МАТЕМАТИЧЕСКИЙ АНАЛИЗ (КОНЕЧНОМЕРНЫЕ ЛИНЕЙНЫЕ НРОСТРАНСТВА).

Г. Е. Шилов.

Книга представляет собой существенно переработанный вариант книги того же автора «Введение в теорию линейных пространств» (Гостехиздат, 1952 и 1956). Издание соответствует в основном программе университетского курса линейной алгебры и рассчитано в первую очередь на студентов математических, физических и других естественнонаучных специальностей. Для ее чтения необходимо, как правило, владение лишь элементарной математикой; в отдельных случаях используются сведения из математического анализа с соответствующими отсылками. В главе 1 излагается теория определителей. В главах 2—7 рассматривается аффинная теория линейных пространств (над произвольным числовым полем), в главах 8—10—теория евклидовых и унитарных пространств. В главе 11 описываются алгебры линейных операторов в конечномерных пространствах и в главе 12—соответствующие категории.

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Эта книга предназначена в качестве учебного пособия для студентов младших курсов математических и физических специальностей. В ней излагается материал, обычно входящий в курс линейной алгебры и в дальнейшем обслуживающий различные разделы математического анализа. Следует заметить, впрочем, что название «линейная алгебра» давно уже не соответствует реальному содержанию курса, который представляет собой синтез идей алгебры, геометрии и анализа. И хотя анализ в точном смысле слова (т. е. отдел математики, связанный с пределами, дифференцированием, интегрированием) присутствует в книге явно лишь на втором плане, на самом деле он-то и является настоящим организатором курса, поскольку проблемы «линейной алгебры» можно считать «конечномерными проекциями» основных проблем анализа и в то же время «опорой» для них.

Книга написана на основе нашей старой книги «Введение в теорию линейных пространств» (Гостехиздат, 1952 и 1956; далее ее именуем ВЛП). Различие между ВЛП и новой книгой, в кратких словах, следующее. В ВЛП речь шла исключительно о вещественных пространствах; в новой книге рассматриваются пространства над произвольным числовым полем, вещественный и комплексный случаи излагаются как специфические случаи общей теории, находящиеся в тесной связи друг с другом.

Введена глава о жордановой форме матрицы линейного оператора в комплексном и вещественном пространстве. Для комплексного пространства со скалярным произведением рассмотрены канонические формы матриц нормальных операторов, из которых, как частные случаи, получаются канонические формы матриц эрмитовых, антиэрмитовых и унитарных операторов и их вещественных аналогов Книга ВЛП заканчивалась большой главой о геометрии бесконечно-мерного гильбертова пространства; здесь такой главы нет (В ряде других книг можно найти более систематическое изложение этого материала, относящегося скорее уже к функциональному анализу). Зато добавлены две новые главы, непосредственно примыкающие к основному содержанию курса: глава о структуре матричных алгебр (написанная по просьбе автора А. Я. Хелемским) и глава о строении матричных категорий (содержание которой взято из статьи И. М. Гельфанда и автора, Вестник МГУ, Математика и механика, 1963, № 4). Эти две главы, хотя и вполне элементарны по методам, все же несколько выше по уровню, чем остальные; они представляют линейную алгебру в ее развитии и могут быть использованы в факультативных занятнях.

Каждая глава книги заканчивается рядом задач. В некоторой, но весьма небольшой мере они помогают выработке необходимых технических навыков (для этой цели лучше использовать ряд распространенных задачников с богатым выбором упражнений). В основном имеющиеся в книге задачи предназначены для иллюстрации и некоторого развития основного текста; циклы из некоторых задач могут служить темами для докладов на семинарах. Для этой же цели предназначены отдельные необязательные параграфы основного текста, выделенные звездочкой.

Автор считает своим приятным долгом отметить тщательную работу М. С. Аграновича и принести ему благодарность за ряд весьма ценных замечаний. Автор благодарит также И. Я. Дорфман за проверку всех задач.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОПРЕДЕЛИТЕЛИ

§ 1.1. Числовые поля

1.11. Как и бóльшая часть математики, линейная алгебра использует числовые системы (числовые поля). *Числовым полем* называют всякую совокупность K объектов, называемых *числами*, в которой можно производить с этими объектами четыре арифметических действия. Перечислим необходимые сведения об этих действиях (аксиомы поля).

а. Каждой паре чисел α и β отвечает число $\alpha+\beta$, называемое *суммой* чисел α и β, причем

1) $\alpha+\beta=\beta+\alpha$ для любых α и β из K (переместительность, или коммутативность сложения);

2) $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ для любых α, β, γ из K (сочетательность, или ассоциативность сложения);

3) существует число 0 (*нуль*) такое, что $0+\alpha=\alpha$ для любого α из K;

4) для любого α из K существует число β из K такое, что $\alpha+\beta=0$ (*противоположный элемент*).

Разрешимость уравнения $\alpha+\beta=0$ при любом α позволяет ввести операцию вычитания: *разность* $\alpha-\beta$ по определению есть сумма числа α и решения γ уравнения $\beta+\gamma=0$.

б. Каждой паре чисел α и β отвечает число $\alpha\cdot\beta$ (или $\alpha\beta$), называемое *произведением* чисел α и β, причем

5) $\alpha\beta=\beta\alpha$ для любых α и β из K (коммутативность умножения);

6) $(\alpha\beta)\gamma=\alpha(\beta\gamma)$ для любых α, β, γ из K (ассоциативность умножения);

7) существует число $1\ (\neq 0)$ такое, что $1\cdot\alpha=\alpha$ для любого числа α из K;

8) для любого $\alpha\neq 0$ существует число γ из K такое, что $\alpha\gamma=1$ (*обратный элемент*).

Разрешимость уравнения $\alpha\gamma = 1$ при любом $\alpha \neq 0$ позволяет ввести операцию деления на число $\alpha \neq 0$: *частное* β/α есть произведение числа β и решения γ уравнения $\alpha\gamma = 1$.

Числа $1+1=2$, $2+1=3$ и т. д. называются *натуральными*; предполагается, что ни одно из них не равно 0 *).

Натуральные числа, им противоположные и 0 образуют, по определению, *совокупность целых чисел* поля K. Частные p/q, где p и q — целые и $q \neq 0$, образуют *совокупность рациональных чисел* поля K.

Два поля K и K' называются *изоморфными*, если между ними можно так установить взаимно однозначное соответствие, что сумма и произведение чисел поля K соответствуют сумме и произведению соответствующих чисел поля K'. (В этом случае результаты остальных операций — разности и частного — также будут соответствовать друг другу.)

1.12. Наиболее часто встречаются следующие примеры конкретных числовых полей.

а. *Поле рациональных чисел*, т. е. отношений p/q, где p и $q \neq 0$ — обычные целые числа с обычными арифметическими правилами действий. (Заметим, что одни только *целые числа не образуют числового поля*, так как в этом случае не выполнена аксиома 8).)

Из сказанного выше следует, что в каждом поле K имеется часть (подполе), изоморфная полю рациональных чисел.

б. *Поле вещественных чисел*, имеющее геометрическим образом совокупность всех точек прямой. Аксиоматика поля вещественных чисел получается добавлением к аксиомам 1)—8) аксиом порядка и аксиомы о точной верхней грани **).

в. *Поле комплексных чисел* $a+ib$, где a и b вещественны (символ i не есть вещественное число), с правилами

---

*) Из двух элементов, положим N и E, можно устроить поле по правилам $N+N=N$, $N+E=E$, $E+E=N$, $N\cdot N=N$, $N\cdot E=N$, $E\cdot E=E$. В соответствии с нашими обозначениями мы должны положить $N=0$, $E=1$, и тогда $2=1+1=0$. Для исключения таких числовых систем мы и требуем, чтобы все натуральные элементы поля были бы отличными от 0.

**) Теория вещественных чисел подробно излагается в нашей книге «Математический анализ (Функции одного переменного)», «Наука», 1969, гл. I. В дальнейшем обозначаем эту книгу через ФОП.

действий (ФОП, *2.71*):

$$(a_1+ib_1)+(a_2+ib_2)=(a_1+a_2)+i\,(b_1+b_2),$$
$$(a_1+ib_1)\cdot(a_2+ib_2)=(a_1a_2-b_1b_2)+i(a_1b_2+a_2b_1).$$

Для чисел вида $a+i0$ эти действия приводятся к одноименным действиям над вещественными числами a; мы пишем коротко $a+i0=a$ и называем эти комплексные числа вещественными. Можно сказать, что поле комплексных чисел содержит часть (подполе), изоморфную полю вещественных чисел. Комплексные числа вида $0+ib$ называются (чисто) мнимыми и короче обозначаются ib.

Из правила умножения следует, что

$$i^2=i\cdot i=(0+i1)\,(0+i1)=-1.$$

1.13. Поле вещественных чисел в дальнейшем обозначается через R. Поле комплексных чисел обозначается через C. Согласно так называемой основной теореме алгебры (ФОП, *4.86*), в поле C не только выполнимы четыре арифметические операции, но и разрешимо любое алгебраическое уравнение

$$z^n+a_1z^{n-1}+\ldots+a_n=0.$$

Поле R вещественных чисел не обладает этим свойством: например, уравнение $x^2+1=0$ не имеет решений в поле R.

Многие из дальнейших построений справедливы для любого числового поля. *Любое* числовое поле мы будем обозначать в дальнейшем буквой K. Если некоторое предложение верно для поля K, оно автоматически верно для поля R и поля C, которые являются частными случаями поля K.

§ 1.2. Основные задачи теории систем линейных уравнений

1.21. В этой и двух следующих главах мы будем заниматься изучением систем линейных уравнений.

В самом общем случае такая система имеет следующий вид:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n=b_k. \end{array}\right\} \qquad (1)$$

Здесь через $x_1, x_2, \ldots, x_n$ обозначены неизвестные (элементы поля K), подлежащие определению (заметим, что число неизвестных не предполагается обязательно равным числу уравнений). Числа $a_{11}, a_{12}, \ldots, a_{kn}$, взятые из поля K, называются *коэффициентами* системы. Первый индекс коэффициента указывает номер уравнения, в котором фигурирует данный коэффициент, а второй индекс — номер неизвестного, при котором этот коэффициент поставлен *). Числа $b_1, b_2, \ldots, b_k$, стоящие в правых частях равенства (1), взятые из того же поля K, называются *свободными членами* системы; как и коэффициенты, они предполагаются известными.

Решением системы называется всякая совокупность чисел $c_1, c_2, \ldots, c_n$ из того же поля K, которая, будучи подставлена в систему (1) на место неизвестных $x_1, x_2, \ldots, x_n$, обращает все уравнения системы в тождества **).

Не всякая система линейных уравнений вида (1) имеет решение. Например, система

$$\left.\begin{aligned} 2x_1+3x_2&=5,\\ 2x_1+3x_2&=6 \end{aligned}\right\} \qquad (2)$$

заведомо не может иметь ни одного решения. Действительно, какие бы числа c_1, c_2 мы ни подставили на место неизвестных x_1, x_2, левые части уравнений системы (2) окажутся совпадающими, в то время как правые части различны. Поэтому оба уравнения системы (2) такой подстановкой не могут быть одновременно обращены в тождества.

Систему уравнений вида (1), имеющую (хотя бы одно) решение, мы будем называть *совместной*; систему, не имеющую решений, будем называть *несовместной*.

Совместная система может иметь одно решение или более чем одно; в последнем случае для различения решений мы будем указывать их номера индексами наверху в скобках; например, первое решение $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}$, второе решение $c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$ и т. д. Решение $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}$ и $c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$ считаются *различными*, если хотя бы одно из чисел $c_i^{(1)}$ не совпадает с соответствующим числом

---

*) Поэтому, например, запись a_{34} должна читаться так: «a — три — четыре» (а не «a — тридцать четыре»).

**) Подчеркнем, что совокупность чисел $c_1, c_2, \ldots, c_n$ составляет одно решение системы (а не n решений).

$c_i^{(2)}$ $(i=1, 2, \ldots, n)$. Например, система

$$\left.\begin{aligned} 2x_1+3x_2=0, \\ 4x_1+6x_2=0 \end{aligned}\right\} \qquad (3)$$

имеет различные решения $c_1^{(1)}=c_2^{(1)}=0$, $c_1^{(2)}=3$, $c_2^{(2)}=-2$ (а также бесконечное множество других решений). Если совместная система имеет единственное решение, она называется *определенной;* если совместная система имеет по крайней мере два различных решения, она называется *неопределенной.*

1.22. Мы можем сформулировать теперь те основные задачи, которые возникают при изучении системы (1):

I. *Выяснить, является система* (1) *совместной или несовместной.*

II. *Если система* (1) *совместна, то выяснить, является ли она определенной.*

III. *Если система* (1) *совместна и определенна, то найти ее единственное решение.*

IV. *Если система* (1) *совместна и неопределенна, то описать совокупность всех ее решений.*

Основным математическим инструментом для изучения линейных систем является теория определителей; мы переходим теперь к ее изложению.

§ 1.3. Определитель n-го порядка

1.31. Пусть дана *квадратная матрица*, т. е. таблица из n^2 чисел (элементов поля K) a_{ij} $(i, j=1, 2, \ldots, n)$:

$$A=\left\| \begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{matrix} \right\|. \qquad (4)$$

Число n, указывающее количество строк и столбцов матрицы (4), называется ее *порядком*. Числа a_{ij} называются *элементами* матрицы A; первый и второй индексы у элемента a_{ij} указывают соответственно номер строки и столбца, в которых расположен этот элемент. Элементы $a_{11}, a_{22}, \ldots, a_{nn}$ образуют *главную диагональ* матрицы A.

Рассмотрим любое произведение n элементов, расположенных в различных строках и различных столбцах матрицы (4), т. е. по одному в каждой строке и в каждом столбце. Такое произведение можно записать в виде

$$a_{\alpha_1 1}\ a_{\alpha_2 2}\ \ldots\ a_{\alpha_n n}. \tag{5}$$

Действительно, в качестве первого сомножителя мы всегда можем взять элемент, стоящий в первом столбце матрицы (4); если обозначить через α_1 номер строки, в которой находится этот элемент, то индексы этого элемента будут α_1 и 1. Аналогично в качестве второго сомножителя можно взять элемент, стоящий во втором столбце; его индексы будут α_2 и 2, где α_2 — номер той строки, в которой расположен этот второй элемент; и т. д. Таким образом, индексы $\alpha_1, \alpha_2, \ldots, \alpha_n$ являются номерами строк, в которых расположены сомножители произведения (5), в соответствии с принятым порядком их записи по возрастанию индексов столбцов.

Так как по условию элементы $a_{\alpha_1 1}, a_{\alpha_2 2}, \ldots, a_{\alpha_n n}$ расположены в различных строках матрицы (4), по одному в каждой строке, то числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ все различны и представляют собой некоторую перестановку чисел $1, 2, \ldots, n$.

Назовем «беспорядком» в этой последовательности $\alpha_1, \alpha_2, \ldots, \alpha_n$ такое расположение индексов, когда старший индекс стоит раньше младшего. Число всех «беспорядков» обозначим через $N(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Например, в перестановке четырех цифр 2, 1, 4, 3 — два «беспорядка» (2 впереди 1, 4 впереди 3); таким образом,

$$N(2, 1, 4, 3) = 2.$$

В перестановке 4, 3, 1, 2 — пять «беспорядков» (4 впереди 3, 4 впереди 1, 4 впереди 2, 3 впереди 1, 3 впереди 2); поэтому

$$N(4, 3, 1, 2) = 5.$$

Если число беспорядков в последовательности $\alpha_1, \ldots, \alpha_n$ четно, поставим перед произведением (5) знак $+$; если это число нечетно, поставим перед этим произведением знак $-$. Иными словами, условимся перед каждым произведением вида (5) писать знак, определяемый выражением

$$(-1)^{N(\alpha_1, \alpha_2, \ldots, \alpha_n)}.$$

Число всех произведений вида (5), которые можно составить из элементов данной матрицы n-го порядка, равно числу всех возможных перестановок чисел $1, 2, \ldots, n$, которое, как известно, равно $n!$.

Теперь введем следующее определение:

Определителем матрицы (4) *называется алгебраическая сумма, состоящая из* $n!$ *всевозможных произведений вида* (5), *перед каждым из которых поставлен знак, определенный по указанному выше правилу:*

$$D=\sum(-1)^{N(\alpha_1, \alpha_2, \ldots, \alpha_n)} a_{\alpha_1 1}a_{\alpha_2 2}\ldots a_{\alpha_n n}. \tag{6}$$

В дальнейшем произведения вида (5) мы будем называть *членами* определителя. Элементы a_{ij} матрицы (4) будем называть *элементами* определителя.

Определитель матрицы (4) обозначается одним из следующих символов:

$$D=\begin{vmatrix} a_{11} & a_{12} \ldots & a_{1n} \\ a_{21} & a_{22} \ldots & a_{2n} \\ \cdot & \cdot \quad \cdot & \cdot \\ a_{n1} & a_{n2} \ldots & a_{nn} \end{vmatrix}=\det\|a_{ij}\|=\det\|a_{ij}\|_{i,j=1,2,\ldots,n}. \tag{7}$$

Например, для определителей 2-го и 3-го порядка мы получаем следующие выражения:

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

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

Роль определителей при решении систем линейных уравнений мы покажем на примере системы из двух уравнений с двумя неизвестными. Если дана система

$$a_{11}x_1+a_{12}x_2=b_1,$$
$$a_{21}x_1+a_{22}x_2=b_2,$$

то, исключая обычным образом одно из неизвестных, можно легко получить формулы

$$x_1=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{21}a_{12}}, \qquad x_2=\frac{a_{11}b_2-a_{21}b_1}{a_{11}a_{22}-a_{21}a_{12}},$$

в предположении, что знаменатели этих отношений отличны от нуля. Числители и знаменатели получающихся дробей представляют собой определители 2-го порядка:

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

$$b_1a_{22}-b_2a_{12}=\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix},$$

$$a_{11}b_2-a_{21}b_1=\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

Оказывается, что аналогичные формулы имеют место и для решения систем с любым числом неизвестных (см. § 1.7).

1.32. Правило для определения знака данного члена определителя можно сформулировать несколько иначе, в геометрических терминах.

В матрице (4) в соответствии с нумерацией элементов естественно выделяются положительные направления: слева направо — вдоль строк, сверху вниз — вдоль столбцов. Вместе с этим и косые отрезки, соединяющие два каких-либо элемента матрицы, можно снабдить указанием направления: будем говорить, что отрезок, соединяющий элемент a_{ij} с элементом a_{km}, имеет *положительный наклон*, если его правый конец расположен ниже левого, и *отрицательный наклон*, если его правый конец лежит выше, чем левый. Теперь проведем мысленно в матрице (4) все отрезки, соединяющие попарно элементы $a_{\alpha_1 1}, a_{\alpha_2 2}, \ldots, a_{\alpha_n n}$ произведения (5) и при этом имеющие о т р и ц а т е л ь н ы й наклон. Будем ставить перед произведением (5) знак $+$, если число всех таких отрезков четно, и знак $-$, если их число нечетно.

Например, в случае матрицы 4-го порядка перед произведением $a_{21}a_{12}a_{43}a_{34}$ должен быть поставлен знак $+$, так как в матрице имеется два отрезка отрицательного наклона, соединяющих элементы данного произведения:

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix},$$

а перед произведением $a_{41}a_{32}a_{13}a_{24}$ должен быть поставлен знак $-$, так как в матрице имеется пять отрезков с отрицательным наклоном, соединяющих его элементы:

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix}.$$

В этих примерах количество отрезков отрицательного наклона, соединяющих элементы данного члена, равно числу «беспорядков» в расположении первых индексов элементов, составляющих в произведении данный член: в первом примере последовательность первых индексов 2, 1, 4, 3 имеет два «беспорядка», во втором примере последовательность первых индексов 4, 3, 1, 2 имеет пять «беспорядков».

Покажем, что *второе определение знака члена определителя равносильно первому*. Для этого достаточно показать, что число «беспорядков» в последовательности первых индексов элементов данного члена (при натуральном порядке вторых индексов) всегда равно числу отрезков отрицательного наклона, соединяющих элементы данного члена в матрице. Но это почти очевидно: наличие отрезка отрицательного наклона, соединяющего элементы $a_{\alpha_i i}$ и $a_{\alpha_j j}$, означает при $i < j$, что $\alpha_i > \alpha_j$, т. е. наличие «беспорядка» в расположении первых индексов.

См. задачи 1—3 (в конце этой главы).

§ 1.4. Свойства определителей

1.41. Операция транспонирования. Определитель

$$\begin{vmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix}, \tag{8}$$

полученный из определителя (7) заменой строк на столбцы с теми же номерами, называется *транспонированным* по отношению к определителю (7). Покажем, что *величина транспонированного определителя совпадает с величиной исходного определителя*. Действительно, определители (7) и (8) состоят, очевидно, из одних и тех же членов; поэтому нам достаточно показать, что одинаковые члены обладают в определителях (7) и (8) и одинаковыми знаками. Транспонирование матрицы определителя, очевидно, есть результат ее поворота (в пространстве) на $180°$ вокруг диагонали $a_{11}, a_{22}, \ldots, a_{nn}$. При этом повороте каждый отрезок с отрицательным наклоном (например, образующий угол $\alpha < 90°$ со строками матрицы) переходит снова в отрезок с отрицательным наклоном (именно, образующий со строками матрицы угол $90° - \alpha$). Поэтому число отрезков с отрицательным наклоном, соединяющих элементы данного члена, после транспонирования не изменится; следовательно, не изменится и знак этого члена. Таким образом, знаки всех членов сохранятся; тем самым величина определителя остается неизменной.

