для эрмитовых форм отличается от четвертого условия для билинейных форм (там $B(x, \alpha y) = \alpha B(x, y)$).

Из (26.1) тотчас следует (по индукции)

$$H\left(\sum_{k=1}^{p} \alpha_k x_k, \sum_{m=1}^{q} \beta_m y_m\right) = \sum_{k=1}^{p} \sum_{m=1}^{q} \alpha_k \bar{\beta}_m H(x_k, y_m) \tag{26.2}$$

для любых $x_k, y_m \in \mathscr{L}$ и любых $\alpha_k, \beta_m \in \mathbf{C}$.

Пусть $\{e_1, \ldots, e_n\}$ — базис $\mathscr{L}_n$ и

$$x = \sum_{k=1}^{n} \xi^k e_k, \quad y = \sum_{m=1}^{n} \eta^m e_m.$$

Тогда, согласно (26.2), имеем

$$H(x, y) = \sum_{k, m=1}^{n} \xi^k \bar{\eta}^m H(e_k, e_m) = \sum_{k, m=1}^{n} h_{km} \xi^k \bar{\eta}^m, \tag{26.3}$$

где h_{km} — значения эрмитовой формы на векторах базиса, называемые *коэффициентами эрмитовой формы* относительно данного базиса. Из них можно построить по обычному правилу квадратную матрицу $H = [h_{km}]$, называемую *матрицей эрмитовой формы*.

Если перейти к другому базису $e_{k'} = \sum_{i=1}^{n} \tau_{k'}^{i} e_i$ с матрицей перехода $T = [\tau_{k'}^{i}]$, то

$$h_{k'm'} = H(e_{k'}, e_{m'}) = \sum_{i, j=1}^{n} \tau_{k'}^{i} \bar{\tau}_{m'}^{j} H(e_i, e_j) = \sum_{i, j=1}^{n} h_{ij} \tau_{k'}^{i} \bar{\tau}_{m'}^{j}. \tag{26.4}$$

В матричном представлении равенство (26.4) имеет вид

$$H' = T^{tr} H \overline{T}. \tag{26.5}$$

В силу невырожденности матрицы перехода и очевидного равенства $\det \overline{T} = \overline{\det T}$ из (26.5) следует, что *ранг r матрицы эрмитовой формы есть инвариант преобразования базисов.* Если $r = n$, то эрмитова форма называется *невырожденной.*

Для приложений нам необходимы не общие эрмитовы формы, а *симметричные эрмитовы формы.*

Определение. Эрмитова форма называется *симметричной*, если $H(x, y) = \overline{H(y, x)}$.

Исходя из этого определения мы получим, что

$$h_{km} = H(e_k, e_m) = \overline{H(e_m, e_k)} = \overline{h}_{mk}, \tag{26.6}$$

и, следовательно, *матрица всякой симметричной эрмитовой формы обладает свойством $H = \overline{H}^{tr}$.*

Определение. Матрица, которая после операции транспонирования и замены элементов на комплексно-сопряженные переходит в себя, называется *эрмитовой матрицей.*

Для сокращения записи обе проводимые операции обозначают символом „ $*$ “, и поэтому эрмитова матрица может быть определена как всякая матрица, обладающая свойством $A = A^*$.

Например, матрица $A = \begin{bmatrix} 1 & 1+i \\ 1-i & 1 \end{bmatrix}$ является эрмитовой, в чем легко убедиться, если i заменить на $-i$ и заменить строки столбцами.

В силу (26.6) *матрица симметричной эрмитовой формы необходимо эрмитова.*

Обратно, *если матрица эрмитовой формы эрмитова, то данная эрмитова форма является симметричной.* В самом деле,

$$H(x, y) = \sum_{i, k=1}^{n} h_{ik} \xi^i \overline{\eta}^k = \sum_{i, k=1}^{n} \overline{h}_{ki} \overline{\eta}^k \xi^i = $$

$$= \sum_{i, k=1}^{n} \overline{h_{ki} \eta^k \overline{\xi}^i} = \overline{H(y, x)}.$$

Таким образом, *между эрмитовыми матрицами и симметричными эрмитовыми формами установлено взаимно-однозначное соответствие.*

Отметим, что в вещественном случае эрмитова форма переходит в билинейную вещественную форму, симметричная эрмитова форма—в симметричную билинейную форму, а эрмитова матрица — в симметрическую матрицу. Эти факты автоматически следуют из (26.1), (26.3), (26.6). Но в отличие от вещественного линейного пространства в комплексном линейном пространстве можно определить 2 вида квадратичных форм — *эрмитово-квадратичную форму и симметричную эрмитово-квадратичную форму*.