Доказанное сейчас свойство определителя устанавливает равноправие его строк и столбцов. Поэтому дальнейшие свойства определителей мы будем формулировать и доказывать только для столбцов.

1.42. Свойство антисимметрии. Под *антисимметрией относительно столбцов* понимают свойство определителя *менять знак при перестановке двух столбцов*. Рассмотрим сначала случай, когда переставляются два соседних столбца определителя, например j-й и $(j+1)$-й. Определитель, полученный после перестановки столбцов, будет состоять, очевидно, из тех же самых членов, что и исходный определитель. Рассмотрим какой-нибудь из членов исходного определителя. Этот член в своем составе имеет элемент из j-го столбца и элемент из $(j+1)$-го столбца. Если отрезок, соединяющий эти два элемента, имел отрицательный наклон, то после перестановки столбцов его наклон станет положительным, и наоборот. Что же касается остальных отрезков, соединяющих попарно элементы выделенного члена, то после перестановки столбцов характер наклона каждого из них останется неизменным. Следовательно, количество отрезков с отрицательным наклоном, соединяющих элементы данного члена, при перестановке столбцов заведомо изменяется на единицу; поэтому каждый член определителя, а следовательно, и сам определитель, при перестановке столбцов меняет знак.

Пусть теперь переставляются не соседние столбцы, а, например, j-й столбец с k-м столбцом, причем между ними находится m других столбцов и $j<k$. Эту перестановку можно осуществить последовательными перестановками соседних столбцов в следующем порядке: сначала j-й столбец переставляется с $(j+1)$-м, далее с $(j+2)$-м, $(j+3)$-м,... ..., k-м, затем получившийся $(k-1)$-й столбец (ранее бывший k-м) переставляется с $(k-2)$-м, $(k-3)$-м, ..., j-м. Всего понадобится $m+1+m=2m+1$ перестановок соседних столбцов; после каждой из них, по доказанному, определитель изменяет знак и, следовательно, после конца процесса будет иметь знак, противоположный начальному (поскольку $2m+1$ при любом целом m есть нечетное число).

1.43. Следствие. *Определитель, имеющий два одинаковых столбца, равен нулю.*

В самом деле, переставляя эти столбцы, мы не изменим определителя; с другой стороны, по доказанному, он должен изменить свой знак. Таким образом, $D=-D$, откуда следует, что $D=0$.

См. задачу 4.

1.44. Линейное свойство определителя. Это свойство формулируется следующим образом:

а. *Если все элементы j-го столбца определителя D представлены в виде «линейной комбинации» двух слагаемых*

$$a_{ij}=\lambda b_i+\mu c_i \qquad (i=1,\ 2,\ \ldots,\ n)$$

(λ и μ — фиксированные числа), то определитель D равен такой же линейной комбинации двух определителей:

$$D=\lambda D_1+\mu D_2, \tag{9}$$

причем у каждого из этих двух определителей все столбцы, кроме j-го, такие же, как у определителя D, а j-й столбец состоит у определителя D_1 из чисел b_i, у определителя D_2 — из чисел c_i.

Действительно, всякий член определителя D можно представить в виде

$$a_{\alpha_1 1}a_{\alpha_2 2}\ldots a_{\alpha_j j}\ldots a_{\alpha_n n}=a_{\alpha_1 1}a_{\alpha_2 2}\ldots(\lambda b_{\alpha_j}+\mu c_{\alpha_j})\ldots a_{\alpha_n n}=$$
$$=\lambda a_{\alpha_1 1}a_{\alpha_2 2}\ldots b_{\alpha_j}\ldots a_{\alpha_n n}+\mu a_{\alpha_1 1}a_{\alpha_2 2}\ldots c_{\alpha_j}\ldots a_{\alpha_n n}.$$

Собирая вместе первые слагаемые (с теми знаками, которые имели соответствующие члены первого определителя) и вынося за скобки число λ, получим в скобках, очевидно, определитель D_1; аналогично, собирая вторые слагаемые и вынося за скобки число μ, получим определитель D_2. Таким образом, формула (9) установлена.

Эту формулу удобнее записать в несколько ином виде. Пусть D — произвольный фиксированный определитель. Обозначим через $D_j(p_i)$ определитель, который получается при замене элементов j-го столбца определителя D на числа p_i ($i=1, 2, \ldots, n$). Исходный определитель можно записать в форме $D(a_{ij})$. Тогда доказанное нами равенство (9) принимает вид $D_j(\lambda b_i + \mu c_i) = \lambda D_j(b_i) + \mu D_j(c_i)$.

б. Линейное свойство определителя без труда распространяется на тот случай, когда каждый элемент j-го столбца есть линейная комбинация любого фиксированного числа слагаемых:

$$a_{ij} = \lambda b_i + \mu c_i + \ldots + \tau f_i;$$

в этом случае

$$\begin{aligned} D_j(a_{ij}) &= D_j(\lambda b_i + \mu c_i + \ldots + \tau f_i) = \\ &= \lambda D_j(b_i) + \mu D_j(c_i) + \ldots + \tau D_j(f_i). \end{aligned} \tag{10}$$

1.45. Следствие. *Общий множитель всех элементов некоторого столбца определителя можно вынести за знак определителя.*

В самом деле, если $a_{ij} = \lambda b_i$, то по формуле (10)

$$D_j(a_{ij}) = D_j(\lambda b_i) = \lambda D_j(b_i),$$

что и утверждается.

1.46. Следствие. *Если некоторый столбец определителя состоит целиком из нулей, то определитель равен нулю.*

В самом деле, 0 есть общий множитель элементов данного столбца; вынося его за знак определителя, получим

$$D_j(0) = D_j(0 \cdot 1) = 0 \cdot D_j(1) = 0.$$

См. задачу 5.

1.47. Прибавление к одному столбцу другого столбца с произвольным множителем.

а. *Определитель не изменится, если к элементам одного из его столбцов прибавить соответствующие элементы любого другого столбца, умноженные на фиксированное число.*

Пусть к j-му столбцу прибавляется k-й ($k \neq j$), умноженный на число λ. В полученном определителе j-й столбец будет состоять из элементов вида $a_{ij}+\lambda a_{ik}$ ($i=1, 2, \ldots, n$). В силу формулы (9)

$$D_j(a_{ij}+\lambda a_{ik})=D_j(a_{ij})+\lambda D_j(a_{ik}).$$

Во втором определителе j-й столбец состоит из элементов a_{ik}, т. е. совпадает с k-м столбцом. По следствию *1.43* $D_j(a_{ik})=0$, откуда

$$D_j(a_{ij}+\lambda a_{ik})=D_j(a_{ij}),$$

что и требуется.

б. Разумеется, свойство *а* можно сформулировать в более общей форме: *определитель D не изменится, если к элементам его j-го столбца прибавить соответствующие элементы k-го столбца, умноженные на число λ, затем элементы l-го столбца, умноженные на число μ, ..., элементы p-го столбца, умноженные на число τ ($k \neq j$, $l \neq j$, ..., $p \neq j$).*

См. задачу 6.

1.48. Все свойства, доказанные нами в этом параграфе для столбцов определителя, в силу неизменности определителя при транспонировании *(1.41)* остаются справедливыми и для его строк.

§ 1.5. Алгебраические дополнения и миноры

1.51. Рассмотрим произвольный, например j-й, столбец определителя D. Пусть a_{ij} — некоторый элемент этого столбца. В правой части равенства (6)

$$D=\sum(-1)^{N(\alpha_1, \alpha_2, \ldots, \alpha_n)}a_{\alpha_1 1}a_{\alpha_2 2}\ldots a_{\alpha_n n},$$

задающего определитель D, соберем все члены, содержащие элемент a_{ij}, заключим их в скобки и вынесем за эти скобки элемент a_{ij}. Величина, оставшаяся в скобках, обозначается через A_{ij}; она называется *алгебраическим дополнением элемента a_{ij} в определителе D.*

Так как в каждый член определителя D входит элемент из j-го столбца, то равенству (6) можно придать теперь вид

$$D=a_{1j}A_{1j}+a_{2j}A_{2j}+\ldots+a_{nj}A_{nj}. \tag{11}$$

Формула (11) называется *формулой разложения определителя D по элементам j-го столбца*. Разумеется, аналогичную формулу можно написать и для любой строки определителя D; например, для i-й строки мы получим такое равенство:

$$D = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}. \tag{12}$$

Мы получили теорему.

Теорема. *Сумма всех произведений элементов какого-нибудь столбца (или строки) определителя D на соответствующие алгебраические дополнения равна самому определителю D.*

Формулы (11) и (12) можно использовать для вычисления определителя. Но при этом необходимо уметь вычислять алгебраические дополнения; правила для их вычисления мы приведем в *1.54*.

1.52. Отметим одно следствие формул (11) и (12), которое будет в дальнейшем использовано.

Равенство (11) выполняется тождественно относительно величин $a_{1j}, a_{2j}, \ldots, a_{nj}$; поэтому оно останется справедливым, если заменить в нем a_{ij} $(i = 1, 2, \ldots, n)$ на любые другие величины. При такой замене величины $A_{1j}, A_{2j}, \ldots, A_{nj}$ остаются неизменными, поскольку они не зависят от элементов a_{ij}. Заменим в правой и левой частях равенства (11) элементы $a_{1j}, a_{2j}, \ldots, a_{nj}$ на соответствующие элементы какого-нибудь другого, например k-го, столбца. Тогда определитель слева в (11) будет иметь два одинаковых столбца и по *1.43* будет равен нулю. Мы получаем равенство (при $k \neq j$)

$$a_{1k}A_{1j} + a_{2k}A_{2j} + \ldots + a_{nk}A_{nj} = 0. \tag{13}$$

Аналогично из формулы (12) при $l \neq i$ получаем

$$a_{l1}A_{i1} + a_{l2}A_{i2} + \ldots + a_{ln}A_{in} = 0. \tag{14}$$

Итак, мы доказали теорему:

Теорема. *Сумма всех произведений элементов какого-нибудь столбца (или какой-нибудь строки) определителя D на алгебраические дополнения соответствующих элементов другого столбца (строки) равна нулю.*

1.53. Если зачеркнуть в матрице n-го порядка некоторую строку и некоторый столбец, то оставшиеся элементы, естественно, образуют некоторую матрицу $(n-1)$-го порядка. Определитель этой матрицы называется *минором* данной матрицы n-го порядка (а также минором ее определителя D). Если были зачеркнуты i-я строка и j-й столбец, то полученный минор обозначается через M_{ij} или $M_{ij}(D)$.

Мы докажем, что имеет место равенство

$$A_{ij}=(-1)^{i+j}M_{ij}, \tag{15}$$

с помощью которого вычисление алгебраических дополнений сводится к вычислению соответствующих миноров.

Доказательство равенства (15) проведем сначала для случая $i=1$, $j=1$. Соберем в правой части равенства (6) все члены, содержащие элемент a_{11}. Рассмотрим один из таких членов. Очевидно, что произведение всех его элементов, за исключением a_{11}, дает некоторый член c минора M_{11}. Так как в матрице определителя D нет отрезков с отрицательным наклоном, соединяющих элемент a_{11} с остальными элементами выделенного члена, то знак, который приписывается члену $a_{11}c$ определителя D, совпадает со знаком, который приписывается члену c в миноре M_{11}. Выбирая должным образом член определителя D, содержащий элемент a_{11}, и зачеркивая a_{11}, можно получить любой член минора M_{11}. Поэтому рассматриваемая алгебраическая сумма всех членов определителя D, содержащих a_{11}, равна произведению $a_{11}M_{11}$. Но согласно *1.51* эта сумма равна произведению $a_{11}A_{11}$. Следовательно, $A_{11}=M_{11}$, что и требуется.

Теперь мы докажем формулу (15) при любых i и j. То обстоятельство, что при $i=j=1$ эта формула справедлива, будет нами существенно использовано. Рассмотрим элемент $a_{ij}=a$, расположенный на пересечении i-й строки и j-го столбца определителя D. Переставляя последовательно соседние строки и столбцы, мы можем перевести элемент a в левый верхний угол матрицы; для этого понадобится $i-1+j-1=i+j-2$ перестановок. В результате мы получим определитель D_1 с теми же членами, какие будет иметь исходный определитель D, если его умножить на $(-1)^{i+j-2}=(-1)^{i+j}$. Минор $M_{11}(D_1)$ определителя D_1, очевидно, совпадает с минором $M_{ij}(D)$ определителя D. По доказанному, в определителе D_1 члены, содержащие элемент a, составляют в сумме величину $aM_{11}(D_1)$. Поэтому в составе исходного определителя D члены, содержащие элемент $a_{ij}=a$, образуют в сумме величину

$$(-1)^{i+j} a M_{11}(D_1) = a_{ij}(-1)^{i+j} M_{ij}(D).$$

Но согласно *1.51* эта же сумма равна произведению $a_{ij}A_{ij}$. Следовательно, $A_{ij} = (-1)^{i+j} M_{ij}$; тем самым формула (15) доказана полностью.

1.54. Формулы (11) и (12) можно теперь записать соответственно в форме

$$D = (-1)^{1+j} a_{1j} M_{1j} + (-1)^{2+j} a_{2j} M_{2j} + \ldots + (-1)^{n+j} a_{nj} M_{nj},$$
$$D = (-1)^{i+1} a_{i1} M_{i1} + (-1)^{i+2} a_{i2} M_{i2} + \ldots + (-1)^{i+n} a_{in} M_{in},$$

в которой они обычно и употребляются.

1.55. Примеры.

а. Определитель третьего порядка допускает шесть различных разложений (три по строкам и три по столбцам). Например, разложение по первой строке имеет вид

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

б. Определитель n-го порядка

$$D_n = \begin{vmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ a_{31} & a_{32} & a_{33} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{vmatrix}$$

называется *треугольным*. Разлагая его по первой строке, получим, что определитель D_n равен произведению элемента a_{11} на треугольный определитель $(n-1)$-го порядка

$$D_{n-1} = \begin{vmatrix} a_{22} & 0 & \ldots & 0 \\ a_{32} & a_{33} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ a_{n2} & a_{n3} & \ldots & a_{nn} \end{vmatrix}.$$

Определитель D_{n-1} снова разложим по первой строке; получим $D_{n-1} = a_{22} D_{n-2}$, где D_{n-2} — треугольный определитель $(n-2)$-го порядка. Продолжая таким образом далее, в конце концов получим

$$D_n = a_{11}\, a_{22} \ldots a_{nn},$$

т. е. треугольный определитель равен произведению элементов, стоящих на его главной диагонали.

§ 1.6. Практическое вычисление определителей

1.61. Формула (12) приобретает особенно простой вид, когда все элементы i-й строки, кроме одного, например a_{ik}, равны нулю. В этом случае

$$D = a_{ik}A_{ik}, \tag{16}$$

и вычисление определителя D n-го порядка непосредственно приводится к вычислению определителя $(n-1)$-го порядка. Но если при $a_{ik} \neq 0$ в i-й строке есть элемент a_{ij}, также не равный нулю, то мы можем вычесть из j-го столбца определителя D k-й столбец, умноженный на $\lambda = \frac{a_{ij}}{a_{ik}}$; в результате мы получим определитель, равный исходному (*1.47*), у которого j-й элемент i-й строки уже равен нулю. Повторяя аналогичные операции, мы можем от любого определителя с фиксированным элементом $a_{ik} \neq 0$ перейти к определителю, у которого все элементы i-й строки, кроме a_{ik}, равны нулю, и вычислить его по формуле (16). Разумеется, аналогичные преобразования можно производить и со столбцами определителя.

1.62. Пример. Вычислим определитель 5-го порядка

$$D = \begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & 0 & 3 & 7 & -2 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 6 & -4 & 1 & 2 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix}.$$

В третьем столбце этого определителя уже имеется два нуля. Чтобы получить в этом столбце еще два нуля, нужно ко второй строке прибавить утроенную пятую, а из четвертой строки вычесть учетверенную пятую. После этой операции и разложения определителя по третьему столбцу мы получаем

$$D = \begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & -9 & 0 & 13 & 7 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 18 & 0 & -7 & -10 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix} =$$

$$= (-1)^{3+5} \cdot (-1) \begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{vmatrix} = - \begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{vmatrix}.$$

Теперь проще всего получить три нуля в первом столбце: для этого мы прибавим к первой строке удвоенную вторую, а из третьей и четвертой строк вычтем вторую, соответственно утроенную и удвоенную:

$$D=-\begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{vmatrix}=-\begin{vmatrix} 0 & -13 & 25 & 17 \\ 1 & -9 & 13 & 7 \\ 0 & 26 & -34 & -26 \\ 0 & 36 & -33 & -24 \end{vmatrix}=$$

$$=-(-1)^{1+2}\begin{vmatrix} -13 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix}.$$

Чтобы легче было вычислить полученный определитель 3-го порядка, постараемся уменьшить абсолютные величины его элементов. Для этого после вынесения из второй строки общего множителя 2 прибавим вторую строку к первой и из третьей строки вычтем удвоенную вторую:

$$D=2\begin{vmatrix} -13 & 25 & 17 \\ 13 & -17 & -13 \\ 36 & -33 & -24 \end{vmatrix}=2\begin{vmatrix} 0 & 8 & 4 \\ 13 & -17 & -13 \\ 10 & 1 & 2 \end{vmatrix}=2\cdot 4\begin{vmatrix} 0 & 2 & 1 \\ 13 & -17 & -13 \\ 10 & 1 & 2 \end{vmatrix}.$$

В первой строке имеется уже один нуль. Чтобы получить еще один нуль, вычтем из второго столбца удвоенный третий; после этого определитель легко вычисляется до конца:

$$D=8\begin{vmatrix} 0 & 2 & 1 \\ 13 & -17 & -13 \\ 10 & 1 & 2 \end{vmatrix}=8\begin{vmatrix} 0 & 0 & 1 \\ 13 & 9 & -13 \\ 10 & -3 & 2 \end{vmatrix}=8(-1)^{1+3}\begin{vmatrix} 13 & 9 \\ 10 & -3 \end{vmatrix}=$$

$$=8\cdot 3\begin{vmatrix} 13 & 3 \\ 10 & -1 \end{vmatrix}=8\cdot 3(-13-30)=-8\cdot 3\cdot 43=-1032.$$

См. задачи 7—10.

§ 1.7. Правило Крамера

1.71. Мы можем перейти теперь к решению систем линейных уравнений. Рассмотрим сначала систему специального вида

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n&=b_n \end{aligned}\right\}\qquad(17)$$

с числом неизвестных, равным числу уравнений. Коэффициенты a_{ij} $(i,\ j=1,\ 2,\ \ldots,\ n)$ образуют *основную матрицу системы*, относительно которой мы предположим, что ее определитель D отличен от нуля. Мы покажем, что такая система всегда совместна и определенна, и получим формулу для вычисления ее единственного решения.

Допустим сначала, что система (17) имеет некоторое решение $c_1, c_2, \ldots, c_n$; справедлива, следовательно, система равенств

$$\left.\begin{aligned} a_{11}c_1 + a_{12}c_2 + \ldots + a_{1n}c_n &= b_1, \\ a_{21}c_1 + a_{22}c_2 + \ldots + a_{2n}c_n &= b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}c_1 + a_{n2}c_2 + \ldots + a_{nn}c_n &= b_n. \end{aligned}\right\} \qquad (18)$$

Умножим первое из равенств (18) на алгебраическое дополнение A_{11} к элементу a_{11} в матрице системы, далее умножим второе равенство на A_{21}, третье — на A_{31} и т. д., пока не дойдем до последнего равенства; затем все полученные равенства сложим. В результате мы получим следующее соотношение:

$$\begin{aligned} (a_{11}A_{11} + a_{21}A_{21} + \ldots + a_{n1}A_{n1})\, c_1 + \\ + (a_{12}A_{11} + a_{22}A_{21} + \ldots + a_{n2}A_{n1})\, c_2 + \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ + (a_{1n}A_{11} + a_{2n}A_{21} + \ldots + a_{nn}A_{n1})\, c_n = \\ = b_1A_{11} + b_2A_{21} + \ldots + b_nA_{n1}. \end{aligned} \qquad (19)$$

В силу теоремы *1.51* коэффициент при c_1 в соотношении (19) равен самому определителю D; в силу теоремы *1.52* коэффициенты при всех остальных c_j $(j \neq 1)$ обращаются в нуль. Выражение в правой части есть разложение определителя

$$D_1 = \begin{vmatrix} b_1 & a_{12} & \ldots & a_{1n} \\ b_2 & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_n & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

по первому его столбцу; поэтому равенство (19) можно теперь записать в виде $D \cdot c_1 = D_1$, откуда

$$c_1 = \frac{D_1}{D}.$$

Совершенно аналогично можно получить выражение

$$c_j = \frac{D_j}{D} \qquad (j = 1, 2, \ldots, n), \qquad (20)$$

где

$$D_j = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1,j-1} & b_1 & a_{1,j+1} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2,j-1} & b_2 & a_{2,j+1} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{n,j-1} & b_n & a_{n,j+1} & \ldots & a_{nn} \end{vmatrix} = D_j(b_i)$$

есть определитель, полученный из определителя D заменой его j-го столбца на столбец из чисел $b_1, b_2, \ldots, b_n$.

Мы получили следующий результат:

Если решение системы (17) *существует, то оно выражается через коэффициенты системы и правые части по формулам* (20). В частности, мы получаем, что решение системы (17), если оно существует, *единственно.*

1.72. Теперь остается показать, что решение системы (17) всегда существует. Подставим величины

$$c_j = \frac{D_j}{D} \qquad (j = 1, 2, \ldots, n)$$

в систему (17) на место неизвестных $x_1, x_2, \ldots, x_n$. Покажем, что все уравнения системы (17) при этом обращаются в тождества. Действительно, для i-го уравнения мы получаем

$$\begin{aligned} a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{in}c_n &= a_{i1}\frac{D_1}{D} + a_{i2}\frac{D_2}{D} + \ldots + a_{in}\frac{D_n}{D} = \\ &= \frac{1}{D}\left[a_{i1}(b_1A_{11} + b_2A_{21} + \ldots + b_nA_{n1}) + \right. \\ &\qquad + a_{i2}(b_1A_{12} + b_2A_{22} + \ldots + b_nA_{n2}) + \\ &\qquad \cdots\cdots\cdots\cdots\cdots \\ &\qquad \left. + a_{in}(b_1A_{1n} + b_2A_{2n} + \ldots + b_nA_{nn})\right] = \\ &= \frac{1}{D}\left[b_1(a_{i1}A_{11} + a_{i2}A_{12} + \ldots + a_{in}A_{1n}) + \right. \\ &\qquad \cdots\cdots\cdots\cdots\cdots \\ &\qquad + b_i(a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}) + \\ &\qquad \cdots\cdots\cdots\cdots\cdots \\ &\qquad \left. + b_n(a_{i1}A_{n1} + a_{i2}A_{n2} + \ldots + a_{in}A_{nn})\right]. \end{aligned}$$

Из всех скобок, служащих коэффициентами при величинах $b_1, b_2, \ldots, b_n$, отлична от нуля в силу теорем *1.51* и *1.52* только одна, именно та, которая стоит при величине b_i;

она равна самому определителю D. Следовательно, полученное выражение приводится к виду

$$\frac{1}{D} b_i D = b_i,$$

т. е. совпадает с правой частью i-го уравнения системы.

1.73. Итак, величины c_j действительно образуют решение системы (17). Тем самым мы установили следующее правило для получения решения системы (17) (*правило Крамера*):

Если определитель системы (17) *отличен от нуля, то она имеет одно и только одно решение: значение неизвестного* x_j *равно дроби, знаменателем которой является определитель системы* (17), *а числителем — определитель, получающийся заменой j-го столбца в определителе системы* (17) *на столбец из правых частей системы* ($j=1, \ldots, n$).

Отыскание решения системы (17), таким образом, сводится к вычислению определителей.

См. задачу 11.

Способы решения более общих систем (с определителем, равным нулю, или с числом уравнений, не равным числу неизвестных) будут даны в двух следующих главах.

1.74. Замечание. Иногда встречаются системы линейных уравнений, свободные члены которых являются не числами, а векторами (в аналитической геометрии, в механике). Теорема Крамера и ее вывод остаются справедливыми и для этого случая; следует иметь только в виду, что и значения неизвестных $x_1, x_2, \ldots, x_n$ также будут не числами, а векторами. Например, система

$$\begin{aligned} \boldsymbol{x}_1 + \boldsymbol{x}_2 &= \boldsymbol{i} - 3\boldsymbol{j}, \\ \boldsymbol{x}_1 - \boldsymbol{x}_2 &= \boldsymbol{i} + 5\boldsymbol{j} \end{aligned}$$

имеет решение (единственное)

$$\boldsymbol{c}_1 = \boldsymbol{i} + \boldsymbol{j}, \quad \boldsymbol{c}_2 = -4\boldsymbol{j}.$$

§ 1.8. Миноры произвольного порядка. Теорема Лапласа

1.81. Теорема *1.51* о разложении определителя по строке или столбцу является частным случаем более общей теоремы о разложении определителя по целой совокупности строк или столбцов. Прежде чем сформулировать эту общую теорему (теорему Лапласа), введем некоторые новые определения.

Пусть в квадратной матрице n-го порядка указаны произвольно $k \leqslant n$ различных строк и столько же различных столбцов. Элементы, стоящие на пересечениях этих строк и столбцов, образуют квадратную матрицу порядка k; ее определитель называется *минором k-го порядка* данной матрицы n-го порядка (а также минором k-го порядка определителя D) и обозначается через

$$M = M_{j_1,\ j_2,\ \ldots,\ j_k}^{i_1,\ i_2,\ \ldots,\ i_k},$$

где $i_1, i_2, \ldots, i_k$ — номера выделенных строк, а $j_1, j_2, \ldots, j_k$ — номера выделенных столбцов.

Если зачеркнуть в исходной матрице строки и столбцы, в которых лежит минор M, то оставшиеся элементы снова образуют квадратную матрицу порядка $n-k$; ее определитель называется минором, *дополнительным* к минору M, и обозначается символом

$$\overline{M} = \overline{M}_{j_1,\ j_2,\ \ldots,\ j_k}^{i_1,\ i_2,\ \ldots,\ i_k},$$

где индексы указывают номера в ы ч е р к н у т ы х строк и столбцов.

В частности, если исходный минор имеет порядок 1, т. е. совпадает с некоторым элементом a_{ij} определителя D, то дополнительный минор совпадает с минором M_{ij}, о котором шла речь в *1.53*.

Рассмотрим минор

$$M_1 = M_{1,\ 2,\ \ldots,\ k}^{1,\ 2,\ \ldots,\ k},$$

лежащий в первых k строках и первых k столбцах определителя D; минор, дополнительный к нему, есть минор

$$M_2 = \overline{M}_1 = \overline{M}_{1,\ 2,\ \ldots,\ k}^{1,\ 2,\ \ldots,\ k}.$$

Выделим в правой части формулы (6) все те члены определителя, в которых первые k элементов принадлежат минору M_1, а следовательно, остальные $n-k$ элементов — минору $\overline{M}_1$. Фиксируем сначала один из таких членов с тем, чтобы определить знак, который должен быть ему приписан; обозначим этот член через c. Первые k элементов этого члена определяют некоторый член c_1 минора M_1; обозначим через N_1 число соответствующих отрезков отрицательного наклона; тогда знак, который должен быть поставлен перед членом c_1 в миноре M_1, определяется выражением $(-1)^{N_1}$. Остальные $n-k$ элементов члена c определяют некоторый член c_2 минора M_2; знак, который должен быть поставлен перед этим членом в миноре M_2, определяется выражением $(-1)^{N_2}$, где N_2 — число соответствующих отрезков отрицательного наклона в миноре M_2. Поскольку в матрице определителя D нет ни одного отрезка с отрицательным наклоном, который соединял бы элемент минора M_1 с элементом минора M_2, общее число отрезков отрицательного наклона, соединяющих элементы члена c, равно сумме N_1+N_2. Поэтому знак, который следует поставить перед членом c в определителе D, определяется выражением $(-1)^{N_1+N_2}$ и, следовательно, равен произведению знаков членов c_1 и c_2 в минорах M_1 и M_2. Заметим, далее, что произведение любого члена минора M_1 на любой член минора M_2 дает нам один из выделенных членов определителя D. Отсюда вытекает, что сумма всех выделенных членов в выражении определителя D по формуле (6) равна *произведению миноров* M_1 и M_2.

Теперь мы получим похожий результат для произвольного минора

$$M_1 = M_{j_1,\, j_2,\, \ldots,\, j_k}^{i_1,\, i_2,\, \ldots,\, i_k}$$

с дополнительным минором M_2. Переставляя последовательно соседние строки и столбцы, мы можем минор M_1 перевести в левый верхний угол определителя D; для этого понадобится всего $(i_1-1)+(i_2-2)+\ \ldots\ +(i_k-k)+(j_1-1)+$ $+(j_2-2)+\ldots+(j_k-k)$ перестановок. В результате мы получим определитель D_1 с теми же членами, какие будет иметь исходный определитель D, если его умножить на $(-1)^{i+j}$, где $i=i_1+i_2+\ldots+i_k$, $j=j_1+j_2+\ldots+j_k$. По доказанному, в определителе D_1 сумма всех тех членов, первые k элементов которых входят в минор M_1, равна произведению M_1M_2. Отсюда следует, что сумма соответствующих членов определителя D равна произведению

$$(-1)^{i+j} M_1 M_2 = M_1 A_2,$$

где величина $A_2 = (-1)^{i+j} M_2$ называется *алгебраическим дополнением минора* M_1 *в определителе* D. Иногда употребляют обозначение

$$A_2 = \overline{A}^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}.$$

Фиксируем теперь в определителе D строки с номерами $i_1, i_2, \ldots, i_k$. В каждый член определителя D входят некоторые элементы из этих строк. Если собрать вместе все такие члены, у которых элементы выделенных строк принадлежат к фиксированным столбцам с номерами $j_1, j_2, \ldots, j_k$, то, по доказанному, сумма всех этих членов будет равна произведению минора

$$M^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}$$

на соответствующее алгебраическое дополнение. Все члены определителя, таким образом, можно разбить на группы, каждая из которых определяется заданием k столбцов. Сумма членов в каждой группе равна произведению соответствующего минора и его алгебраического дополнения. Поэтому весь определитель представляется в виде суммы

$$D = \sum (-1)^{i+j} M^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k} \overline{M}^{i_1, i_2, \ldots, i_k}_{j_1, j_2, \ldots, j_k}, \tag{21}$$

причем суммирование производится при фиксированных индексах $i_1, i_2, \ldots, i_k$ (индексы выбранных строк) по всем возможным значениям индексов столбцов $j_1, j_2, \ldots, j_k$ $(1 \leqslant j_1 < j_2 < \ldots < j_k \leqslant n)$. Свойство определителя D, выражающееся равенством (21), и называется *теоремой Лапласа*. Очевидно, формула (21) действительно является обобщением формулы разложения определителя по одной строке, полученной в *1.54*.

Аналогичная формула справедлива для разложения определителя D по фиксированной системе столбцов.

1.82. Пример. Определитель вида

$$D=\begin{vmatrix} a_{11} & \dots & a_{1k} & a_{1,k+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & a_{2,k+1} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & \dots & a_{kk} & a_{k,k+1} & \dots & a_{kn} \\ 0 & \dots & 0 & a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & a_{n,k+1} & \dots & a_{nn} \end{vmatrix},$$

все элементы которого, находящиеся в первых k столбцах и последних $n-k$ строках, равны нулю, называется *квазитреугольным*. Для его вычисления разложим определитель по первым k строкам с помощью теоремы Лапласа. В сумме (21) останется только одно слагаемое, и мы получим

$$D=\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \cdot & \cdot & \cdot \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \cdot \begin{vmatrix} a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot \\ a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

См. задачу 12.

§ 1.9. О линейной зависимости между столбцами

1.91. Пусть нам дано несколько, например m, числовых столбцов по n чисел в каждом:

$$A_1=\begin{Vmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{n1} \end{Vmatrix}, \quad A_2=\begin{Vmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{n2} \end{Vmatrix}, \quad \dots, \quad A_m=\begin{Vmatrix} a_{1m} \\ a_{2m} \\ \cdot \\ \cdot \\ \cdot \\ a_{nm} \end{Vmatrix}.$$

Умножим каждый элемент первого столбца на некоторое число λ_1, каждый элемент второго столбца — на число λ_2, ..., каждый элемент последнего, m-го, столбца на число λ_m и сложим соответствующие элементы полученных столбцов. В результате получится некоторый новый числовой столбец, элементы которого мы обозначим буквами $c_1, c_2, \dots, c_n$. Все эти действия можно наглядно представить с помощью следующей схемы:

$$\lambda_1 \left\| \begin{matrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{n1} \end{matrix} \right\| + \lambda_2 \left\| \begin{matrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{n2} \end{matrix} \right\| + \ldots + \lambda_m \left\| \begin{matrix} a_{1m} \\ a_{2m} \\ \cdot \\ \cdot \\ \cdot \\ a_{nm} \end{matrix} \right\| = \left\| \begin{matrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{matrix} \right\|,$$

или, короче,

$$\lambda_1 A_1 + \lambda_2 A_2 + \ldots + \lambda_m A_m = C,$$

где через C обозначен получившийся столбец. Этот столбец C называется *линейной комбинацией взятых столбцов* $A_1, A_2, \ldots, A_m$; числа $\lambda_1, \lambda_2, \ldots, \lambda_m$ называются *коэффициентами* этой линейной комбинации.

Частными случаями линейной комбинации являются *сумма столбцов* (когда коэффициенты $\lambda_1, \ldots, \lambda_m$ равны 1) и произведение столбца на число (когда $m = 1$).

Теперь представим себе, что наши столбцы взяты не сами по себе, а входят в состав некоторого определителя D n-го порядка. Докажем следующую теорему.

Теорема. *Если один из столбцов определителя D является линейной комбинацией других столбцов, то $D = 0$.*

Доказательство. Пусть, например, q-й столбец определителя D является линейной комбинацией j-го, k-го, ..., p-го столбцов этого определителя с коэффициентами, соответственно, $\lambda_j, \lambda_k, \ldots, \lambda_p$. Тогда, вычитая из q-го столбца j-й столбец, умноженный на λ_j, затем k-й столбец, умноженный на λ_k, ..., наконец, p-й столбец, умноженный на λ_p, мы согласно *1.47б* не изменим величины определителя D; но в результате q-й столбец будет состоять из одних нулей, откуда вытекает, что $D = 0$.

Замечательно, что справедлива и обратная теорема: *если заданный определитель D равен нулю, то один* (по меньшей мере) *из его столбцов является линейной комбинацией других столбцов*. Доказательство этой теоремы требует некоторых предварительных построений, к которым мы и переходим.

1.92. Пусть опять имеется m числовых столбцов по n элементов в каждом. Мы можем записать их в виде матрицы с n строками и m столбцами:

$$A = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{Vmatrix}.$$

Если фиксировать некоторое число k столбцов этой матрицы и такое же число k ее строк, то элементы, стоящие на пересечениях указанных столбцов и строк, образуют квадратную матрицу k-го порядка. Ее определитель называется *минором k-го порядка матрицы A*; он может быть равен нулю или отличен от нуля. Если среди чисел a_{ij} есть отличные от нуля (что мы всегда будем предполагать), то всегда можно указать натуральное число r, обладающее такими свойствами:

а) у матрицы A имеется минор r-го порядка, отличный от нуля;

б) всякий минор матрицы A, имеющий порядок $r+1$ или выше (если вообще таковые существуют), равен нулю.

Число r, обладающее указанными свойствами, называется *рангом* матрицы A. Если все a_{ik} равны нулю, то ранг матрицы считается равным нулю ($r=0$). В дальнейшем мы предполагаем, что $r>0$. Тот минор r-го порядка, который отличен от нуля, называется *базисным минором* матрицы A. (Разумеется, у матрицы A может быть и несколько базисных миноров; но все они имеют один и тот же порядок r.) Столбцы, на которых построен базисный минор, называются *базисными столбцами*.

1.93. Имеет место следующая важная теорема:

Теорема (о базисном миноре). *Любой столбец матрицы A является линейной комбинацией ее базисных столбцов.*

Доказательство. Предположим для определенности, что базисный минор матрицы A расположен в первых r строках и первых r столбцах матрицы A. Пусть s — любое целое число от 1 до m, а k — любое целое число от 1 до n.

Рассмотрим определитель $(r+1)$-го порядка

$$D=\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} & a_{1s} \\ a_{21} & a_{22} & \ldots & a_{2r} & a_{2s} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{r1} & a_{r2} & \ldots & a_{rr} & a_{rs} \\ a_{k1} & a_{k2} & \ldots & a_{kr} & a_{ks} \end{vmatrix}.$$

Если $k\leqslant r$, определитель D, очевидно, равен нулю, так как в нем имеются две одинаковые строки. Аналогично $D=0$ при $s\leqslant r$. Если $k>r$ и $s>r$, то определитель D также равен нулю как минор $(r+1)$-го порядка матрицы ранга r. Следовательно, $D=0$ при любых значениях k и s.

Разложим определитель D по последней строке; мы получим равенство

$$a_{k1}A_{k1}+a_{k2}A_{k2}+\ldots+a_{kr}A_{kr}+a_{ks}A_{ks}=0, \tag{22}$$

где числа A_{k1}, A_{k2}, ..., A_{kr}, A_{ks} означают соответственно алгебраические дополнения элементов a_{k1}, a_{k2}, ..., a_{kr}, a_{ks}, находящихся в нижней строке определителя D. Эти алгебраические дополнения не зависят от числа k, так как образуются с помощью элементов a_{ij} с $i\leqslant r$; поэтому мы можем ввести обозначения

$$A_{k1}=c_1,\quad A_{k2}=c_2,\quad \ldots,\quad A_{kr}=c_r,\quad A_{ks}=c_s.$$

Подставляя в равенство (22) последовательно значения $k=1, 2, \ldots, n$, получим систему равенств

$$\left.\begin{array}{l} c_1a_{11}+c_2a_{12}+\ldots+c_ra_{1r}+c_sa_{1s}=0, \\ c_1a_{21}+c_2a_{22}+\ldots+c_ra_{2r}+c_sa_{2s}=0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ c_1a_{n1}+c_2a_{n2}+\ldots+c_ra_{nr}+c_sa_{ns}=0. \end{array}\right\} \tag{23}$$

Число $c_s=A_{ks}$ отлично от нуля, так как A_{ks} есть базисный минор матрицы A. Разделив каждое из равенств (23) на c_s, перенеся все слагаемые, кроме последнего, в правую часть и обозначив $-\dfrac{c_j}{c_s}$ через λ_j $(j=1, 2, \ldots, r)$, мы получим

$$\left.\begin{array}{l} a_{1s}=\lambda_1a_{11}+\lambda_2a_{12}+\ldots+\lambda_ra_{1r}, \\ a_{2s}=\lambda_1a_{21}+\lambda_2a_{22}+\ldots+\lambda_ra_{2r}, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{ns}=\lambda_1a_{n1}+\lambda_2a_{n2}+\ldots+\lambda_ra_{nr}. \end{array}\right\} \tag{24}$$

Эти равенства показывают, что s-й столбец матрицы A является линейной комбинацией первых r столбцов этой матрицы (с коэффициентами $\lambda_1, \lambda_2, \ldots, \lambda_r$). Поскольку s может быть любым числом от 1 до m, наша теорема полностью доказана.

1.94. Мы можем теперь доказать сформулированное в конце *1.91* обращение теоремы *1.91.*

Теорема. *Если некоторый определитель D порядка n равен нулю, то у него имеется столбец, который является линейной комбинацией других столбцов.*

Доказательство. Рассмотрим матрицу определителя D. Поскольку $D=0$, базисный минор этой матрицы имеет порядок $r<n$. Поэтому после выделения r базисных столбцов мы сможем найти еще по меньшей мере один столбец, не попавший в число базисных. В силу теоремы о базисном миноре он представляет собой некоторую линейную комбинацию базисных столбцов. Итак, мы нашли в определителе D столбец, который является линейной комбинацией других столбцов, что и утверждалось.

Заметим, что в состав этой линейной комбинации можно включить и все оставшиеся столбцы определителя D, поставив перед ними, например, нулевые коэффициенты.

1.95. Полученные результаты можно сформулировать в несколько более симметричном виде.

Если коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_m$ линейной комбинации m числовых столбцов $A_1, A_2, \ldots, A_m$ (*1.91*) взять равными нулю, то очевидно, что в результате получится *нулевой столбец*, т. е. столбец, состоящий из одних нулей. Но возможно, что нулевой столбец получается из заданных столбцов не только таким способом, а и с помощью коэффициентов $\lambda_1, \lambda_2, \ldots, \lambda_m$, из которых не все равны нулю. В этом случае взятые столбцы $A_1, A_2, \ldots, A_m$ называются *линейно зависимыми.*

Например, столбцы

$$A_1=\begin{Vmatrix}1\\2\\3\\4\end{Vmatrix}, \quad A_2=\begin{Vmatrix}2\\4\\6\\8\end{Vmatrix}, \quad A_3=\begin{Vmatrix}1\\1\\1\\1\end{Vmatrix}$$

линейно зависимы, так как нулевой столбец равен линейной комбинации

$$2\cdot A_1-1\cdot A_2+0\cdot A_3.$$

Определение линейной зависимости можно сформулировать более подробно так: столбцы

$$A_1=\left\|\begin{matrix}a_{11}\\a_{21}\\ \cdot\\ \cdot\\ \cdot\\ a_{n1}\end{matrix}\right\|,\quad A_2=\left\|\begin{matrix}a_{12}\\a_{22}\\ \cdot\\ \cdot\\ \cdot\\ a_{n2}\end{matrix}\right\|,\quad \ldots,\quad A_m=\left\|\begin{matrix}a_{1m}\\a_{2m}\\ \cdot\\ \cdot\\ \cdot\\ a_{nm}\end{matrix}\right\|$$

называются *линейно зависимыми*, если существуют величины $\lambda_1, \lambda_2, \ldots, \lambda_m$, не все равные нулю, такие, что удовлетворяется система уравнений

$$\left.\begin{matrix}\lambda_1a_{11}+\lambda_2a_{12}+\ldots+\lambda_ma_{1m}=0,\\ \lambda_1a_{21}+\lambda_2a_{22}+\ldots+\lambda_ma_{2m}=0,\\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\ \lambda_1a_{n1}+\lambda_2a_{n2}+\ldots+\lambda_ma_{nm}=0,\end{matrix}\right\}$$

или, что то же самое,

$$\lambda_1A_1+\lambda_2A_2+\ldots+\lambda_mA_m=0\ ^{*)}.$$

Если один из столбцов $A_1, A_2, \ldots, A_m$, например последний, *является линейной комбинацией остальных,*

$$A_m=\lambda_1A_1+\lambda_2A_2+\ldots+\lambda_{m-1}A_{m-1}, \tag{25}$$

то столбцы $A_1, A_2, \ldots, A_m$ *линейно зависимы.* Действительно, соотношение (25) равносильно соотношению

$$\lambda_1A_1+\lambda_2A_2+\ldots+\lambda_{m-1}A_{m-1}+(-1)A_m=0;$$

следовательно, существует линейная комбинация столбцов $A_1, A_2, \ldots, A_m$, коэффициенты которой не все равны нулю (в частности, последний коэффициент равен -1) и которая дает в результате нулевой столбец; это и означает линейную зависимость столбцов $A_1, A_2, \ldots, A_m$.