Определение. *Эрмитово-квадратичной формой* в комплексном пространстве $\mathscr{L}$ называется функция аргумента $x \in \mathscr{L}$, получаемая из эрмитовой формы $H(x, y)$ заменой y на x.

Ее общий вид следует из (26.3)

$$H(x, x) = \sum_{k, m=1}^{n} h_{km} \xi^k \bar{\xi}^m. \tag{26.7}$$

Определение. *Симметричной эрмитово-квадратичной формой* называется эрмитово-квадратичная форма, полученная на основе симметричной эрмитовой формы.

Ее общий вид дается формулой (26.7), но с условием — матрица $H = [h_{km}]$ должна быть эрмитовой. Учитывая взаимно-однозначное соответствие между эрмитовыми матрицами и симметричными эрмитовыми формами, мы можем заключить, что *между симметричными эрмитово-квадратичными формами и эрмитовыми матрицами также существует взаимно-однозначное соответствие*. Для приложений основную роль играют симметричные эрмитово-квадратичные формы, поскольку по своим свойствам они близки к квадратичным формам в вещественном пространстве $\mathscr{L}$. Поэтому в дальнейшем под символом $H(x, x)$ понимается только симметричная эрмитово-квадратичная форма.

Первое свойство, которое мы сразу же отметим, заключается в том, что *симметричная эрмитово-квадратичная форма принимает только вещественные значения*, поскольку из равенства $H(x, y) = \overline{H(y, x)}$ тотчас следует

$$H(x, x) = \overline{H(x, x)}. \tag{26.8}$$

Для симметричных эрмитово-квадратичных форм имеют место теоремы, аналогичные теоремам, доказанным для квадратичных форм в вещественном линейном пространстве.

Теорема 1. *В комплексном пространстве $\mathcal{L}_n$ существует базис, в котором симметричная эрмитово-квадратичная форма имеет канонический вид*

$$H(x, x)=\sum_{k=1}^{n} \alpha_k |\xi^{k'}|^2, \tag{26.9}$$

где α_k — вещественные числа, среди которых r чисел отлично от нуля ($r =$ ранг H).

Доказательство этой теоремы совпадает почти дословно с доказательством теоремы о приведении вещественной квадратичной формы к каноническому виду методом Лагранжа (необходимо только учитывать эрмитовость матрицы $H=[h_{km}]$, у которой из-за условия $h_{km}=\overline{h}_{mk}$ диагональные элементы ($k=m$) необходимо вещественные числа).

В самом деле, симметричную эрмитово-квадратичную форму можно всегда представить в виде (считаем, не ограничивая общности рассуждений, что $h_{11}\neq 0$)

$$H(x, x)=\frac{1}{a_{11}}[a_{11}\xi^1+(a_{12}-ib_{12})\xi^2+\ldots+(a_{1n}-ib_{1n})\xi^n]\times$$
$$\times[a_{11}\overline{\xi}^1+(a_{12}+ib_{12})\overline{\xi}^2+\ldots+(a_{1n}+ib_{1n})\overline{\xi}^n]+$$
$$+\sum_{k, m=2}^{n} h'_{km}\xi^k\overline{\xi}^m,$$

где $a_{11}=h_{11}$ — вещественное число, $a_{12}+ib_{12}=h_{12}$, $a_{13}+ib_{13}=h_{13}, \ldots, a_{1n}+ib_{1n}=h_{1n}$.

В качестве первого шага совершаем преобразование координат

$$\begin{cases} \xi^{1'}=a_{11}\xi^1+(a_{12}-ib_{12})\xi^2+\ldots+(a_{1n}-ib_{1n})\xi^n \\ \xi^{p'}=\xi^p \quad (p=\overline{2, n}), \end{cases}$$

после которого форма $H(x, x)$ записывается следующим образом:

$$H(x, x)=\frac{1}{a_{11}}\xi^{1'}\overline{\xi}^{1'}+\sum_{k, m=2}^{n} h'_{km}\xi^{k'}\overline{\xi}^{m'}. \tag{26.10}$$

Применяя данный алгоритм к форме $\sum\limits_{k,\,m=2}^{n} h'_{km}\xi^{k'}\xi^{m'}$, через конечное число шагов придем к каноническому виду (26.9).