Обратно, *если между столбцами* $A_1, A_2, \ldots, A_m$ *имеется линейная зависимость, то (по меньшей мере) один из этих*

*) Здесь в правой части символ 0 означает нулевой столбец.

столбцов является линейной комбинацией остальных. В самом деле, пусть в равенстве

$$\lambda_1 A_1 + \lambda_2 A_2 + \ldots + \lambda_{m-1} A_{m-1} + \lambda_m A_m = 0, \qquad (26)$$

выражающем линейную зависимость столбцов $A_1, A_2, \ldots, A_m$, отличен от нуля, например, коэффициент λ_m. Тогда соотношение (26) равносильно соотношению

$$A_m = -\frac{\lambda_1}{\lambda_m} A_1 - \frac{\lambda_2}{\lambda_m} A_2 - \ldots - \frac{\lambda_{m-1}}{\lambda_m} A_{m-1},$$

которое показывает, что столбец A_m является линейной комбинацией столбцов $A_1, A_2, \ldots, A_{m-1}$.

Итак, *столбцы $A_1, A_2, \ldots, A_m$ линейно зависимы тогда и только тогда, когда один из этих столбцов является линейной комбинацией остальных столбцов.*

1.96. Теоремы *1.91* и *1.94* показывают, что определитель D равен нулю тогда и только тогда, когда один из его столбцов является линейной комбинацией остальных столбцов.

Используя полученный в *1.95* результат, мы можем сформулировать следующую теорему:

Теорема. *Определитель D равен нулю тогда и только тогда, когда между его столбцами существует линейная зависимость.*

1.97. Так как при транспонировании определителя его величина не меняется (*1.41*), а столбцы заменяются строками, то во всех формулировках настоящего параграфа столбцы можно заменить на строки. В частности, имеет место следующий результат:

Определитель D равен нулю тогда и только тогда, когда между его строками существует линейная зависимость.

См. задачи 13—14.

ЗАДАЧИ

1. С каким знаком в определитель 6-го порядка входят члены:
а) $a_{23}a_{31}a_{42}a_{56}a_{14}a_{65}$,
б) $a_{32}a_{43}a_{14}a_{51}a_{66}a_{25}$?

2. Выписать все члены, входящие в состав определителя 4-го порядка со знаком — и содержащие множителем a_{23}.

3. С каким знаком входит в определитель n-го порядка член $a_{1n}\, a_{2,\, n-1} \ldots a_{n1}$?

4. Показать, что из $n!$ членов определителя ровно половина $\left(\frac{n!}{2}\right)$ получает по определению § 1.3 знак $+$, а вторая половина — знак $-$.

5. Вычислить определитель

$$\Delta = \begin{vmatrix} am+bp & an+bq \\ cm+dp & cn+dq \end{vmatrix},$$

разложив его на слагаемые.

6. Числа 20 604, 53 227, 25 755, 20 927 и 78 421 делятся на 17. Доказать, что определитель

$$\begin{vmatrix} 2 & 0 & 6 & 0 & 4 \\ 5 & 3 & 2 & 2 & 7 \\ 2 & 5 & 7 & 5 & 5 \\ 2 & 0 & 9 & 2 & 7 \\ 7 & 8 & 4 & 2 & 1 \end{vmatrix}$$

также делится на 17.

7. Вычислить определители

$$\Delta_1 = \begin{vmatrix} 246 & 427 & 327 \\ 1014 & 543 & 443 \\ -342 & 721 & 621 \end{vmatrix}, \qquad \Delta_2 = \begin{vmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 & 1 \\ 1 & 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 1 & 6 \end{vmatrix}.$$

8. Вычислить определитель

$$P(x) = \begin{vmatrix} 1 & 1 & 2 & 3 \\ 1 & 2-x^2 & 2 & 3 \\ 2 & 3 & 1 & 5 \\ 2 & 3 & 1 & 9-x^2 \end{vmatrix}.$$

9. Вычислить определитель n-го порядка

$$\Delta = \begin{vmatrix} x & a & a & \ldots & a \\ a & x & a & \ldots & a \\ a & a & x & \ldots & a \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a & a & a & \ldots & x \end{vmatrix}.$$

10. Вычислить *определитель Вандермонда*

$$\Delta(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{vmatrix}.$$

11. Решить систему уравнений

$$\begin{aligned}
x_1+2x_2+3x_3+4x_4+5x_5&=13,\\
2x_1+x_2+2x_3+3x_4+4x_5&=10,\\
2x_1+2x_2+x_3+2x_4+3x_5&=11,\\
2x_1+2x_2+2x_3+x_4+2x_5&=6,\\
2x_1+2x_2+2x_3+2x_4+x_5&=3.
\end{aligned}$$

12. Сформулировать и доказать теорему, находящуюся в таком же логическом отношении к теореме Лапласа, в каком теорема *1.52* находится к теореме *1.51*.

13. Построить четыре столбца из четырех чисел в каждом, которые не были бы линейно зависимыми.

14. Показать, что если строки некоторого определителя n-го порядка линейно зависимы, то линейно зависимы и его столбцы.

ЛИНЕЙНЫЕ ПРОСТРАНСТВА

§ 2.1. Определение

2.11. В аналитической геометрии и в механике используются направленные отрезки — векторы. Для векторов установлены по определенным правилам действия: известно, что означает сумма двух векторов и что означает произведение вектора на вещественное число*). При этом выполняются обычные законы арифметики.

Определение линейного пространства обобщает определение совокупности всех векторов. Обобщение производится, во-первых, путем отвлечения от конкретной природы объектов (направленных отрезков) с сохранением свойств действий над ними, во-вторых, путем отвлечения от конкретной природы допустимых множителей (вещественных чисел). Таким образом, получается следующее определение:

Множество $\mathbf{K}$ называется *линейным* (*аффинным*) *пространством над полем* K, если а) имеется правило (правило сложения), которое позволяет для каждых двух элементов x и y из $\mathbf{K}$ построить третий элемент $z \in \mathbf{K}$**), называемый *суммой элементов* x и y и обозначаемый $x+y$; б) име-

---

*) Мы не касаемся пока других векторных операций — скалярного и векторного произведений. Во всяком случае оба эти произведения не могут играть той роли, которую играет обычное произведение в поле вещественных чисел; скалярное произведение векторов уже не есть вектор; векторное произведение векторов хоть и есть вектор, но эта операция, в отличие от умножения вещественных чисел, некоммутативна.

**) Мы используем здесь и в дальнейшем некоторые обозначения теории множеств. Запись $a \in A$ означает, что элемент a входит в множество A; запись $B \subset A$ означает, что множество B является частью множества A (причем B может и совпадать с A). Два соотношения $B \subset A$, $A \subset B$ равносильны утверждению, что множества A и B совпадают. Знаки $\in$, $\subset$ называются знаками *включения*.

ется правило (правило умножения на число), которое позволяет построить для каждого элемента $x \in \mathrm{K}$ и любого числа $\lambda \in K$ элемент $u \in \mathrm{K}$, называемый *произведением элемента* x *на число* λ и обозначаемый λx; в) правила а) и б) удовлетворяют аксиомам, перечисленным в *2.12—2.13.*

Элементы линейного пространства мы будем называть *векторами*, невзирая на то, что по своей конкретной природе они могут быть вовсе и не похожи на привычные нам направленные отрезки. Геометрические представления, связанные с названием «векторы», помогут нам уяснить и часто предвидеть нужные результаты, а также находить прямой геометрический смысл в различных фактах из алгебры и анализа, который без того не был бы очевидным. В частности, в следующей главе мы получим простую геометрическую характеристику всех решений однородной или неоднородной системы линейных уравнений.

2.12. Предполагается, что правило сложения обладает следующими свойствами:

1) $x+y=y+x$ для любых x и y из K;

2) $(x+y)+z=x+(y+z)$ для любых x, y, z из K;

3) существует элемент 0 (*нуль-вектор*) такой, что $x+0=x$ для любого $x \in \mathrm{K}$;

4) для каждого $x \in \mathrm{K}$ существует элемент $y \in \mathrm{K}$ такой, что $x+y=0$ (*противоположный элемент*).

2.13. Предполагается, что правило умножения на число обладает следующими свойствами:

5) $1 \cdot x=x$ для любого $x \in \mathrm{K}$;

6) $\alpha(\beta x)=(\alpha\beta)x$ для любого $x \in \mathrm{K}$ и любых α и β из K;

7) $(\alpha+\beta)x=\alpha x+\beta x$ для любого $x \in \mathrm{K}$ и любых α и β из K;

8) $\alpha(x+y)=\alpha x+\alpha y$ для любых x и y из K и любого $\alpha \in K$.

2.14. Из аксиом 1)—8) можно получить в первую очередь следующие теоремы:

а. Теорема. *В любом линейном пространстве существует единственный нуль.*

Доказательство. Существование хотя бы одного нуля утверждается в аксиоме 3). Допустим, что в пространстве K

имеются два нуля: 0_1 и 0_2. Полагая в аксиоме 3) $x=0_1$, $0=0_2$, мы получаем

$$0_1+0_2=0_1.$$

Полагая в той же аксиоме $x=0_2$, $0=0_1$, мы получаем

$$0_2+0_1=0_2.$$

Сравнивая первое из полученных равенств со вторым и пользуясь аксиомой 1), находим $0_1=0_2$, что и требовалось.

б. Теорема. *В любом линейном пространстве для каждого элемента существует единственный противоположный элемент.*

Доказательство. Существование хотя бы одного противоположного элемента утверждается в аксиоме 4). Допустим, что для некоторого элемента x имеется два противоположных элемента y_1 и y_2. Прибавим к обеим частям равенства $x+y_1=0$ элемент y_2; используя аксиомы 2) и 3), мы получаем

$$y_2+(x+y_1)=(y_2+x)+y_1=0+y_1=y_1,$$
$$y_2+(x+y_1)=y_2+0=y_2,$$

откуда $y_1=y_2$, что и требовалось.

в. Теорема. *Для всякого элемента x в любом линейном пространстве имеет место равенство*

$$0\cdot x=0$$

(в правой части равенства 0 означает нуль-вектор, в левой — число 0).

Доказательство. Рассмотрим элемент $0\cdot x+1\cdot x$; используя аксиомы 7) и 5), мы получаем

$$0\cdot x+1\cdot x=(0+1)\,x=1\cdot x=x,\quad 0\cdot x+1\cdot x=0\cdot x+x,$$

откуда

$$x=0\cdot x+x;$$

прибавляя к обеим частям равенства противоположный к x элемент y, находим

$$0=x+y=(0\cdot x+x)+y=0\cdot x+(x+y)=0\cdot x+0=0\cdot x,$$

откуда

$$0 = 0 \cdot x,$$

что и требуется.

г. Теорема. *Для всякого элемента x в любом линейном пространстве противоположным элементом служит*

$$y = (-1) \cdot x.$$

Доказательство. Составим сумму $x + y$; используя аксиомы и теорему *в*, находим

$$x + y = 1 \cdot x + (-1) \cdot x = (1 - 1) \cdot x = 0 \cdot x = 0,$$

что и требуется.

д. Мы будем обозначать теперь элемент, противоположный к данному элементу x, через $-x$; доказанная теорема *г* делает естественным это обозначение.

Наличие противоположного элемента позволяет ввести операцию вычитания. Именно, *разность* $x - y$ определяется как сумма x и $-y$. Это определение согласуется с определением вычитания в арифметике.

2.15. Линейное пространство над полем R вещественных чисел мы будем называть *вещественным* и обозначать через R. Линейное пространство над полем C комплексных чисел мы будем называть *комплексным* и обозначать через C. Если указаны природа элементов $x, y, z, \ldots$ и правила действий над ними [причем должны быть выполнены аксиомы 1)—8)], мы будем называть линейное пространство *конкретным* и использовать для него, как правило, индивидуальное обозначение.

В дальнейшем для нас будут особенно важны следующие три типа конкретных пространств:

а. Пространство V_3. Элементы этого пространства — свободные векторы в пространстве, рассматриваемые в аналитической геометрии. Каждый вектор характеризуется длиной и направлением*). Сложение векторов определено обычным образом по правилу параллелограмма. Умножение вектора на вещественное число λ определено также обычным образом (именно, длина вектора умножается на $|\lambda|$,

*) За исключением нуль-вектора, длина которого равна нулю, а направление произвольно.

направление при $\lambda>0$ остается неизменным, при $\lambda<0$ заменяется на противоположное). Легко проверить, что все аксиомы 1)—8) здесь выполнены. Аналогичные совокупности векторов на плоскости и на прямой, также представляющие собой линейные пространства, обозначим соответственно через V_2 и V_1; V_1, V_2, V_3 — линейные пространства над полем R.

б. Пространство K_n. Элемент этого пространства — любая совокупность $x=(\xi_1, \xi_2, \ldots, \xi_n)$ n чисел из поля K. Эти числа $\xi_1, \xi_2, \ldots, \xi_n$ будем называть *координатами* элемента x. Действия сложения и умножения на число $\lambda\in K$ производятся по следующим правилам:

$$(\xi_1, \xi_2, \ldots, \xi_n)+(\eta_1, \eta_2, \ldots, \eta_n)=\\ =(\xi_1+\eta_1, \xi_2+\eta_2, \ldots, \xi_n+\eta_n), \quad (1)$$

$$\lambda(\xi_1, \xi_2, \ldots, \xi_n)=(\lambda\xi_1, \lambda\xi_2, \ldots, \lambda\xi_n). \quad (2)$$

Легко проверить, что аксиомы 1)—8) удовлетворены. В частности, элемент 0 есть совокупность n нулей: $0=(0, 0, \ldots, 0)$. Фактически мы имели дело с элементами этого пространства в § 1.9; только мы записывали тогда их не в форме числовой строки, а в форме столбца. Если K есть поле R вещественных чисел, обозначение K_n заменяется на R_n. Если K есть поле C комплексных чисел, обозначение K_n заменяется на C_n.

в. Пространство $R(a, b)$. Элемент этого пространства — любая вещественная непрерывная функция $x=x(t)$, определенная на отрезке $a\leqslant t\leqslant b$. Действия сложения функций и умножения их на вещественные числа определяются по правилам анализа; выполнение аксиом 1)—8) очевидно. При этом элемент 0 есть функция, тождественно равная нулю. Пространство $R(a, b)$ есть линейное пространство над полем R вещественных чисел.

г. Пространство $C(a, b)$ соответственно есть пространство всех комплекснозначных непрерывных функций на отрезке $a\leqslant t\leqslant b$. Это пространство есть линейное пространство над полем комплексных чисел.

2.16. Заметим, что все свойства элементов конкретных пространств (например, векторов пространства V_3), основанные только на аксиомах 1)—8), справедливы и для элементов любых линейных пространств. Например, анализируя доказательство теоремы Крамера о решении системы линейных уравнений

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n, \end{array}\right\}$$

мы можем заметить, что в той части, которая касалась величин $b_1, b_2, \ldots, b_n$, оно основывалось только на том факте, что эти величины можно было складывать и умножать на числа из K, причем использовались правила 1)—8). Это позволило обобщить теорему Крамера на системы, в которых величины $b_1, b_2, \ldots, b_n$ суть векторы (элементы пространства V_3), как мы уже указывали в *1.74*. Это позволяет, далее, утверждать, что теорема Крамера справедлива и для систем, в которых величины $b_1, b_2, \ldots, b_n$ являются элементами произвольного линейного пространства K. Отметим только, что значения неизвестных $x_1, x_2, \ldots, x_n$ будут тогда также элементами этого пространства K, линейно выражающимися через величины $b_1, b_2, \ldots, b_n$.

См. задачи 1—3.

2.17. Замечание. В аналитической геометрии иногда бывает удобно рассматривать векторы не свободные, а закрепленные своим началом в начале координат. Такое рассмотрение удобно тем, что при этом каждый вектор ассоциируется с некоторой точкой пространства — своим концом и каждая точка пространства может быть определена соответствующим вектором — так называемым *радиусом-вектором* этой точки. Имея в виду эту картину, мы будем иногда называть элементы линейного пространства не векторами, а *точками*. Разумеется, такая перемена названия не сопровождается никакими изменениями в определениях и апеллирует лишь к нашим геометрическим представлениям.

§ 2.2. Линейная зависимость

2.21. Пусть $x_1, x_2, \ldots, x_k$ — векторы линейного пространства K над полем K и $\alpha_1, \alpha_2, \ldots, \alpha_k$ — числа из поля K. Вектор

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k$$

называется *линейной комбинацией* векторов $x_1, x_2, \ldots, x_k$; числа $\alpha_1, \alpha_2, \ldots, \alpha_k$ — *коэффициенты* этой линейной комбинации.

Если $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, то в силу теоремы *2.14 в* мы получаем, что $y = 0$. Но может быть и так, что существует линейная комбинация векторов $x_1, x_2, \ldots, x_k$, в которой не все коэффициенты равны нулю и которая, тем не менее, дает в результате нуль-вектор; в этом случае векторы $x_1, x_2, \ldots, x_k$ называются линейно зависимыми. Иными словами, векторы $x_1, x_2, \ldots, x_k$ называются *линейно зависимыми*, если существуют числа $\alpha_1, \alpha_2, \ldots, \alpha_k$, не все равные нулю и такие, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0. \tag{3}$$

Если же равенство (3) возможно в единственном случае, когда $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, векторы $x_1, x_2, \ldots, x_k$ называются *линейно независимыми*.

2.22. Примеры.

а. В линейном пространстве V_3 линейная зависимость двух векторов означает, что они параллельны одной и той же прямой; линейная зависимость трех векторов — что они параллельны одной и той же плоскости. Всякие четыре вектора линейно зависимы.

б. Выясним, что означает линейная зависимость векторов $x_1, x_2, \ldots, x_k$ линейного пространства K_n. Пусть вектор x_i имеет координаты $\xi_1^{(i)}, \xi_2^{(i)}, \ldots, \xi_n^{(i)}$ $(i = 1, 2, \ldots, k)$; тогда линейная зависимость

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0$$

означает, что выполняется n равенств

$$\left.\begin{array}{l} \alpha_1 \xi_1^{(1)} + \alpha_2 \xi_1^{(2)} + \ldots + \alpha_k \xi_1^{(k)} = 0, \\ \alpha_1 \xi_2^{(1)} + \alpha_2 \xi_2^{(2)} + \ldots + \alpha_k \xi_2^{(k)} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_1 \xi_n^{(1)} + \alpha_2 \xi_n^{(2)} + \ldots + \alpha_k \xi_n^{(k)} = 0, \end{array}\right\} \tag{4}$$

причем среди постоянных $\alpha_1, \alpha_2, \ldots, \alpha_k$ имеются отличные от нуля; это — то самое определение линейной зависимости, которое мы дали в *1.95* для числовых столбцов.

Таким образом, вопрос о линейной зависимости векторов $x_1, x_2, \ldots, x_k$ в общем случае сводится к вопросу о существовании ненулевого решения у однородной системы уравнений с коэффициентами, равными соответствующим координатам данных векторов. В следующей главе этот вопрос будет полностью решен (*3.21*), и тем самым будет получено правило, позволяющее судить о линейной зависимости или независимости данных векторов пространства K_n по их координатам.

в. Но в некоторых случаях мы можем и теперь уже делать выводы о линейной зависимости или независимости данной системы векторов. Пусть, например, в пространстве K_n взяты n векторов

$$e_1=(1, 0, 0, \ldots, 0), \quad e_2=(0, 1, 0, \ldots, 0), \quad \ldots \\ \ldots, \quad e_n=(0, 0, 0, \ldots, 0, 1).$$

Система (4) для этих векторов принимает вид

$$\left.\begin{array}{l} \alpha_1\cdot 1+\alpha_2\cdot 0+\alpha_3\cdot 0+\ldots+\alpha_n\cdot 0=0, \\ \alpha_1\cdot 0+\alpha_2\cdot 1+\alpha_3\cdot 0+\ldots+\alpha_n\cdot 0=0, \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ \alpha_1\cdot 0+\alpha_2\cdot 0+\alpha_3\cdot 0+\ldots+\alpha_n\cdot 1=0 \end{array}\right\}$$

и, очевидно, допускает единственное решение $\alpha_1=\alpha_2=\ldots \ldots=\alpha_n=0$. Таким образом, векторы $e_1, e_2, \ldots, e_n$ в пространстве K_n линейно независимы.

См. задачу 4.

г. Линейная зависимость векторов $x_1=x_1(t), x_2= =x_2(t), \ldots, x_k=x_k(t)$ пространства $R(a, b)$ (или $C(a, b)$) означает, что между функциями $x_1(t), x_2(t), \ldots, x_k(t)$ имеется соотношение

$$\alpha_1 x_1(t)+\alpha_2 x_2(t)+\ldots+\alpha_k x_k(t)\equiv 0,$$

причем вещественные (комплексные) постоянные $\alpha_1, \alpha_2, \ldots \ldots, \alpha_k$ не все равны нулю. Например, функции $x_1(t)=\cos^2 t$, $x_2(t)=\sin^2 t$, $x_3(t)=1$ линейно зависимы, так как имеет место соотношение

$$x_1(t)+x_2(t)-x_3(t)\equiv 0.$$

Проверим, что функции $1, t, t^2, \ldots, t^k$ линейно независимы. Допустим, что существует соотношение

$$\alpha_0 \cdot 1 + \alpha_1 t + \ldots + \alpha_k t^k \equiv 0. \tag{5}$$

Тогда, последовательно дифференцируя k раз равенство (5), мы получим систему $k+1$ уравнений относительно величин $\alpha_0, \alpha_1, \ldots, \alpha_k$ с определителем, заведомо отличным от нуля (*1.55б*); решая систему по правилу Крамера (*1.73*), мы получаем, что

$$\alpha_0 = \alpha_1 = \ldots = \alpha_k = 0.$$

Следовательно, функции $1, t, t^2, \ldots, t^k$ линейно независимы в пространстве $R(a, b)$, что мы и утверждали.

См. задачу 5.

2.23. Отметим два простых свойства систем векторов, связанных с линейной зависимостью.

а. Лемма. *Если некоторые из векторов $x_1, x_2, \ldots, x_k$ линейно зависимы, то и вся система $x_1, x_2, \ldots, x_k$ линейно зависима.*

Доказательство. Без ограничения общности можно принять, что линейно зависимы векторы $x_1, x_2, \ldots, x_j$ $(j < k)$; таким образом, имеет место соотношение

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_j x_j = 0,$$

где среди постоянных $\alpha_1, \alpha_2, \ldots, \alpha_j$ имеются отличные от нуля. В силу теоремы *2.14в* и аксиомы *2.12* 3) справедливо равенство

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_j x_j + 0 \cdot x_{j+1} + \ldots + 0 \cdot x_k = 0;$$

оно показывает, что векторы $x_1, x_2, \ldots, x_k$ также линейно зависимы, так как среди постоянных $\alpha_1, \alpha_2, \ldots, \alpha_j, 0, \ldots, 0$ имеются отличные от нуля.

б. Лемма. *Векторы $x_1, x_2, \ldots, x_k$ линейно зависимы тогда и только тогда, когда можно представить один из этих векторов в виде линейной комбинации других.*

Предложение, аналогичное сформулированному, уже встречалось нам прежде: мы доказывали его в *1.95* для числовых столбцов. Если мы просмотрим еще раз это доказательство, то увидим, что оно основано только на возможности производить со столбцами операции сложения и умножения на число. Следовательно, это доказательство можно провести для элементов любого линейного пространства. Вместе с этим и наша лемма оказывается справедливой для любого линейного пространства, что нам и требуется.

§ 2.3. Базис, координаты, размерность

2.31. Система линейно независимых векторов $e_1, e_2, \ldots, e_n$ некоторого линейного пространства K образует, по определению, *базис* пространства K, если для всякого вектора $x \in$ K существует разложение

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n \qquad (\xi_j \in K,\ j = 1, \ldots, n). \tag{6}$$

Легко видеть, что при указанных условиях *коэффициенты разложения* (6) *определяются единственным образом.* Действительно, если бы для некоторого вектора x можно было написать два разложения

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n,$$
$$x = \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n,$$

то, вычитая почленно, мы получили бы равенство

$$0 = (\xi_1 - \eta_1) e_1 + (\xi_2 - \eta_2) e_2 + \ldots + (\xi_n - \eta_n) e_n,$$

из которого в силу предположенной линейной независимости векторов $e_1, e_2, \ldots, e_n$ мы получили бы, что

$$\xi_1 = \eta_1, \quad \xi_2 = \eta_2, \ \ldots, \ \xi_n = \eta_n.$$

Эти единственным образом определяемые числа $\xi_1, \xi_2, \ldots, \xi_n$ называются *координатами вектора x относительно базиса* $e_1, e_2, \ldots, e_n$.

2.32. Примеры.

а. В пространстве V_3 хорошо известный базис образует тройка единичных взаимно ортогональных векторов $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$. Координаты ξ_1, ξ_2, ξ_3 вектора $\boldsymbol{x}$ относительно этого базиса суть проекции вектора $\boldsymbol{x}$ на координатные оси.

б. В пространстве K_n примером базиса служит система векторов $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, ..., $e_n = (0, 0, \ldots, 1)$, рассмотренная уже нами в *2.22в*. Действительно, для любого вектора $x = (\xi_1, \xi_2, \ldots, \xi_n) \in K_n$, очевидно, имеет место равенство

$$x=\xi_1(1, 0, \ldots, 0)+\xi_2(0, 1, \ldots, 0)+\ldots+\xi_n(0, 0, \ldots, 1),$$

которое и доказывает в соединении с уже известной линейной независимостью векторов $e_1, e_2, \ldots, e_n$, что эти векторы образуют базис в пространстве K_n.

В частности, оказывается, что числа $\xi_1, \xi_2, \ldots, \xi_n$ являются как раз координатами вектора x относительно базиса $e_1, e_2, \ldots, e_n$.

в. В пространстве $R(a, b)$ базиса — в том смысле, как он нами здесь определен, — не существует; доказательство этого утверждения будет дано в *2.36в*.

2.33. Основное значение базиса линейного пространства состоит в том, что линейные операции в пространстве, вначале заданные абстрактно, при задании базиса становятся обычными линейными операциями с числами — координатами взятых векторов относительно этого базиса. Именно, имеет место следующая теорема:

Теорема. *При сложении двух векторов пространства* К *их координаты (относительно любого базиса) складываются. При умножении вектора на число все его координаты умножаются на это число.*

Действительно, пусть

$$x=\xi_1 e_1+\xi_2 e_2+\ldots+\xi_n e_n, \qquad y=\eta_1 e_1+\eta_2 e_2+\ldots+\eta_n e_n.$$

Тогда в силу аксиом *2.12—2.13*

$$x+y=(\xi_1+\eta_1)e_1+(\xi_2+\eta_2)e_2+\ldots+(\xi_n+\eta_n)e_n,$$
$$\lambda x=\lambda\xi_1 e_1+\lambda\xi_2 e_2+\ldots+\lambda\xi_n e_n,$$

что и требуется.

См. задачи 6—7.

2.34. Если в линейном пространстве К можно найти n линейно независимых векторов, а всякие $n+1$ векторов этого пространства линейно зависимы, то число n называют *размерностью* пространства K; само же пространство K называют *n-мерным*. Мы будем в дальнейшем для n-мерного пространства над полем K использовать обозначение K_n (над полем R — соответственно R_n, над полем C — соответственно C_n). Линейное пространство, в котором можно указать сколь угодно большое число линейно независимых векторов, называется *бесконечномерным*.

Теорема. *В пространстве* K *размерности* n *существует базис из* n *векторов; более того, любая совокупность из* n *линейно независимых векторов пространства* K *является базисом этого пространства.*

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ — система из n линейно независимых векторов заданного n-мерного пространства K. Если x — некоторый вектор пространства K, то совокупность из $n+1$ векторов $x, e_1, e_2, \ldots, e_n$ линейно зависима; существует соотношение вида

$$\alpha_0 x + \alpha_1 e_1 + \ldots + \alpha_n e_n = 0, \tag{7}$$

причем среди коэффициентов $\alpha_0, \alpha_1, \ldots, \alpha_n$ имеются отличные от нуля. Можно утверждать, что коэффициент α_0 заведомо отличен от нуля: действительно, в противном случае мы получили бы линейную зависимость между векторами $e_1, e_2, \ldots, e_n$, которая, по предположению, не имеет места. Но в таком случае обычным путем, т. е. разделив уравнение на α_0 и перенеся все остальные члены вправо, мы получим, что x линейно выражается через векторы $e_1, e_2, \ldots, e_n$. Поскольку x — любой вектор пространства K, мы доказали, что векторы $e_1, e_2, \ldots, e_n$ образуют базис в этом пространстве, что и требовалось.

2.35. Следующая теорема является обратной по отношению к теореме *2.34*.

Теорема. *Если в пространстве* K *имеется базис, то размерность этого пространства равна числу базисных векторов.*

Доказательство. Пусть векторы $e_1, e_2, \ldots, e_n$ образуют базис пространства K. По самому определению базиса векторы $e_1, e_2, \ldots, e_n$ линейно независимы; следовательно, у нас уже имеется n линейно независимых векторов. Покажем, что всякие $n+1$ векторов пространства K линейно зависимы.

Пусть в пространстве К заданы $n+1$ векторов

$$\begin{aligned} x_1 &= \xi_1^{(1)}e_1 + \xi_2^{(1)}e_2 + \ldots + \xi_n^{(1)}e_n, \\ x_2 &= \xi_1^{(2)}e_1 + \xi_2^{(2)}e_2 + \ldots + \xi_n^{(2)}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{n+1} &= \xi_1^{(n+1)}e_1 + \xi_2^{(n+1)}e_2 + \ldots + \xi_n^{(n+1)}e_n. \end{aligned}$$

Выписывая в отдельный столбец координаты каждого из этих векторов, составим матрицу с n строками и $n+1$ столбцами

$$A = \left\| \begin{matrix} \xi_1^{(1)} & \xi_1^{(2)} & \ldots & \xi_1^{(n+1)} \\ \xi_2^{(1)} & \xi_2^{(2)} & \ldots & \xi_2^{(n+1)} \\ \cdot & \cdot & \cdot & \cdot \\ \xi_n^{(1)} & \xi_n^{(2)} & \ldots & \xi_n^{(n+1)} \end{matrix} \right\|.$$

Базисный минор (*1.92*) матрицы A имеет порядок $r \leqslant n$. Если $r=0$, то линейная зависимость очевидна. Пусть $r>0$. После указания r базисных столбцов мы сможем найти еще по меньшей мере один столбец, не попавший в число базисных. Но тогда согласно теореме о базисном миноре этот столбец является линейной комбинацией базисных столбцов. Соответствующий вектор пространства К является линейной комбинацией других векторов (из числа заданных $x_1, x_2, \ldots, x_{n+1}$). Но в таком случае векторы $x_1, x_2, \ldots$ $\ldots, x_{n+1}$, согласно *2.23б*, линейно зависимы, что и требовалось.

2.36. Примеры.

а. Пространство V_3 трехмерно, поскольку оно обладает базисом из трех векторов $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ (*2.32а*); соответственно V_2 двумерно, V_1 одномерно.

б. Пространство K_n n-мерно, поскольку оно обладает базисом из n векторов $e_1, e_2, \ldots, e_n$ (*2.32б*).

в. В пространствах $R(a, b)$ и $C(a, b)$ имеется сколь угодно большое число линейно независимых векторов (*2.22г*), и, следовательно, эти пространства бесконечномерны. Поэтому они не имеют базиса: наличие базиса привело бы к противоречию с теоремой *2.35*.

г. Всякое комплексное линейное пространство С является, очевидно, и вещественным, поскольку область комплексных чисел включает в себя область вещественных чисел. Однако размерность пространства C как комплексного пространства не совпадает с размерностью того же C как вещественного пространства: если векторы $e_1, \ldots, e_n$ линейно независимы в C как в комплексном пространстве, то в C как вещественном пространстве будут линейно независимы векторы $e_1, ie_1, \ldots, e_n, ie_n$, так что размерность C как вещественного пространства (если она конечна), вдвое больше, чем размерность C как комплексного пространства.

См. задачу 8.

§ 2.4. Подпространства

2.41. Допустим, что некоторая совокупность L элементов линейного пространства K обладает следующими свойствами:

а) если $x \in L$, $y \in L$, то $x+y \in L$;

б) если $x \in L$, λ — элемент поля K, то $\lambda x \in L$.

Таким образом, нам задана совокупность элементов, и в ней определены линейные операции. Покажем, что мы получаем здесь также линейное пространство. Для этого нужно проверить для совокупности L с операциями а)—б) выполнение аксиом *2.12—2.13*. Аксиомы 1), 2), 5)—8) удовлетворяются, поскольку они удовлетворяются вообще для всех элементов пространства K. Остается проверить аксиомы 3) и 4). Пусть x — любой элемент из L; тогда по условию $\lambda x \in L$ при любом вещественном λ.

Возьмем $\lambda = 0$; тогда, поскольку по теореме *2.14в* $0 \cdot x = 0$, нуль-вектор принадлежит совокупности L. Тем самым выполнена аксиома 3). Возьмем теперь $\lambda = -1$; поскольку по теореме *2.14г* $(-1)\,x$ есть элемент, противоположный элементу x, совокупность L вместе с каждым элементом x содержит и противоположный элемент. Таким образом, выполнена и аксиома 4), и наше утверждение полностью доказано.

Поэтому всякая совокупность $L \subset K$, удовлетворяющая условиям а) и б), называется *линейным подпространством* (или просто *подпространством*) пространства K.

2.42. Примеры.

а. Нуль-вектор пространства K образует, очевидно, наименьшее возможное подпространство пространства K.

б. Все пространство K — наибольшее возможное подпространство пространства K.

Эти два подпространства — нуль-вектор и все пространство — называют иногда *тривиальными* подпространствами; и тогда все остальные подпространства называют *нетривиальными*.

в. Пусть L_1 и L_2 — два подпространства одного и того же линейного пространства K. Совокупность всех векторов $x \in K$, принадлежащих одновременно к L_1 и L_2, образует подпространство, называемое *пересечением* подпространств L_1 и L_2. Совокупность всех векторов вида $y+z$, где $y \in L_1$, $z \in L_2$, образует подпространство, называемое *суммой* подпространств L_1 и L_2.

г. В пространстве V_3 все векторы, параллельные какой-либо плоскости (или какой-либо прямой), образуют подпространство. Если говорить не о векторах, а о точках *(2.17)*, то подпространствами пространства V_3 являются совокупности точек, расположенных на плоскости (или прямой), проходящей через начало координат.

д. В пространстве K_n рассмотрим совокупность L тех векторов $x=(\xi_1, \xi_2, \ldots, \xi_n)$, координаты которых удовлетворяют системе линейных уравнений

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n=0 \end{array}\right\} \tag{8}$$

с коэффициентами из поля K и со свободными членами, равными 0. Такая система называется *однородной* линейной системой. Однородная линейная система всегда совместна, так как имеет очевидное «нулевое» решение $x_1=x_2=\ldots=$ $=x_n=0$.

Пусть $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}$ и $c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$ — два решения такой системы; образуем числа

$$c_1=c_1^{(1)}+c_1^{(2)}, \quad c_2=c_2^{(1)}+c_2^{(2)}, \quad \ldots, \quad c_n=c_n^{(1)}+c_n^{(2)}.$$

Утверждается, что $c_1, c_2, \ldots, c_n$ *снова образуют решение* системы (8). Действительно, подставляя эти числа в i-е уравнение этой системы, мы получаем

$$\begin{aligned}a_{i1}c_1 + a_{i2}c_2 + \ldots + a_{in}c_n &= \\ = a_{i1}(c_1^{(1)} + c_1^{(2)}) + a_{i2}(c_2^{(1)} + c_2^{(2)}) + \ldots + a_{in}(c_n^{(1)} + c_n^{(2)}) &= \\ = (a_{i1}c_1^{(1)} + a_{i2}c_2^{(1)} + \ldots + a_{in}c_n^{(1)}) &+ \\ + (a_{i1}c_1^{(2)} + a_{i2}c_2^{(2)} + \ldots + a_{in}c_n^{(2)}) &= 0,\end{aligned}$$

что и требовалось.

Это решение мы будем называть *суммой* решений $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}$ и $c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$. Аналогично, если $c_1, c_2, \ldots, c_n$ — произвольное решение системы (8), то числа $\lambda c_1, \lambda c_2, \ldots, \lambda c_n$ при фиксированном $\lambda \in K$ образуют также решение системы (8); это решение мы будем называть *произведением решения $c_1, c_2, \ldots, c_n$ на число λ*.

Таким образом, *решения однородной системы с коэффициентами из поля K можно складывать друг с другом и умножать на числа из того же поля K*.

Тем самым совокупность L есть подпространство пространства K_n и, следовательно, является линейным пространством. Мы будем называть его *пространством решений системы* (8). В *3.41* мы вычислим размерность этого пространства и построим его базис.

См. задачу 9.

2.43. Отметим некоторые свойства подпространств, связанные с определениями §§ 2.2—2.3.

Прежде всего, всякое линейное соотношение, связывающее векторы $x, y, \ldots, z$ в подпространстве L, справедливо и во всем пространстве K, и обратно; в частности, факт линейной зависимости векторов $x, y, \ldots, z \in \mathrm{L}$ выполняется одновременно в подпространстве L и в пространстве K. Если, например, в пространстве K всякие $n+1$ векторов линейно зависимы, то это утверждение и подавно будет выполнено в подпространстве L. Отсюда вытекает, что *размерность любого подпространства L в n-мерном пространстве K не превосходит числа n*. В этом случае согласно теореме *2.34* *в каждом подпространстве $\mathrm{L} \subset \mathrm{K}$ можно построить базис из такого числа векторов, какова размерность подпространства L.*

Если в пространстве K выбран базис $e_1, e_2, \ldots, e_n$, то в общем случае, разумеется, нельзя выбрать базисные векторы подпространства L прямо из числа векторов $e_1, e_2, \ldots, e_n$ хотя бы потому, что в подпространство L может не входить ни один из них. Но можно утверждать, что *если выбран базис $f_1, f_2, \ldots, f_l$ в подпространстве* L (*имеющем для определенности размерность $l < n$*), *то всегда можно дополнительно выбрать векторы $f_{l+1}, \ldots, f_n$ во всем пространстве* K *так, что система $f_1, f_2, \ldots, f_l, \ldots, f_n$ будет базисом во всем* K.

Для доказательства будем рассуждать так. В пространстве K существуют векторы, которые линейно не выражаются через $f_1, f_2, \ldots, f_l$; действительно, если бы таких векторов не было, то векторы $f_1, f_2, \ldots, f_l$ — по условию линейно независимые — составляли бы базис пространства K и по теореме *2.35* размерность K была бы равна l, а не n. Обозначим через f_{l+1} любой из векторов, не выражающихся линейно через $f_1, f_2, \ldots, f_l$. Система $f_1, f_2, \ldots, f_l, f_{l+1}$ линейно независима; в самом деле, если бы существовало соотношение вида

$$\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_l f_l + \alpha_{l+1} f_{l+1} = 0,$$

то при $\alpha_{l+1} \neq 0$ мы получили бы, что вектор f_{l+1} можно линейно выразить через $f_1, f_2, \ldots, f_l$, а при $\alpha_{l+1} = 0$ получили бы, что векторы $f_1, f_2, \ldots, f_l$ линейно зависимы; оба эти вывода противоречат построению. Если теперь всякий вектор пространства K линейно выражается через $f_1, f_2, \ldots, f_l, f_{l+1}$, то система $f_1, f_2, \ldots, f_l, f_{l+1}$ образует базис в K (и $l + 1 = n$), и наше построение закончено. Если $l + 1 < n$, то имеется вектор f_{l+2}, не выражающийся линейно через $f_1, f_2, \ldots, f_l, f_{l+1}$. Таким образом можно продолжить построение; в конечном счете (через $n - l$ шагов) мы получим базис пространства K.

См. задачу 10.

2.44. Будем говорить, что векторы $g_1, \ldots, g_k$ *линейно независимы над подпространством* $L \subset K$, если из соотношения

$$\alpha_1 g_1 + \ldots + \alpha_k g_k \in L, \quad \alpha_1, \ldots, \alpha_k \in K,$$

следует $\alpha_1 = \ldots = \alpha_k = 0$. Если L — нулевое подпространство, то линейная независимость над L означает обычную линейную независимость. Линейная зависимость векторов $g_1, \ldots, g_k$ над подпространством L означает, что существует линейная комбинация $\alpha_1 g_1 + \ldots + \alpha_k g_k$, лежащая в L, причем среди коэффициентов $\alpha_1, \ldots, \alpha_k$ имеются отличные от 0.

Наибольшее возможное число векторов пространства K, линейно независимых над подпространством $\mathrm{L} \subset \mathrm{K}$, называется *размерностью* K *над* L.

Если векторы $g_1, \ldots, g_k$ линейно независимы над пространством $\mathrm{L} \subset \mathrm{K}$, а $f_1, \ldots, f_l$ — векторы, линейно независимые в подпространстве L, то векторы $g_1, \ldots, g_k, f_1, \ldots, f_l$ линейно независимы в пространстве K. Действительно, если бы имело место равенство

$$\alpha_1 f_1 + \ldots + \alpha_l f_l + \beta_1 g_1 + \ldots + \beta_k g_k = 0,$$

то, написанное в форме

$$\beta_1 g_1 + \ldots + \beta_k g_k = -(\alpha_1 f_1 + \ldots + \alpha_l f_l) \in \mathrm{L},$$

оно привело бы к заключению, что $\beta_1 = \ldots = \beta_k = 0$ в силу предположенной линейной независимости векторов $g_1, \ldots, g_k$ над L; отсюда $\alpha_1 = \ldots = \alpha_l = 0$ в силу линейной независимости векторов $f_1, \ldots, f_l$.