Теорема 2. *Число s положительных коэффициентов и число $(r-s)$ отрицательных коэффициентов в наборе $\alpha_1, \ldots, \alpha_n$ из (26.9) не зависит от выбора канонического базиса.*

Доказательство теоремы полностью повторяет доказательство теоремы инерции квадратичных форм в вещественном случае.

Определение. Симметричная эрмитово-квадратичная форма называется *положительно определенной*, если $H(x, x) > 0$ для любого $x \neq 0$.

Из канонического вида (26.9) и теоремы 2 вытекает, что *симметричная эрмитово-квадратичная форма положительно определена тогда и только тогда, когда $s = r = n$.* Таким образом, *положительно определенная симметричная эрмитово-квадратичная форма всегда является невырожденной формой.*

Заметим, что детерминант любой эрмитовой матрицы есть вещественное число, поскольку из равенства $A = \bar{A}^{tr}$ следует, что $\det A = \overline{\det A^{tr}} = \overline{\det A}$. Из закона преобразования (26.5) для эрмитовой матрицы имеем

$$\det H' = |\det T|^2 \det H.$$

Поскольку $\det H$ и $\det H'$ — вещественные числа, то *знак детерминанта матрицы симметричной эрмитово-квадратичной формы есть инвариант преобразования базисов.* Вследствие этого факта мы можем доказать критерий Сильвестра для положительно определенных симметричных эрмитово-квадратичных форм, повторяя из слова в слово доказательство критерия Сильвестра в случае вещественных квадратичных форм: *форма $H(x, x)$ положительно определена тогда и только тогда, когда все главные миноры*

$$h_{11}, \quad \begin{vmatrix} h_{11} & h_{12} \\ \bar{h}_{12} & h_{22} \end{vmatrix}, \ldots, \begin{vmatrix} h_{11} & h_{12} \ldots h_{1n} \\ \bar{h}_{12} & h_{22} \ldots h_{2n} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot \\ \bar{h}_{1n} & \bar{h}_{2n} \ldots h_{nn} \end{vmatrix}$$

строго положительны.

Для положительно определенной формы $H(x, x)$ канонический вид (26.9) упрощается, поскольку все $\alpha_k > 0$ $(k = \overline{1, n})$. За счет преобразования $\xi^k = \sqrt{\alpha_k}\,\xi^{k'}$ $H(x, x)$ примет окончательный канонический вид

$$H(x, x) = \sum_{k=1}^{n} |\xi^k|^2 = \sum_{i,\,k=1}^{n} \delta_{ik}\xi^i \bar{\xi}^k, \tag{26.11}$$

т. е. *в каноническом базисе матрица положительно определенной симметричной эрмитово-квадратичной формы совпадает с единичной матрицей.*

Сейчас мы можем перейти к определению *унитарного пространства*, которое является обобщением собственно евклидова пространства на случай линейных комплексных пространств.

На первый взгляд кажется естественным в качестве скалярного произведения векторов с комплексными координатами $\xi^1, \ldots, \xi^n$ и $\eta^1, \ldots, \eta^n$ взять обычное выражение $\sum_{k=1}^{n} \xi^k\eta^k$. В определенных случаях так и поступают. Пространство, которое при этом получается, называется *комплексным евклидовым пространством.* Оно в этом случае, как уже выше указывалось, теряет ряд свойств, и среди них важнейшее свойство $(x, x) > 0$ для любого $x \neq 0$. По этой причине комплексное евклидово пространство по своим свойствам отстоит очень далеко от вещественных евклидовых пространств, изученных нами в предыдущих лекциях. Поэтому в комплексном линейном пространстве необходимо так ввести операцию скалярного произведения, чтобы свойство $(x, x) > 0$ (для любого $x \neq 0$) сохранилось в данном случае.

Но таким свойством, как нами установлено, обладает симметричная положительно определенная эрмитово-квадратичная форма. В связи с этим мы примем следующее определение.

Определение. Комплексное линейное пространство $\mathscr{L}$ называется *унитарным*, если в $\mathscr{L}$ зафиксирована некоторая симметричная эрмитова форма такая, что соответствующая ей симметричная эрмитово-квадратичная форма является положительно определенной. При этом данная симметричная эрмитова форма называется *скалярным произведением* векторов x, y и обозначается символом (x, y).