Векторы $f_{l+1}, \ldots, f_n$, *построенные в 2.43, линейно независимы над подпространством* L: действительно, если бы имело место равенство

$$\alpha_{l+1} f_{l+1} + \ldots + \alpha_n f_n = \alpha_1 f_1 + \ldots + \alpha_l f_l,$$

причем среди чисел $\alpha_{l+1}, \ldots, \alpha_n$ были бы отличные от 0, то векторы $f_1, \ldots, f_n$ оказались бы линейно зависимыми в противоречии с построением. Размерность пространства K над L, следовательно, не меньше, чем $n-l$. С другой стороны, она не может быть и больше, чем $n-l$, так как если бы нашлось $n-l+1$ векторов, например, $h_1, \ldots, h_{n-l+1}$, линейно независимых над L, то в пространстве K были бы линейно независимы векторы $h_1, \ldots, h_{n-l+1}, f_1, \ldots, f_l$, число которых больше, чем n. Таким образом, размерность K над L в данном случае равна $n-l$.

2.45. Прямая сумма. Говорят, что пространство L является *прямой суммой* своих подпространств $\mathrm{L}_1, \ldots, \mathrm{L}_m$, если:

а. Для всякого $x \in \mathrm{L}$ существует разложение

$$x = x_1 + \ldots + x_m, \quad x_1 \in \mathrm{L}_1, \ \ldots, \ x_m \in \mathrm{L}_m.$$

б. Это разложение единственно: если

$$x = x_1 + \ldots + x_m = y_1 + \ldots + y_m,$$

где $x_j \in L_j$, $y_j \in \mathrm{L}_j$, $j = 1, \ldots, m$, то $x_1 = y_1, \ \ldots, \ x_m = y_m$.

Впрочем, справедливость последнего условия следует из более простого требования:

б′. Если имеется разложение

$$0 = z_1 + \ldots + z_m, \quad \text{где} \quad z_1 \in \mathrm{L}_1, \ \ldots, \ z_m \in \mathrm{L}_m,$$

то $z_1 = \ldots = z_m = 0$.

Действительно, пусть выполнено условие *б′* и дано разложение *б*. Вычитая, находим

$$0 = (x_1 - y_1) + \ldots + (x_m - y_m);$$

применяя *б′*, получаем $x_1 = y_1, \ \ldots, \ x_m = y_m$. Обратно, *б′* следует из *б*, если положить $x = 0$, $x_1 = \ldots = x_m = 0$.

Из *б* следует, что всякие два из подпространств $\mathrm{L}_1, \ldots, \mathrm{L}_m$ имеют общим один лишь элемент 0. Действительно, если бы мы имели $z \in \mathrm{L}_k$ и $z \in \mathrm{L}_j$, то из сравнения двух разложений

$$z = z + 0, \quad z \in \mathrm{L}_j, \quad 0 \in \mathrm{L}_k,$$

$$z = 0 + z, \quad 0 \in \mathrm{L}_j, \quad z \in \mathrm{L}_k,$$

и условия *б* вытекало бы, что $z = 0$.

Так, n-мерное пространство K_n есть прямая сумма n одномерных подпространств, определенных любыми n линейно независимыми векторами. Кроме того, пространство K_n можно представить разными способами в форме прямой суммы и неодномерных подпространств.

2.46. Пусть в n-мерном пространстве K_n фиксировано подпространство L. Покажем, что *всегда существует подпространство* $\mathrm{M} \subset \mathrm{K}_n$, *которое в прямой сумме с* L *дает все* K_n. Для доказательства используем векторы $f_{l+1}, \ \ldots, \ f_n$, построенные в *2.43*, линейно независимые над подпространством L. Пусть M — подпространство, составленное из всех линейных комбинаций векторов $f_{l+1}, \ldots, f_n$; покажем, что оно удовлетворяет нашему требованию. Действительно, поскольку векторы $f_1, \ldots, f_n$ образуют базис в K_n (*2.43*), каждый вектор $x \in \mathrm{L}$ допускает разложение

$$x = \alpha_1 f_1 + \ldots + \alpha_l f_l + \alpha_{l+1} f_{l+1} + \ldots + \alpha_n f_n = y + z,$$

где $y = \alpha_1 f_1 + \ldots + \alpha_l f_l \in \mathrm{L}$, $z = \alpha_{l+1} f_{l+1} + \ldots + \alpha_n f_n \in \mathrm{M}$. При этом из $x = 0$ следует $\alpha_j = 0$ $(j = 1, \ldots, n)$ в силу линейной независимости векторов $f_1, \ldots, f_n$. Следовательно, условия *2.45а,б′* выполнены, и K_n есть прямая сумма L и M.

2.47. а. Если размерность пространства L_k равна r_k $(k = 1, \ldots, m)$ и в каждом пространстве L_k выделены r_k линейно независимых векторов $f_{k1}, \ldots f_{kr_k}$, то каждый вектор x суммы $\mathrm{L} = \mathrm{L}_1 + \ldots + \mathrm{L}_k$ можно линейно выразить через все эти векторы. Следовательно, *размерность суммы пространств* $\mathrm{L}_1, \ldots, \mathrm{L}_k$ *не больше суммы их размерностей. Если сумма* $\mathrm{L}_1 + \ldots + \mathrm{L}_k$ *прямая, то* все векторы $f_{k1}, \ldots, f_{kr_k}$ $(k = 1, \ldots, m)$ линейно независимы, так что в этом случае *размерность суммы равна сумме размерностей.*

б. В общем случае размерность суммы определяется через размерности слагаемых более сложным образом; мы рассмотрим только вопрос о размерности суммы *двух* конечномерных подпространств P и Q пространства K.

Пусть p и q обозначают размерности этих подпространств. Обозначим через L пересечение подпространств P и Q и через l его размерность. Выберем в L базис $e_1, e_2, \ldots, e_l$ и, используя соображения *2.43,* дополним его векторами $f_{l+1}, f_{l+2}, \ldots, f_p$ до полного базиса подпространства P и векторами $g_{l+1}, g_{l+2}, \ldots, g_q$ до полного базиса подпространства Q. Каждый вектор суммы $\mathrm{P} + \mathrm{Q}$ по самому определению есть сумма вектора из P и вектора из Q и поэтому может быть линейно выражен через векторы $e_1, e_2, \ldots, e_l, f_{l+1}, \ldots, f_p, g_{l+1}, \ldots, g_q$. Покажем, что *эти векторы образуют базис подпространства* $\mathrm{P} + \mathrm{Q}$. Для этого нам остается проверить их линейную независимость. Допустим, что существует линейное соотношение вида

$$\alpha_1 e_1 + \ldots + \alpha_l e_l + \beta_{l+1} f_{l+1} + \ldots + \beta_p f_p + \\ + \gamma_{l+1} g_{l+1} + \ldots + \gamma_q g_q = 0, \quad (9)$$

причем среди коэффициентов $\alpha_1, \ldots, \gamma_q$ имеются отличные от нуля. Мы можем тогда утверждать, что имеются отличные от нуля числа в совокупности $\gamma_{l+1}, \ldots, \gamma_q$, так как в противном случае векторы $e_1, e_2, \ldots, e_l, f_{l+1}, \ldots, f_p$ оказались бы линейно зависимыми, что невозможно ввиду того, что они образуют базис подпространства P. Следовательно, вектор

$$x = \gamma_{l+1} g_{l+1} + \ldots + \gamma_q g_q \neq 0, \tag{10}$$

иначе векторы $g_{l+1}, \ldots, g_q$ оказались бы линейно зависимыми. Но из (9) вытекает, что

$$-x = \alpha_1 e_1 + \ldots + \beta_p f_p \in \mathrm{P},$$

в то время как (10) показывает, что $x \in \mathrm{Q}$. Таким образом, x принадлежит и P и Q и, следовательно, входит в подпространство L. Но тогда

$$x = \gamma_{l+1} g_{l+1} + \ldots + \gamma_q g_q = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_l e_l.$$

Поскольку векторы $e_1, e_2, \ldots, e_l, g_{l+1}, \ldots, g_q$ линейно независимы, $\gamma_{l+1} = \ldots = \gamma_q = 0$. Полученное противоречие показывает, что векторы $e_1, e_2, \ldots, e_l, f_{l+1}, \ldots, f_p, g_{l+1}, \ldots, g_q$ действительно линейно независимы.

Согласно теореме *2.35* размерность подпространства $\mathrm{P} + \mathrm{Q}$ равна числу базисных векторов $e_1, \ldots, e_l, f_{l+1}, \ldots, f_p, g_{l+1}, \ldots, g_q$; но это число равно $p + q - l$. Итак, *размерность суммы двух подпространств равна сумме их размерностей за вычетом размерности пересечения.*

в. Следствие. *Если в n-мерном пространстве R_n выделены два подпространства R_p и R_q, размерности которых p и q в сумме превышают число n, то пересечение R_p и R_q имеет размерность не меньшую, чем $p + q - n$.*

2.48. Фактор-пространства.

а. Пусть в линейном пространстве K выделено подпространство L. Назовем элемент $x \in \mathrm{K}$ *сравнимым с элементом* $y \in \mathrm{K}$ (точнее, *сравнимым относительно* L), если $x - y \in \mathrm{L}$. Очевидно, в этом случае и y сравним с x, так что отношение сравнения симметрично. Всякий $x \in \mathrm{K}$ сравним сам с собою. Далее, если x сравним с y, а y сравним с z, то и x сравним с z, так как

$$x - z = (x - y) + (y - z) \in \mathrm{L}.$$

б. Совокупность всех элементов $y \in \mathrm{K}$, сравнимых с данным элементом $x \in \mathrm{K}$, называется *классом* и обозначается через X. В силу сказанного класс X содержит сам элемент x, и всякие два элемента $y \in \mathrm{X}$, $z \in \mathrm{X}$ сравнимы друг с другом. Наконец, если $u \notin \mathrm{X}$, то u не сравним ни с одним элементом из класса X. Поэтому два класса X и Y или не имеют общих элементов, или полностью совпадают. Одним из классов является все подпространство L; поскольку оно содержит нулевой элемент пространства K, этот класс обозначается через 0.

в. Все пространство K разбивается в совокупность непересекающихся классов X, Y, ... Эту совокупность классов мы обозначим через K/L. Введем в множестве K/L линейные операции следующим образом. Пусть X и Y — классы, α и β — элементы поля K; мы желаем определить класс $\alpha\mathrm{X}+\beta\mathrm{Y}$. Для этого выберем произвольно элементы $x \in \mathrm{X}$ и $y \in \mathrm{Y}$ и найдем класс Z, который содержит элемент $z=\alpha x+\beta y$; этот класс и обозначим $\alpha\mathrm{X}+\beta\mathrm{Y}$. Проверим, что он определен однозначно. Если в классе X мы возьмем элемент x_1, а в классе Y — элемент y_1, то

$$(\alpha x_1+\beta y_1)-(\alpha x+\beta y)=\alpha(x_1-x)+\beta(y_1-y)$$

лежит в подпространстве L вместе с x_1-x и y_1-y; это означает, что $\alpha x_1+\beta y_1$ лежит в том же классе, что и $\alpha x+\beta y$.

В частности, мы определили сложение классов X, Y и умножение их на числа $\alpha \in K$. Покажем, что эти операции удовлетворяют аксиомам линейного пространства *2.12—2.13*. Действительно, из справедливости *2.12* 1)—2) и *2.13* 5)—8) для элементов пространства K следует сразу выполнение этих же свойств для классов. Нулем пространства K/L является класс 0 (состоящий из всех элементов пространства L). Противоположным к классу X является класс, состоящий из элементов, противоположных к элементам класса X. Таким образом, для совокупности классов выполнены и аксиомы *2.12* 3)—4).

Построенное здесь линейное пространство K/L называется *фактор-пространством* пространства K по подпространству L.

2.49. Теорема. *Пусть* $\mathrm{K}=\mathrm{K}_n$ *есть* n*-мерное линейное пространство над полем* K, $\mathrm{L}=\mathrm{L}_l\subset\mathrm{K}$ *есть* l*-мерное подпро-*

странство в K. *Тогда фактор-пространство* K/L *имеет размерность* $n-l$.

Доказательство. Выберем произвольно базис $f_1, \ldots, f_l$ в подпространстве L и дополним его, как в *2.43*, векторами $f_{l+1}, \ldots, f_n$ до базиса всего K. Рассмотрим классы $X_{l+1} \ni f_{l+1}, \ldots, X_n \ni f_n$ и покажем, что они образуют базис в пространстве K/L. Для любого $x \in \mathrm{K}$ существует представление

$$x = \sum_{k=1}^{n} \alpha_k f_k,$$

поэтому для класса $X \ni x$ существует представление

$$X = \sum_{k=l+1}^{n} \alpha_k X_k.$$

Покажем, что классы $X_{l+1}, \ldots, X_n$ линейно независимы. Если бы мы имели при некоторых $\alpha_{l+1}, \ldots, \alpha_n$ из K

$$\alpha_{l+1} X_{l+1} + \ldots + \alpha_n X_n = 0 \in \mathrm{K}/\mathrm{L},$$

то, в частности, выполнялось бы соотношение

$$\alpha_{l+1} f_{l+1} + \ldots + \alpha_n f_n \in \mathrm{L};$$

но так как $f_{l+1}, \ldots, f_n$ линейно независимы над L *(2.44)*, то $\alpha_{l+1} = \ldots = \alpha_n = 0$, что и требовалось. Таким образом, $X_{l+1}, \ldots, X_n$ образуют базис в K/L; но тогда их число $n-l$ есть размерность пространства K/L *(2.35)*. Теорема доказана.

§ 2.5. Линейные оболочки

2.51. Важным способом построения подпространств является образование линейной оболочки заданной системы векторов.

Пусть $x, y, z, \ldots$ — некоторая система векторов линейного пространства K; *линейной оболочкой* системы $x, y, z, \ldots$ называется совокупность всех (конечных) линейных комбинаций

$$\alpha x + \beta y + \gamma z + \ldots \tag{11}$$

с коэффициентами $\alpha, \beta, \gamma, \ldots$ из поля K. Легко проверить, что для этой совокупности выполнены условия *2.41 а, б*; поэтому линейная оболочка системы $x, y, z, \ldots$ есть под пространство пространства K. Это подпространство, очевидно, содержит векторы $x, y, z, \ldots$ С другой стороны, всякое подпространство, содержащее векторы $x, y, z, \ldots$, содержит и все их линейные комбинации (11); следовательно, *линейная оболочка векторов $x, y, z, \ldots$ есть наименьшее подпространство, содержащее эти векторы*. Линейная оболочка векторов $x, y, z, \ldots$ обозначается через $\mathrm{L}(x, y, z, \ldots)$.

2.52. Примеры.

а. Линейная оболочка векторов $e_1, e_2, \ldots, e_n$, образующих базис некоторого пространства K, очевидно, совпадает со всем пространством K.

б. Линейная оболочка пары (неколлинеарных) векторов пространства V_3 состоит из всех векторов, параллельных плоскости этих векторов.

в. Линейная оболочка системы функций $1, t, t^2, \ldots, t^k$ пространства $K(a, b)$ (K есть R или C) совпадает с совокупностью всех многочленов от t не выше k-й степени. Линейная оболочка бесконечной системы функций $1, t, t^2, \ldots$ состоит из всех многочленов (любой степени) от переменного t с коэффициентами из поля K.

2.53. Отметим два простых свойства линейных оболочек.

а. Лемма. *Если векторы $x', y', \ldots$ принадлежат к линейной оболочке векторов $x, y, \ldots$, то вся линейная оболочка $\mathrm{L}(x', y', \ldots)$ содержится в линейной оболочке $\mathrm{L}(x, y, \ldots)$.*

Действительно, поскольку векторы $x', y', \ldots$ принадлежат к подпространству $\mathrm{L}(x, y, \ldots)$, все их линейные комбинации, совокупность которых и составляет линейную оболочку $\mathrm{L}(x', y', \ldots)$, также принадлежат к подпространству $\mathrm{L}(x, y, \ldots)$.

б. Лемма. *Всякий вектор системы $x, y, \ldots$, линейно зависящий от остальных векторов этой системы, можно из нее исключить без изменения линейной оболочки.*

Действительно, если, например, вектор x линейно зависит от векторов $y, z, \ldots$, то это означает, что $x \in \mathrm{L}(y, z, \ldots)$. Отсюда и из леммы *а* вытекает, что $\mathrm{L}(x, y, z, \ldots) \subset \subset \mathrm{L}(y, z, \ldots)$. С другой стороны, очевидно, $\mathrm{L}(y, z, \ldots) \subset \subset \mathrm{L}(x, y, z, \ldots)$. Оба включения вместе показывают, что $\mathrm{L}(y, z, \ldots) = \mathrm{L}(x, y, z, \ldots)$, что и требуется.

2.54. Поставим вопрос о *построении базиса линейной оболочки и определении ее размерности.* При решении этого вопроса мы будем предполагать, что число векторов $x, y, \ldots$, порождающих линейную оболочку $\mathrm{L}(x, y, \ldots)$, конечно (хотя некоторые из заключений и не будут существенно требовать такого предположения).

Допустим, что среди векторов $x, y, \ldots$, порождающих линейную оболочку $\mathrm{L}(x, y, \ldots)$, мы смогли найти r линейно независимых векторов,— обозначим их через $x_1, x_2, \ldots, x_r$,— через которые может быть линейно выражен любой вектор из системы $x, y, \ldots$ В этом случае мы можем утверждать, что *векторы $x_1, x_2, \ldots, x_r$ образуют базис пространства* $\mathrm{L}(x, y, \ldots)$. Действительно, всякий вектор $z \subset \mathrm{L}(x, y, \ldots)$ по самому определению линейной оболочки линейно выражается через конечное число векторов из системы $x, y, \ldots$; но каждый из векторов этой системы по условию линейно выражается через $x_1, x_2, \ldots, x_r$; поэтому в конечном счете и вектор z может быть линейно выражен непосредственно через векторы $x_1, x_2, \ldots, x_r$. В соединении с предположенной линейной независимостью векторов $x_1, x_2, \ldots, x_r$ мы получаем выполнение для них обоих условий, содержащихся в определении базиса *(2.31)*, что и требовалось.

Согласно теореме *2.35* размерность пространства $\mathrm{L}(x, y, \ldots)$ равна числу r. Поскольку в r-мерном пространстве не может существовать более чем r линейно независимых векторов, мы можем сделать относительно размерности r пространства $\mathrm{L}(x, y, \ldots)$ следующие выводы:

а. *Если число порождающих векторов $x, y, \ldots$ больше числа r, то векторы $x, y, \ldots$ линейно зависимы; если их число равно числу r, то они линейно независимы.*

б. *Всякие $r+1$ векторов из системы $x, y, \ldots$ линейно зависимы.*

в. *Размерность пространства $\mathrm{L}(x, y, \ldots)$ можно определить как максимальное число линейно независимых векторов в системе $x, y, \ldots$*

§ 2.6. Гиперплоскости

2.61. Как мы уже видели в *2.42г*, геометрический образ, соответствующий понятию подпространства, для пространства V_3 в «точечной» (а не «векторной») интерпретации есть плоскость (или прямая), проходящая через начало координат. Но плоскости и прямые, не проходящие через начало координат, желательно было бы также включить в круг наших рассмотрений. Замечая, что такие плоскости и прямые получаются из плоскостей и прямых, проходящих через начало координат, параллельным перемещением в пространстве, т. е. сдвигом, мы естественно приходим к следующему общему построению.

Пусть L — некоторое подпространство линейного пространства K и x_0 — фиксированный вектор, вообще говоря, не принадлежащий L. Рассмотрим совокупность H всех векторов x, которые получаются по формуле

$$x = x_0 + y,$$

когда вектор y пробегает все подпространство L. Совокупность H называется *результатом сдвига подпространства* L *вдоль вектора* x_0 или *гиперплоскостью*.

Заметим, что гиперплоскость сама, вообще говоря, не образует подпространства.

2.62. Примеры.

а. В пространстве V_3 совокупность всех векторов, выходящих из начала координат и кончающихся на некоторой плоскости γ, образует гиперплоскость. Легко проверить, что эта гиперплоскость является подпространством в том и только в том случае, когда плоскость γ проходит через начало координат.

б. В пространстве K_n рассмотрим совокупность H тех векторов $x = (\xi_1, \xi_2, \ldots, \xi_n)$, координаты которых удовлетворяют совместной системе линейных уравнений

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n = b_k, \end{array}\right\} \qquad (12)$$

и совокупность L тех векторов $y = (\eta_1, \eta_2, \ldots, \eta_n)$, координаты которых удовлетворяют однородной системе линейных уравнений с теми же коэффициентами:

$$\left.\begin{array}{l} a_{11}y_1+a_{12}y_2+\ldots+a_{1n}y_n=0, \\ a_{21}y_1+a_{22}y_2+\ldots+a_{2n}y_n=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{k1}y_1+a_{k2}y_2+\ldots+a_{kn}y_n=0. \end{array}\right\} \qquad (13)$$

Как мы уже знаем, совокупность L есть подпространство пространства K_n. Пусть $x_0=(\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_n^{(0)})$ — некоторое решение системы (12). Покажем, что совокупность H совпадает с совокупностью всех сумм x_0+y, где y пробегает все подпространство L. Действительно, если $y=(\eta_1, \eta_2, \ldots, \eta_n)$ есть некоторое решение системы (13), то вектор

$$x=x_0+y=(\xi_1^{(0)}+\eta_1, \xi_2^{(0)}+\eta_2, \ldots, \xi_n^{(0)}+\eta_n),$$

очевидно, является решением системы (12), т. е. входит в совокупность H. Обратно, если x — любой вектор совокупности H, то разность $y=x-x_0$ удовлетворяет уже системе (13), т. е. вектор y входит в подпространство L. В силу данного нами выше определения совокупность H есть гиперплоскость, именно, результат сдвига подпространства L на вектор x_0.

2.63. Каждой гиперплоскости, хотя она и не представляет собой подпространства, можно приписать некоторую *размерность*. Именно, размерность гиперплоскости H будем считать равной размерности того подпространства L, сдвигом которого была получена эта гиперплоскость.

Чтобы это определение было корректным, нужно показать, что данная гиперплоскость H может быть получена сдвигом т о л ь к о одного подпространства. Для доказательства этого утверждения допустим, что гиперплоскость H есть результат сдвига подпространства L на вектор x_0 и в то же время результат сдвига подпространства L' на вектор x_0'.

Тогда для любого $z\in\mathrm{H}$ имеем $z=x_0+y$, где $y\in\mathrm{L}$, и в то же время $z=x_0'+y'$, где $y'\in\mathrm{L}'$. Отсюда следует, что L' есть совокупность векторов, определяемых формулой $y'=(x_0-x_0')+y$, где y — произвольный вектор из L, т. е. что подпространство L' есть результат сдвига подпространства L на вектор $x_1=x_0-x_0'$. Покажем, что вектор x_1 входит в подпространство L. Нуль-вектор, как и всякий элемент подпространства L′, можно представить в виде x_1+y_1, где $y_1 \in$ L (поскольку L′ есть сдвиг подпространства L на вектор x_1); отсюда следует, что $x_1=-y_1$ и поэтому x_1 входит в L, что мы и утверждали. Но в таком случае и всякий вектор $y' \in$ L′ входит в подпространство L, так как представляется суммой вектора $x_1 \in$ L и некоторого вектора $y \in$ L. Следовательно, имеет место включение L′ ⊂ L. В силу полной симметрии условия аналогично можно доказать, что L ⊂ L′, откуда вытекает, что L ≡ L′, что и утверждалось.

В дальнейшем гиперплоскости размерности 1 будем называть *прямыми линиями*, гиперплоскости размерности 2 — *плоскостями*.

См. задачи 11—14.

§ 2.7. Морфизмы линейных пространств

2.71. Пусть каждому вектору x' линейного пространства K′ по некоторому правилу ω поставлен в соответствие вектор x'' линейного пространства K″. Правило ω называется *морфизмом* (или *линейным оператором*), если выполняются следующие соотношения:

а) $\omega(x'+y')=\omega(x')+\omega(y')$ для любых x', y' из K′;

б) $\omega(\alpha x')=\alpha\,\omega(x')$ для любых $x' \in$ K′ и любого $\alpha \in K$.

Если морфизм ω отображает пространство K′ на все пространство K″, он называется *эпиморфизмом*. Если морфизм ω отображает пространство K′ хотя бы и не на все K″, но взаимно однозначно (так что из $x' \neq y'$ следует $\omega(x') \neq \omega(y')$), он называется *мономорфизмом*. Если морфизм ω отображает пространство K′ взаимно однозначно на все пространство K″, т. е. является одновременно моно- и эпиморфизмом, он называется *изоморфизмом*, а сами пространства K′ и K″ называются изоморфными (точнее, *K-изоморфными*).

Общепринятое обозначение морфизма:

$$\omega\colon \mathrm{K}' \to \mathrm{K}''.$$

2.72. Примеры.

а. Пусть L есть подпространство пространства K. Отображение ω, которое каждому вектору $x \in$ L ставит в соответствие этот же вектор x в пространстве K, есть морфизм

L в K, и именно мономорфизм (причем, если $L \neq K$, не эпиморфизм). Этот морфизм ω называется *вложением* L в K.

б. Пусть L есть подпространство пространства K и K/L — фактор-пространство пространства K по подпространству L *(2.48)*. Отображение ω, которое каждому вектору $x \in K$ ставит в соответствие класс $X \subset K/L$, содержащий элемент x, есть морфизм K в K/L, и именно эпиморфизм (причем, если $L \neq 0$, не мономорфизм). Этот морфизм ω называется *каноническим отображением* K *на* K/L.

2.73. а. Пусть пространство K′ n-мерно и обладает базисом $e'_1, \ldots, e'_n$. В пространстве K″ выберем произвольно векторы $e''_1, \ldots, e''_n$. Поставим в соответствие любому вектору $x' = \sum_{k=1}^{n} \xi_k e'_k \in K'$ вектор

$$\omega(x') = x'' = \sum_{k=1}^{n} \xi_k e''_k$$

с теми же коэффициентами ξ_k $(k = 1, \ldots, n)$.

Покажем, что *отображение* $\omega(x') = x''$ *есть морфизм пространства* K′ *в пространство* K″.

Пусть в пространстве K′ выбраны два вектора

$$x' = \sum_{k=1}^{n} \xi_k e'_k \quad \text{и} \quad y' = \sum_{k=1}^{n} \eta_k e'_k;$$

тогда по *2.33*

$$x' + y' = \sum_{k=1}^{n} (\xi_k + \eta_k)\, e'_k.$$

Согласно определению отображения ω

$$\omega(x') = \sum_{k=1}^{n} \xi_k e''_k, \quad \omega(y') = \sum_{k=1}^{n} \eta_k e''_k.$$

Далее,

$$\omega(x' + y') = \sum_{k=1}^{n} (\xi_k + \eta_k)\, e''_k = \sum_{k=1}^{n} \xi_k e''_k + \sum_{k=1}^{n} \eta_k e''_k = \omega(x') + \omega(y'),$$

так что условие *2.71а*) выполнено. Аналогично при любом $\alpha \in K$

$$\omega(\alpha x') = \omega\left(\alpha \sum_{k=1}^{n} \xi_k e'_k\right) = \omega\left(\sum_{k=1}^{n} \alpha \xi_k e'_k\right) = \\ = \sum_{k=1}^{n} \alpha \xi_k e''_k = \alpha \sum_{k=1}^{n} \xi_k e''_k = \alpha\, \omega(x'),$$

так что выполнено и условие *2.71б*). Следовательно, $\omega(x') = x''$ есть морфизм K' в K'', что и утверждалось.

б. Выясним, при каком условии морфизм ω, описанный в *а*, является *эпиморфизмом*. Очевидно, необходимым и достаточным условием для этого является возможность представления в форме $\sum_{k=1}^{n} \xi_k e''_k$ любого вектора $x'' \in \mathrm{K}''$, иными словами, тот факт, что K'' *совпадает с линейной оболочкой векторов* $e''_1, \ldots, e''_n$.

в. Выясним, при каком условии морфизм ω, описанный в *а*, является *мономорфизмом*. Для этого необходимо и достаточно, чтобы векторы $\sum \xi_k e''_k$, $\sum \eta_k e''_k$, различающиеся хотя бы по одной паре координат (т. е. такие, что $\xi_k \neq \eta_k$ хотя бы при одном значении k), были бы различными векторами пространства K''. Но это равносильно линейной независимости векторов $e''_1, \ldots, e''_n$, следовательно, *морфизм ω тогда и только тогда является мономорфизмом, когда векторы $e''_1, \ldots, e''_n$ линейно независимы.*

г. Как следствие получаем: морфизм ω, описанный в *а*, является изоморфизмом тогда и только тогда, когда векторы $e''_1, \ldots, e''_n$ линейно независимы и их линейная оболочка совпадает со всем пространством K''. Иначе говоря, морфизм ω *является изоморфизмом тогда и только тогда, когда векторы $e''_1, \ldots, e''_n$ образуют базис пространства* K''.

2.74. Теорема. *Любые два n-мерных пространства K' и K'' (над одним и тем же полем K) K-изоморфны.*

Доказательство. Пусть $e'_1, \ldots, e'_n$ — базис в пространстве K' и $e''_1, \ldots, e''_n$ — базис в пространстве K''. С помощью этих систем векторов построим морфизм ω, как указано в *2.73а*. В силу *2.73г* он является изоморфизмом, что и требуется.

2.75. Следствие. Всякое n-мерное линейное пространство над полем K K-изоморфно пространству K_n *(2.15б)*. В частности, всякое n-мерное комплексное пространство C-изоморфно пространству C_n; всякое n-мерное вещественное пространство R-изоморфно пространству R_n.

2.76. Дальнейшие свойства моно- и эпиморфизмов.

а. Пусть имеется морфизм $\omega\colon \mathrm{K}' \to \mathrm{K}''$. Рассмотрим совокупность всех векторов $\omega(x') \in \mathrm{K}''$, когда x' пробегает все K'. Эта совокупность есть, очевидно, подпространство $\mathrm{L}'' \subset \mathrm{K}''$. Оно называется *областью значений* морфизма ω. Ясно, что отображение ω пространства K' в L'' является эпиморфизмом. Если морфизм $\omega\colon \mathrm{K}' \to \mathrm{K}''$ был мономорфизмом, то морфизм $\omega\colon \mathrm{K}' \to \mathrm{L}''$ есть изоморфизм.

б. Пусть имеется морфизм $\omega\colon \mathrm{K}' \to \mathrm{K}''$. Рассмотрим совокупность L' всех векторов $x' \in \mathrm{K}'$, для которых $\omega(x') = 0$. Совокупность L' есть, очевидно, подпространство в пространстве K'; оно называется *ядром*, или *нуль-многообразием* морфизма ω.

Построим фактор-пространство K'/L' *(2.48)*. Все элементы x', лежащие в одном и том же классе $\mathrm{X}' \in \mathrm{K}'/\mathrm{L}'$ переводятся морфизмом ω в один и тот же элемент пространства K''; действительно, для двух таких элементов x' и y' мы имеем $x' - y' = z' \in \mathrm{L}'$, откуда $\omega(x') - \omega(y') = \omega(z') = 0$, $\omega(x') = \omega(y')$. Поставим в соответствие каждому классу $\mathrm{X}' \in \mathrm{K}'/\mathrm{L}'$ элемент $x'' = \omega(x') \in \mathrm{K}''$, где $x' \in \mathrm{X}'$ — любой элемент; мы только что видели, что x'' определен при этом однозначно. Положим $x'' = \Omega(\mathrm{X}')$. Отображение Ω, как легко видеть, есть морфизм пространства K'/L' в K''; он является мономорфизмом, так как из $\mathrm{X}' \neq \mathrm{Y}'$, $x' \in \mathrm{X}'$, $y' \in \mathrm{Y}'$ следует

$$\Omega(\mathrm{X}') - \Omega(\mathrm{Y}') = \omega(x') - \omega(y') = \omega(x' - y') \neq 0.$$

Таким образом, любой морфизм $\omega\colon \mathrm{K}' \to \mathrm{K}''$ порождает мономорфизм $\Omega\colon \mathrm{K}'/\mathrm{L}' \to \mathrm{K}''$. Если морфизм ω был эпиморфизмом, то, очевидно, и мономорфизм Ω является эпиморфизмом, так что эпиморфизм $\omega\colon \mathrm{K}' \to \mathrm{K}''$ порождает изоморфизм $\Omega\colon \mathrm{K}'/\mathrm{L}' \to \mathrm{K}''$.

Мы продолжим изучение морфизмов в гл. 4.

ЗАДАЧИ

1. Образует ли совокупность векторов на плоскости, начала которых находятся в начале координат, а концы — в пределах первой четверти, линейное пространство (с обычными операциями)?

2. Образует ли линейное пространство совокупность всех векторов на плоскости с исключением векторов, параллельных некоторой заданной прямой?

3. Рассмотрим совокупность P одних положительных вещественных чисел. Введем операции по следующим правилам: под «сложением» двух чисел будем понимать их (обычное) умножение, а под произведением элемента $r \in P$ на вещественное число λ будем понимать (обычное) возведение числа r в степень λ. Является ли P с указанными операциями линейным пространством?

4. Показать, что в случае n заданных векторов в пространстве K_n критерием линейной независимости их служит неравенство нулю определителя, составленного из координат этих векторов.

5. Показать, что в пространстве $R(a, b)$, где $0 < a < b$, функции $t^{\alpha_1}, t^{\alpha_2}, \ldots, t^{\alpha_k}$ линейно независимы, если $\alpha_1, \alpha_2, \ldots, \alpha_k$ — различные вещественные числа.

6. Относительно системы векторов $e_1, e_2, \ldots, e_n$ линейного пространства K известно следующее:

а) Каждый вектор $x \in \mathrm{K}$ допускает разложение

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n.$$

б) Для некоторого фиксированного вектора $x_0 \in \mathrm{K}$ это разложение единственно.

Доказать, что система $e_1, \ldots, e_n$ образует базис пространства K.

7. Существует ли базис у пространства P (задача 3)?

8. Какова размерность пространства P (задача 3)?

9. В пространстве V_3 взяты два различных двумерных подпространства L_1 и L_2 (две различные плоскости, проходящие через начало координат). Что собой представляют их пересечение и сумма?

10. Доказать теорему: если размерность подпространства $\mathrm{L} \subset \mathrm{K}$ совпадает с размерностью пространства K, то $\mathrm{L} \equiv \mathrm{K}$.

11. Определен ли однозначно самой гиперплоскостью вектор сдвига x_0, фигурирующий в ее построении?

12. Показать, что всякая гиперплоскость $\mathrm{H} \subset \mathrm{K}$ обладает следующим свойством: если $x \in \mathrm{H}$, $y \in \mathrm{H}$, то $\alpha x + (1-\alpha) y \in \mathrm{H}$ при любом вещественном α. Обратно, если некоторое подмножество $\mathrm{H} \subset \mathrm{K}$ обладает сформулированным свойством, то H есть гиперплоскость. Какая геометрическая характеристика гиперплоскости содержится в этом свойстве?

13. Гиперплоскости H_1 и H_2 имеют размерности соответственно p и q. Какой (наименьшей) размерности нужно взять гиперплоскость H_3, чтобы она, будучи должным образом расположена в пространстве, содержала в себе и H_1 и H_2?

14. Аналогичная задача для трех гиперплоскостей $\mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3$ с размерностями p, q и r.

15. Согласно теореме *2.74* одномерные пространства R_1 и P (задача 3) изоморфны. Как можно осуществить этот изоморфизм?

СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

§ 3.1. Еще о ранге матрицы

3.11. Мы уже несколько раз встречались с матрицами в нашем изложении. В этом параграфе мы остановимся более детально на тех свойствах матрицы, которые связаны с понятием ранга (*1.92*). Это позволит нам дать общее решение задачи о системах линейных уравнений, сформулированной в § 1.2.

Напомним основные определения из § 1.9.

Пусть дана матрица с n строками и k столбцами, заполненными числами a_{ij} (i — индекс строки, j — индекс столбца, $i=1, 2, \ldots, n$, $j=1, 2, \ldots, k$) из поля K:

$$A = \left\| \begin{matrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ . & . & \ldots & . \\ a_{n1} & a_{n2} & \ldots & a_{nk} \end{matrix} \right\|. \qquad (1)$$

Если в этой матрице выбрать произвольным образом m строк и m столбцов, то элементы, стоящие на пересечениях этих строк и столбцов, образуют квадратную матрицу порядка m.

Определитель этой матрицы называется *минором m-го порядка матрицы A*. Натуральное число r называется *рангом* матрицы A, если у нее имеется минор порядка r, отличный от нуля, а все имеющиеся миноры порядка $r+1$ и выше равны нулю. Если матрица A имеет ранг $r>0$, то всякий ее минор r-го порядка, отличный от нуля, называется *базисным минором.* Столбцы и строки матрицы, на пересечениях которых находятся элементы базисного минора, называются *базисными столбцами и строками.*

Наши дальнейшие рассмотрения будут основаны на возможности придать любому столбцу из n чисел геометрический смысл вектора в n-мерном пространстве K_n (*2.15б*). Сама матрица A в этой геометрической трактовке отвечает определенной совокупности k векторов пространства K_n; обозначим через x_j вектор, соответствующий j-му столбцу матрицы A ($j = 1, 2, \ldots, k$). Любое линейное соотношение между столбцами матрицы мы можем истолковать как такое же линейное соотношение между соответствующими векторами (*2.22б*).

Образуем в пространстве K_n линейную оболочку векторов $x_1, x_2, \ldots, x_k$ (*2.51*). Покажем, что *векторы, отвечающие базисным столбцам матрицы* A, — пусть для определенности первые r ее столбцов являются базисными — *образуют базис этой линейной оболочки*. Для этого достаточно показать, во-первых, что векторы $x_1, x_2, \ldots, x_r$ линейно независимы и, во-вторых, что через них может быть линейно выражен любой из оставшихся векторов $x_{r+1}, \ldots, x_k$ (*2.54*). Проверим сначала первое из этих утверждений. Линейная зависимость векторов $x_1, x_2, \ldots, x_r$ была бы равносильна линейной зависимости r первых столбцов матрицы A. Но тогда в силу теоремы *1.96* любой определитель r-го порядка, построенный на этих столбцах и каких-нибудь r строках матрицы A, был бы равен нулю. В частности, был бы равен нулю базисный минор матрицы A, что противоречит его определению. Таким образом, первое утверждение доказано. Второе утверждение мы фактически доказали в *1.93*; сформулированное там для столбцов матрицы A, оно составило содержание «теоремы о базисном миноре». Тем самым мы полностью доказали, что векторы $x_1, x_2, \ldots, x_r$ образуют базис пространства $\mathrm{L}(x_1, x_2, \ldots, x_k)$. В силу теоремы *2.35* размерность этого пространства равна числу r, т. е. равна рангу матрицы A. Мы получили следующий важный результат:

Теорема. *Размерность линейной оболочки векторов, определенных столбцами матрицы A, равна рангу этой матрицы. Векторы, отвечающие базисным столбцам матрицы A, образуют базис линейной оболочки.*

3.12. Дальнейшие предложения представляют собой очевидные следствия из *2.54а — в*.

а. Теорема. *Если ранг матрицы A меньше, чем число ее столбцов ($r < k$), то столбцы матрицы A линейно зависимы.*

Если ранг матрицы A равен числу столбцов ($r=k$), *то столбцы матрицы A линейно независимы.*

б. Теорема. *Всякие $r+1$ столбцов матрицы A ранга r линейно зависимы.*

в. Теорема. *Ранг любой матрицы равен максимальному числу ее линейно независимых столбцов.*

Это последнее предложение имеет большое принципиальное значение, так как оно содержит в себе *новое определение ранга матрицы.*

3.13. Если транспонировать матрицу A, т. е. перейти от матрицы A к матрице A', строки которой являются столбцами матрицы A, то ранг транспонированной матрицы A' будет, очевидно, таким же, каков был ранг матрицы A. Но согласно теореме *3.12в* ранг матрицы A равен максимальному количеству ее линейно независимых столбцов, или, что то же самое, строк матрицы A'. Мы приходим к несколько неожиданному заключению:

Теорема. *Максимальное количество линейно независимых строк любой матрицы совпадает с максимальным количеством ее линейно независимых столбцов.*

Заметим, что эта теорема нетривиальна; любое прямое ее доказательство потребовало бы цепи рассуждений, эквивалентной доказательству теорем *1.93* и *3.11*.

3.14. Отметим еще следующий результат, который вытекает из теоремы *3.11* и леммы *2.53б:*

Теорема. *Если один из столбцов матрицы A является линейной комбинацией других ее столбцов, то его можно вычеркнуть из этой матрицы, не меняя ее ранга.*

См. задачи 1—2.

§ 3.2. Нетривиальная совместность однородной линейной системы

3.21. Пусть дана однородная линейная система

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n &= 0. \end{aligned}\right\} \qquad (2)$$

Как мы уже знаем, эта система всегда совместна, так как обладает нулевым решением $x_1 = x_2 = \ldots = x_n = 0$. Основная задача, с которой приходится встречаться при изучении однородных линейных систем, следующая: *при каких условиях однородная система «нетривиально совместна»*, т. е. *имеет, кроме нулевого, еще другие решения*? Результаты § 3.1 позволяют сразу решить эту задачу. Действительно, существование ненулевого решения системы (2), как мы видели в *2.22б*, равносильно линейной зависимости столбцов матрицы

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{Vmatrix},$$

а по теореме *3.12а* эта линейная зависимость имеет место тогда и только тогда, когда ранг матрицы A меньше числа столбцов. Итак, мы получаем следующую теорему:

Теорема. *Если ранг матрицы A равен числу n, система* (2) *не имеет ненулевых решений; если же ранг матрицы A меньше n, ненулевые решения системы* (2) *заведомо существуют; в этом и только в этом случае наша система нетривиально совместна.*

3.22. В частности, если число уравнений в системе (2) меньше числа неизвестных ($k < n$), то ранг матрицы A заведомо меньше n и ненулевые решения всегда существуют. Если $k = n$, то решение вопроса о существовании ненулевых решений зависит от величины $\det A$: если $\det A \neq 0$, ненулевых решений нет ($r = n$), если $\det A = 0$, ненулевые решения есть ($r < n$). При $k > n$ нужно рассмотреть все возможные определители n-го порядка, которые получаются при выборе n произвольных строк матрицы A; если все эти определители равны нулю, то $r < n$ и имеются ненулевые решения; если же среди этих определителей есть хотя бы один отличный от нуля, то $r = n$ и имеется только нулевое решение.