Учитывая свойства симметричных эрмитовых форм, мы получим, что скалярное произведение обладает следующими свойствами:

$$\text{(а) } (x, y) = \overline{(y, x)}; \quad \text{(б) } (x, y + z) = (x, y) + (x, z);$$

$$\text{(в) } (\lambda x, y) = \lambda (x, y); \quad \text{(г) } (x, x) > 0 \ (\forall x \neq 0). \qquad (26.12)$$

Если в унитарном пространстве выбран общий базис $\{e_1, \ldots, e_n\}$, то

$$(x, y) = \sum_{k, m=1}^{n} h_{km} \xi^k \bar{\eta}^m \ \left(\text{для} \quad \sum_{k, m=1}^{n} h_{km} \xi^k \bar{\xi}^m > 0\right). \qquad (26.13)$$

Длиной вектора x, как и в вещественном случае, назовем величину $|x| = \sqrt{(x, x)}$. Вектор, длина которого равна единице, называется *нормированным* (или единичным). Очевидно, каждый вектор $x \neq 0$ в унитарном пространстве можно нормировать, умножая его на $\lambda = \frac{1}{|x|}: x^0 = \frac{x}{|x|}$.

Определение. Говорят, что вектор x *ортогонален* к y, если $(x, y) = 0$.

По свойству (26.12а) в этом случае и $(y, x) = 0$, т. е. y также ортогонален к x.

Теорема. *В любом n-мерном унитарном пространстве существует ортонормированный базис.*

Действительно, согласно (26.11) для положительно определенной симметричной эрмитово-квадратичной формы (x, x) существует такой канонический базис $\{i_1, \ldots, i_n\}$, в котором $(x, x) = \sum_{k, m=1}^{n} \delta_{km} \xi^k \bar{\xi}^m$, где $\delta_{km} = (i_k, i_m)$. Поэтому симметричная эрмитова форма (x, y) в этом базисе $\{i_1, \ldots, i_n\}$ будет иметь вид

$$(x, y) = \sum_{k, m=1}^{n} \delta_{km} \xi^k \bar{\eta}^m = \sum_{k=1}^{n} \xi^k \bar{\eta}^k, \qquad (26.14)$$

где $(i_k, i_m) = \delta_{km} = \begin{cases} 1, & k = m \\ 0, & k \neq m \end{cases}$. Теорема доказана.

Отметим, что процесс ортогонализации системы векторов полностью повторяется как и в вещественном случае.

Определение. Линейное преобразование в n-мерном унитарном пространстве, соответствующее переходу от одного ортонормированного базиса к другому ортонормированному базису, называется *унитарным преобразованием.*

Если $\{i_1, \ldots, i_n\}$ и $\{i_{1'}, \ldots, i_{n'}\}$ — ортонормированные базисы в унитарном пространстве с матрицей перехода $U=[u^m_{k'}]$

$$i_{k'}=\sum_{m=1}^{n} u^m_{k'} i_m,$$

то в силу единичности и ортогональности векторов $i_{1'}, \ldots, i_{n'}$ имеем

$$\delta_{km}=(i_{k'}, i_{m'})=\left(\sum_{p=1}^{n} u^p_{k'} i_p, \quad \sum_{q=1}^{n} u^q_{m'} i_q\right)=$$

$$=\sum_{p,\,q=1}^{n} u^p_{k'} \bar{u}^q_{m'} (i_p, i_q)=\sum_{p,\,q=1}^{n} u^p_{k'} \bar{u}^q_{m'} \delta_{pq}=\sum_{p=1}^{n} u^p_{k'} \bar{u}^p_{m'}. \qquad (26.15)$$

Равенство (26.15), если ввести в рассмотрение транспонированную матрицу $\bar{U}^{tr}=[\bar{u}^{m'}_{k}]$, может быть записано в виде

$$U^* U=E. \qquad (26.16)$$

Всякая матрица U, удовлетворяющая условию (26.16), называется *унитарной.*

Таким образом, *в унитарном пространстве переход от одного ортонормированного базиса к другому ортонормированному базису совершается всегда с помощью унитарной матрицы перехода.*

Обратно, если задана унитарная матрица $U=[u^m_{k'}]$, то в силу (26.16) она *невырожденная* ($|\det U|^2=1$), и она может служить матрицей перехода. Пусть $e_{k'}=\sum_{m=1}^{n} u^m_{k'} i_m$ — векторы нового базиса. Очевидно, что с учетом (26.16) имеем

$$(e_{k'}, e_{m'})=\left(\sum_{p=1}^{n} u^p_{k'} i_p, \quad \sum_{q=1}^{n} u^q_{m'} i_q\right)=\sum_{p=1}^{n} u^p_{k'} \bar{u}^p_{m'}=\delta_{km}.$$

Новый базис $\{e_{1'}, \ldots, e_{n'}\}$ — ортонормированный базис.

Другими словами, *между вращениями базиса в унитарном пространстве и унитарными матрицами установлено взаимно-однозначное соответствие.*

Легко заметить, что унитарные матрицы обобщают понятие ортогональных матриц в вещественном пространстве E_n. Из условия (26.16) тотчас следует эквивалентное равенство

$$U^* = U^{-1}, \tag{26.17}$$

которое часто принимают за *новое определение унитарной матрицы*. Унитарный оператор, как и изометрический в E_n, сохраняет скалярное произведение векторов в унитарном пространстве. В самом деле, согласно определению унитарного преобразования, имеем:

$$(\boldsymbol{U}x, \boldsymbol{U}y) = \sum_{k, m=1}^{n} \xi^k \overline{\eta}^m (\boldsymbol{U}i_k, \boldsymbol{U}i_m) = \sum_{k, m=1}^{n} \xi^k \overline{\eta}^m (f_k, f_m),$$

где $(f_k, f_m) = \delta_{km}$, поскольку унитарное преобразование переводит ортонормированный базис в ортнормированный. Таким образом

$$(\boldsymbol{U}x, \boldsymbol{U}y) = \sum_{k, m=1}^{n} \xi^k \overline{\eta}^m \delta_{km} = \sum_{k=1}^{n} \xi^k \overline{\eta}^k = (x, y).$$

Поэтому унитарный оператор называют также *изометрическим*.

Определение. Самосопряжённый оператор $\boldsymbol{A}$, действующий в унитарном пространстве, называется *эрмитовым оператором*.

Какова матрица самосопряженного оператора в ортонормированном базисе? Согласно определению самосопряженности ($(\boldsymbol{A}x, y) = (x, \boldsymbol{A}y)$ для любых векторов x и y), учитывая, что $(\boldsymbol{A}i_k, i_m) = \sum_{s=1}^{n} \alpha_k^s (i_s, i_m) = \alpha_k^m$ и $(i_k, \boldsymbol{A}i_m) = \sum_{s=1}^{n} \overline{\alpha}_m^s (i_k, i_s) = \overline{\alpha}_m^k$, имеем $\alpha_k^m = \overline{\alpha}_m^k$. Итак, *в ортонормированном базисе матрица самосопряженного оператора является эрмитовой матрицей.*

Обратно, *если оператор $\boldsymbol{A}$ в ортонормированном базисе имеет эрмитову матрицу, то он необходимо самосопряженный оператор*. В самом деле,

$$(\boldsymbol{A}x, y) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \alpha_i^k \xi^i \right) \overline{\eta}^k = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \overline{\alpha}_k^i \overline{\eta}^k \right) \xi^i =$$

$$= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \alpha_k^i \eta^k \right) \bar{\xi}^i = \overline{(Ay, x)} = (x, Ay).$$

Таким образом, *между эрмитовыми операторами и эрмитовыми матрицами существует взаимно-однозначное соответствие.*

Этим свойством и объясняется термин „эрмитов оператор" для самосопряженного оператора в унитарном пространстве.

В то же время существует взаимно-однозначное соответствие между эрмитовыми матрицами и симметричными эрмитовыми формами (эквивалентно симметричными эрмитово-квадратичными формами). Поэтому между эрмитовыми операторами и соответствующими симметричными эрмитовыми формами также может быть установлено взаимно-однозначное соответствие.

Действительно, пусть $\{i_1, \dots, i_n\}$ — ортонормированный базис в унитарном пространстве. Пусть $H(x, y) = \sum_{k, m=1}^{n} h_{km} \xi^k \bar{\eta}^m$ — симметричная эрмитова форма ($h_{km} = \bar{h}_{mk}$). Построим матрицу $A = [\alpha_m^k]$, положив $\alpha_m^k = h_{mk}$ (т. е. $A = H^{tr}$). Из-за условия $h_{km} = \bar{h}_{mk}$ матрица A эрмитова, и она соответствует эрмитову оператору $\boldsymbol{A}$. Подсчитаем далее $(x, \boldsymbol{A}y)$ и устанавливаем следующую связь оператора и формы $H(x, y)$:

$$(x, \boldsymbol{A}y) = \sum_{k=1}^{n} \xi^k \overline{\left(\sum_{m=1}^{n} \alpha_m^k \eta^m \right)} = \sum_{m=1}^{n} \left(\sum_{k=1}^{n} \alpha_k^m \xi^k \right) \bar{\eta}^m =$$

$$= \sum_{k, m=1}^{n} h_{km} \xi^k \bar{\eta}^m = H(x, y). \tag{26.18}$$

Эрмитовы операторы, как и симметричные операторы в вещественном E_n, обладают рядом замечательных свойств.

Теорема 1. *Собственные значения эрмитова оператора — вещественные числа.*

В самом деле, если λ_0 — собственное значение эрмитова оператора $\boldsymbol{A}$ и x — соответствующий собственный вектор оператора, то $(\boldsymbol{A}x, x) = (\lambda_0 x, x) = \lambda_0 (x, x)$ и $(x, \boldsymbol{A}x) = (x, \lambda_0 x) = \bar{\lambda}_0 (x, x)$. В силу самосопряженности $\lambda_0 (x, x) = \bar{\lambda}_0 (x, x)$, откуда следует, что $\lambda_0 = \bar{\lambda}_0$. Теорема доказана.

Теорема 2. *Собственные векторы эрмитова оператора, принадлежащие различным собственным значениям, ортогональны между собой.*

Действительно, из равенств $(Ax, y) = \lambda(x, y)$, $(x, Ay) = \bar{\mu}(x, y) = \mu(x, y)$ вытекает, что $(\lambda - \mu)(x, y) = 0$. При $\lambda \neq \mu$ следует утверждение теоремы.

Поскольку метод ортогонализации системы векторов и их нормировки в унитарном пространстве полностью совпадает с аналогичным методом в E_n, то имеет место

Теорема 3. *В унитарном пространстве для каждого эрмитова оператора существует канонический ортонормированный базис, построенный из единичных взаимно-ортогональных собственных векторов оператора, относительно которого матрица оператора имеет диагональный вид*

$$\begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & 0 & \\ & & & \ddots & & \\ & & & & \lambda_s & \\ & 0 & & & & \ddots \\ & & & & & & \lambda_s \end{bmatrix}, \tag{26.19}$$

где $\lambda_1, \ldots, \lambda_s$ — *собственные значения эрмитова оператора, взятые столько раз, какова их кратность в характеристическом уравнении.*

Учитывая, что матрица A эрмитова оператора связана с матрицей симметричной эрмитовой формы (симметричной эрмитово-квадратичной формы) соотношением $A = H^{tr}$, то из закона (26.5) преобразования матрицы эрмитовой формы следует $H'^{tr} = T^* H^{tr} T$, а переход от одного ортонормированного базиса к другому совершается с помощью унитарной матрицы $T = U$, для которой $U^* = U^{-1}$, мы можем на основании (26.19) сформулировать следующий результат.

Теорема 4. *Для всякой эрмитовой матрицы A существует такая унитарная матрица U, что матрица $U^{-1}AU$ имеет вещественный диагональный вид.*

На этом мы закончим изучение унитарных пространств и свойств эрмитовых и унитарных операторов, многие из которых переносятся на бесконечномерные линейные пространства и находят свое приложение в квантовой механике.

ЛИТЕРАТУРА

1. Александров П. С. Курс аналитической геометрии и линейной алгебры.— М.: Наука, 1979.

2. Беклемишев Д. В. Курс аналитической геометрии и линейной алгебры.— М.: Наука, 1976.

3. Ефимов Н. В. Краткий курс аналитической геометрии.— М.: Наука, 1975.

4. Ильин В. А., Позняк Э. Г. Аналитическая геометрия.— М.: Наука, 1971.

5. Ильин В. А., Позняк Э. Г. Линейная алгебра.—М.: Наука, 1978.

6. Клетеник Д. В. Сборник задач по аналитической геометрии.— М.: Наука, 1975.

7. Мальцев А. И. Основы линейной алгебры.—М.: Наука, 1975.

8. Проскуряков И. В. Сборник задач по линейной алгебре.— М.: Наука, 1978.

9. Цубербиллер О. Н. Задачи и упражнения по аналитической геометрии.— М.: Наука, 1970.

10. Шилов Г. Е. Конечномерные линейные пространства.— М.: Наука, 1969.

СОДЕРЖАНИЕ

Владимир Романович Кайгородов

КУРС АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ И ЛИНЕЙНОЙ АЛГЕБРЫ

Редактор Ф. М. Абубакирова
Техн. редактор Е. Ф. Гаврилова
Обложка худож. Э. Я. Зарипова
Корректор Л. С. Губанова