См. задачу 3.

§ 3.3. Условие совместности общей линейной системы

3.31. Пусть дана общая линейная система уравнений

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n=b_k. \end{array}\right\} \qquad (3)$$

С этой системой мы сопоставим две матрицы: матрицу

$$A=\left\| \begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{array} \right\|,$$

называемую *основной матрицей системы* (3), и матрицу

$$A_1=\left\| \begin{array}{ccccc} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kn} & b_k \end{array} \right\|,$$

называемую *расширенной матрицей системы* (3). Важная теорема о совместности системы (3) гласит:

Теорема (Кро́некер и Капе́лли). *Система* (3) *совместна тогда и только тогда, когда ранг расширенной матрицы этой системы равен рангу основной матрицы.*

Доказательство. Допустим сначала, что система (3) совместна; если $c_1, c_2, \ldots, c_n$ — некоторое ее решение, то имеют место равенства

$$\begin{array}{l} a_{11}c_1+a_{12}c_2+\ldots+a_{1n}c_n=b_1, \\ a_{21}c_1+a_{22}c_2+\ldots+a_{2n}c_n=b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{k1}c_1+a_{k2}c_2+\ldots+a_{kn}c_n=b_k. \end{array}$$

Но эти равенства означают, что последний столбец матрицы A_1 есть линейная комбинация остальных столбцов этой матрицы (с коэффициентами $c_1, c_2, \ldots, c_n$). В силу теоремы *3.14* последний столбец матрицы A_1 можно вычеркнуть без изменения ее ранга. Но при вычеркивании последнего столбца матрица A_1 переходит как раз в матрицу A. Следовательно, если система (3) совместна, матрицы A и A_1 имеют одинаковый ранг.

Допустим теперь, что матрицы A и A_1 имеют одинаковый ранг, и покажем, что система (3) совместна. Пусть r — ранг матрицы A (следовательно, и матрицы A_1). Рассмотрим r базисных столбцов матрицы A; они будут базисными столбцами и матрицы A_1. По теореме *1.93* последний столбец матрицы A_1 есть линейная комбинация базисных столбцов, а следовательно, и линейная комбинация всех столбцов матрицы A. Если мы коэффициенты этой последней линейной комбинации обозначим через $c_1, c_2, \ldots, c_n$, то получим, что выполняются равенства

$$\left.\begin{array}{l} a_{11}c_1+a_{12}c_2+\ldots+a_{1n}c_n=b_1, \\ a_{21}c_1+a_{22}c_2+\ldots+a_{2n}c_n=b_2, \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{k1}c_1+a_{k2}c_2+\ldots+a_{kn}c_n=b_k. \end{array}\right\} \qquad (4)$$

Таким образом, система (4) удовлетворяется значениями

$$x_1=c_1,\quad x_2=c_2,\quad \ldots,\quad x_n=c_n$$

и, следовательно, совместна. Теорема доказана.

§ 3.4. Общее решение линейной системы

3.41. Теорема Кронекера — Капелли, устанавливая общее условие совместности линейной системы, не дает способа получения решений этой системы. В этом параграфе мы выведем формулу, заключающую в себе общее решение линейной системы.

Общим решением системы (3) называется совокупность выражений вида

$$x_j=f_j(a_{11}, a_{12}, \ldots, a_{kn}, b_1, \ldots, b_k, q_1, \ldots, q_s) \qquad (j=1, \ldots, n),$$

где в правых частях стоят функции, зависящие от коэффициентов a_{ij} системы (3), ее свободных членов b_j и неопределенных параметров $q_1, \ldots, q_s$, такие, что при произвольно фиксированных значениях параметров q_j (из поля K) получающиеся величины $x_j=c_j$ $(j=1, \ldots, n)$ образуют решение системы (3), причем любое решение системы (3) получается этим путем при некотором выборе значений параметров $q_1, \ldots, q_s$ из K. В *2.62б* мы видели, что совокупность всех сумм вида x_0+y, где x_0 — какое-либо частное решение этой системы, а y пробегает совокупность всех решений соответствующей однородной системы, есть совокупность всех решений системы (3). Теперь этот факт можно выразить следующим образом: *общее решение неоднородной системы* (3) *есть сумма какого-либо частного решения этой же системы и общего решения соответствующей однородной системы.*

Пусть дана совместная линейная система (3) с основной матрицей $A=\|a_{ij}\|$ ранга r. Можно предположить, что базисный минор M матрицы A расположен в ее левом верхнем углу; если бы это было не так, мы смогли бы достичь желаемого расположения некоторой перестановкой строк и столбцов матрицы A, что равносильно некоторой перестановке уравнений и неизвестных в системе (3). Рассмотрим первые r уравнений системы (3), причем перепишем их следующим образом:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1r}x_r=b_1-a_{1,\,r+1}x_{r+1}-\ldots-a_{1n}x_n, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2r}x_r=b_2-a_{2,\,r+1}x_{r+1}-\ldots-a_{2n}x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1+a_{r2}x_2+\ldots+a_{rr}x_r=b_r-a_{r,\,r+1}x_{r+1}-\ldots-a_{rn}x_n. \end{array}\right\} \qquad (5)$$

Придадим неизвестным $x_{r+1}, \ldots, x_n$ совершенно произвольные значения $c_{r+1}, \ldots, c_n$. Тогда система (5) превращается в систему r уравнений для r неизвестных $x_1, x_2, \ldots, x_r$ с определителем M, отличным от нуля (базисным минором матрицы A). Эту систему можно разрешить по правилу Крамера *1.73*; существуют, следовательно, числа $c_1, c_2, \ldots, c_n$, которые при подстановке их в систему (5) на место неизвестных $x_1, x_2, \ldots, x_n$ обращают все уравнения этой системы в тождества. Покажем, что эти значения $c_1, c_2, \ldots, c_n$ удовлетворяют и всем прочим уравнениям системы (3).

Первые r строк расширенной матрицы A_1 системы (3) являются базисными строками этой матрицы, так как ранг расширенной матрицы в силу условия совместности равен r и в первых r строках матрицы A_1 содержится отличный от нуля минор M. В силу теоремы *1.93* (в применении к строкам) каждая из последующих строк матрицы A_1 есть линейная комбинация ее первых r строк. Для системы (3) это означает, что каждое уравнение системы (3), начиная с $(r+1)$-го, есть линейная комбинация первых r уравнений этой системы. Следовательно, если первые r уравнений системы (3) удовлетворяются значениями $x_1=c_1, \ldots, x_n=c_n$, то все остальные ее уравнения также удовлетворяются этими значениями.

3.42. Чтобы записать построенное решение системы (3) в виде некоторой формулы, обозначим через $M_j(\alpha_i)$ определитель, получающийся из базисного минора $M=\det\|a_{ij}\|$ $(i, j=1, 2, \ldots, r)$ заменой его j-го столбца на столбец из величин $\alpha_1, \alpha_2, \ldots, \alpha_i, \ldots, \alpha_r$. Тогда, записывая решение системы (5) с помощью формул Крамера, мы получим

$$\begin{gathered} x_j=\frac{1}{M} M_j(b_i-a_{i,r+1}c_{r+1}-\ldots-a_{in}c_n)= \\ =\frac{1}{M}\left[M_j(b_i)-c_{r+1}M_j(a_{i,r+1})-\ldots-c_nM_j(a_{in})\right] \qquad (6) \\ (j=1, \ldots, r), \\ x_j=c_j\,(j=r+1, \ldots, n). \end{gathered}$$

Эти формулы выражают значения неизвестных $x_j=c_j$ $(j=1, 2, \ldots, r)$ через коэффициенты системы, свободные члены и произвольные величины (параметры) $c_{r+1}, c_{r+2}, \ldots, c_n$.

Покажем, что в формулах (6) содержится любое решение системы (3). Действительно, пусть $c_1^0, c_2^0, \ldots, c_{r+1}^0, \ldots, c_n^0$ — произвольное решение системы (3). Очевидно, оно является также решением системы (5). Но по правилу Крамера из системы (5) величины $c_1^0, c_2^0, \ldots, c_r^0$ определяются через величины $c_{r+1}^0, \ldots, c_n^0$ однозначно и именно по формулам (6). Таким образом, при $c_{r+1}=c_{r+1}^0, \ldots, c_n=c_n^0$ формулы (6) дают нам как раз взятое решение $c_1^0, c_2^0, \ldots, c_n^0$, что и требовалось. Таким образом, формулы (6) дают *общее решение* системы (3).

См. задачи 4—7.

§ 3.5. Геометрические свойства совокупности решений линейной системы

3.51. Рассмотрим сначала случай однородной линейной системы (2); мы уже знаем, что совокупность всех решений такой системы образует линейное пространство (*2.42д*). Обозначая это пространство через L, вычислим его размерность и построим его базис.

Формулы (6) в данном случае приобретают вид

$$-Mc_j = c_{r+1}M_j(a_{i,r+1}) + \ldots + c_nM_j(a_{in}) \qquad (j = 1, 2, \ldots, r), \quad (7)$$

так как $M_j(b_i) = M_j(0) = 0$.

Каждому решению $(c_1, c_2, \ldots, c_r, c_{r+1}, \ldots, c_n)$ системы (2) поставим в соответствие вектор $(c_{r+1}, \ldots, c_n)$ пространства K_{n-r}. Так как числа $c_{r+1}, \ldots, c_n$ могут быть взяты произвольно и однозначно определяют решение системы (2), то соответствие между пространством решений системы (2) и пространством K_{n-r} получается взаимно однозначным. Поскольку при этом соответствии, как легко проверить, сохраняются линейные операции, это соответствие является изоморфным. Итак, *пространство L решений однородной системы линейных уравнений с n неизвестными и с рангом r матрицы коэффициентов изоморфно пространству K_{n-r}*. В частности, размерность пространства L равна числу $n-r$.

3.52. Любая система из $n-r$ линейно независимых решений однородной линейной системы уравнений, являющаяся вследствие теоремы *2.34* базисом в пространстве всех решений, называется *фундаментальной системой решений*. Для построения фундаментальной системы решений мы можем воспользоваться любым базисом пространства K_{n-r}; в силу изоморфизма соответствующие решения системы (2) будут образовывать базис и в пространстве всех решений этой системы.

Простейший базис пространства K_{n-r} образуется векторами $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, $\ldots$, $e_{n-r} = (0, 0, \ldots, 1)$ (*2.32б*). Чтобы получить решение системы (2), соответствующее, например, вектору e_1, нужно в формулах (7) подставить $c_{r+1} = 1$, $c_{r+2} = \ldots = c_n = 0$ и определить соответствующие значения

$$c_i = c_i^{(1)} \qquad (i = 1, 2, \ldots, n).$$

Аналогично строится решение, соответствующее любому другому вектору e_j $(j = 2, \ldots, n-r)$.

Построенная так совокупность решений системы (2) называется *нормальной фундаментальной системой решений*. Если обозначить эти решения через $x^{(1)}, x^{(2)}, \ldots, x^{(n-r)}$, то по определению базиса для любого решения x мы будем иметь равенство

$$x=\alpha_1 x^{(1)}+\alpha_2 x^{(2)}+\ldots+\alpha_{n-r} x^{(n-r)}. \qquad (8)$$

Поскольку в формуле (8) заключено любое решение системы (2), эта формула дает *общее решение системы (2)*.

См. задачу 8.

3.53. Переходим теперь к рассмотрению неоднородной системы (3) в общем случае. Как было показано в *2.62б*, геометрический образ H, отвечающий совокупности всех решений неоднородной системы, есть гиперплоскость в n-мерном пространстве K_n, полученная сдвигом подпространства L решений соответствующей однородной системы (по доказанному изоморфного пространству R_{n-r}) на некоторый вектор x_0, являющийся произвольным частным решением неоднородной системы. Из этого мы выводим прежде всего, что *размерность гиперплоскости H совпадает с размерностью подпространства L.*

Далее, если r — ранг основной матрицы системы (3), то любой вектор y подпространства L можно представить в виде суммы

$$y=\alpha_1 y^{(1)}+\alpha_2 y^{(2)}+\ldots+\alpha_{n-r} y^{(n-r)},$$

где $y^{(1)}, y^{(2)}, \ldots, y^{(n-r)}$ — базисные векторы подпространства L (фундаментальная система решений). Следовательно, любой вектор x гиперплоскости H можно представить в виде суммы

$$x=x_0+y=x_0+\alpha_1 y^{(1)}+\alpha_2 y^{(2)}+\ldots+\alpha_{n-r} y^{(n-r)}.$$

Этот результат на языке решений систем (3) и (2) согласуется с принципом, установленным в *3.41*: *Общее решение неоднородной системы* (3) *равно сумме любого частного решения этой системы и общего решения соответствующей однородной системы* (2).

См. задачу 9.

§ 3.6. Методы вычисления ранга матрицы

3.61. Для практического использования развитых в предыдущих параграфах методов решения систем линейных уравнений необходимо уметь вычислять ранг матрицы и находить ее базисный минор. Очевидно, что определение ранга матрицы, данное в *1.92*, само по себе не может служить разумным способом практического вычисления ранга; например, в квадратной матрице 5-го порядка можно выделить один минор 5-го порядка, 25 миноров 4-го порядка, 100 миноров 3-го порядка и 100 миноров 2-го порядка; понятно, что если бы мы пожелали найти ранг этой матрицы с помощью прямого вычисления величины всех ее миноров, это была бы весьма трудоемкая задача. В этом пункте будут даны простые способы подсчета ранга матрицы и определения ее базисного минора. Эти способы основаны на изучении некоторых операций со столбцами и строками матрицы, которые не изменяют ее ранга; мы будем называть эти операции *элементарными операциями.* Поскольку ранг матрицы, как мы уже указывали, не меняется при транспонировании, мы будем определять эти операции только для столбцов матрицы. В связи с этим при доказательствах мы будем использовать геометрическую интерпретацию матрицы с n строками и k столбцами как матрицы из координат некоторой системы k векторов $x_1, x_2, \ldots, x_k$ n-мерного пространства R_n и теорему *3.11*, согласно которой ранг этой матрицы равен размерности линейной оболочки векторов $x_1, x_2, \ldots, x_k$.

а. Перестановка столбцов. Пусть в матрице A произвольно переставлены ее столбцы; покажем, что эта операция не изменяет ее ранга. Действительно, размерность линейной оболочки векторов $x_1, x_2, \ldots, x_k$ не зависит от того, в каком порядке они записаны; следовательно, и ранг матрицы не зависит от порядка ее столбцов.

б. Отбрасывание ненулевого общего множителя элементов данного столбца. Допустим, что речь идет об отбрасывании общего множителя $\lambda \neq 0$ у элементов первого столбца матрицы A. Эта операция равносильна замене системы векторов $\lambda x_1, x_2, \ldots, x_k$ на систему $x_1, x_2, \ldots, x_k$; но очевидно, что размерности линейных оболочек этих систем одинаковы (так как сами линейные оболочки совпадают). Поэтому ранг матрицы A не меняется в результате этой элементарной операции.

в. Прибавление к одному столбцу другого столбца с произвольным множителем. Пусть к j-му столбцу матрицы A прибавлен m-й столбец этой же матрицы, умноженный на число λ. Это означает, что система векторов $x_1, \ldots, x_j, \ldots, x_m, \ldots, x_k$ заменена на систему $x_1, \ldots, x_j+\lambda x_m, \ldots, x_m, \ldots, x_k$. Покажем, что линейные оболочки L_1 и L_2 обеих этих систем совпадают. Действительно, все векторы второй системы входят в линейную оболочку векторов первой системы; поэтому в силу леммы *2.53а* $L_2 \subset L_1$. С другой стороны, равенство

$$x_j=(x_j+\lambda x_m)-\lambda x_m$$

показывает, что вектор x_j входит в линейную оболочку векторов второй системы; так как все остальные векторы первой системы, очевидно, также входят в эту линейную оболочку, то $L_1 \subset L_2$. Отсюда вытекает, что $L_1=L_2$. Поэтому ранг матрицы A не меняется в результате рассматриваемой элементарной операции.

г. Зачеркивание столбца, состоящего из одних нулей. Столбец из одних нулей отвечает нулевому вектору пространства R_n. Очевидно, что от зачеркивания нулевого вектора в системе $x_1, x_2, \ldots, x_k$ линейная оболочка $L(x_1, x_2, \ldots, x_k)$ не может измениться; вместе с нею не может измениться и ранг матрицы A.

д. Зачеркивание столбца, являющегося линейной комбинацией других столбцов. Законность этой элементарной операции была доказана в *3.14*.

Подчеркнем еще раз, что все предложения, доказанные в *а — д* для столбцов матрицы A, справедливы также и для ее строк.

3.62. Подсчет ранга матрицы и отыскание базисного минора. Покажем, как можно подсчитать ранг заданной матрицы A и найти ее базисный минор, используя операции, перечисленные в *а — д*. Если матрица A состоит из одних нулей, то ее ранг, очевидно, равен нулю. Допустим, что в матрице A имеется элемент, отличный от нуля; тогда, переставляя строки и столбцы, можно перевести этот элемент в левый верхний угол матрицы. Затем, вычитая из каждого столбца первый столбец с некоторым коэффициентом, можно обратить в нуль все остальные элементы первой строки. Больше мы не будем менять элементы первой строки и первого столбца *). Если среди остальных элементов (т. е. не принадлежащих к первой строке или к первому столбцу) нет

*) Но, может быть, будем переставлять их.

элементов, отличных от нуля, то ранг матрицы A, очевидно, равен 1. Если среди них имеется элемент, отличный от нуля, то, переставляя строки и столбцы, можно перевести его на пересечение второй строки и второго столбца и, так же как и раньше, обратить в нуль все следующие за ним элементы второй строки; отметим, что указанные операции не затрагивают первой строки и первого столбца. После этого вторую строку и второй столбец мы оставляем в покое. Продолжая таким образом далее, мы приведем матрицу A к одному из следующих двух видов (считая, что количество столбцов матрицы A не больше количества ее строк, чего всегда можно добиться транспонированием):

$$A_1 = \left\| \begin{matrix} \alpha_1 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ c_{21} & \alpha_2 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ c_{31} & c_{32} & \alpha_3 & \ldots & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & c_{k2} & c_{k3} & \ldots & \alpha_k & 0 & \ldots & 0 \\ c_{k+1,1} & c_{k+1,2} & c_{k+1,3} & \ldots & c_{k+1,k} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n1} & c_{n2} & c_{n3} & \ldots & c_{nk} & 0 & \ldots & 0 \end{matrix} \right\|$$

или

$$A_2 = \left\| \begin{matrix} \alpha_1 & 0 & 0 & \ldots & 0 \\ c_{21} & \alpha_2 & 0 & \ldots & 0 \\ c_{31} & c_{32} & \alpha_3 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{m1} & c_{m2} & c_{m3} & \ldots & \alpha_m \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n1} & c_{n2} & c_{n3} & \ldots & c_{nm} \end{matrix} \right\|,$$

причем числа α_1, α_2, ... отличны от нуля. В первом случае ранг матрицы A_1 равен k и базисный минор (в преобразованной матрице) стоит в левом верхнем углу; во втором случае ранг матрицы A_2 равен m (числу столбцов) и базисный минор (в преобразованной матрице) стоит в первых m строках. Ранг матрицы A, таким образом, определен; положение базисного минора можно восстановить, если проследить в обратном порядке за всеми операциями, которые производились с матрицей A.

Для примера рассмотрим следующую матрицу с пятью столбцами и шестью строками:

$$A=\begin{Vmatrix} 1 & 2 & 6 & -2 & -1 \\ -2 & -1 & 0 & -5 & -1 \\ 3 & 1 & -1 & 8 & 1 \\ -1 & 0 & 2 & -4 & -1 \\ -1 & -2 & -7 & 3 & 2 \\ -2 & -2 & -5 & -1 & 1 \end{Vmatrix}.$$

Во второй строке матрицы A имеется один нуль; используя общий метод, мы можем получить в ней еще три нуля. Для удобства вначале переставим вторую строку с первой, а затем первый столбец со вторым (чтобы в левом верхнем углу оказался наименьший по абсолютной величине элемент -1). Мы получаем *):

$$A\sim\begin{Vmatrix} -2 & -1 & 0 & -5 & -1 \\ 1 & 2 & 6 & -2 & -1 \\ 3 & 1 & -1 & 8 & 1 \\ -1 & 0 & 2 & -4 & -1 \\ -1 & -2 & -7 & 3 & 2 \\ -2 & -2 & -5 & -1 & 1 \end{Vmatrix}\sim\begin{Vmatrix} -1 & -2 & 0 & -5 & -1 \\ 2 & 1 & 6 & -2 & -1 \\ 1 & 3 & -1 & 8 & 1 \\ 0 & -1 & 2 & -4 & -1 \\ -2 & -1 & -7 & 3 & 2 \\ -2 & -2 & -5 & -1 & 1 \end{Vmatrix}.$$

Теперь для получения трех новых нулей в первой строке вычтем из второго, четвертого и пятого столбцов первый столбец с множителями 2, 5 и 1:

$$A\sim\begin{Vmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & -3 & 6 & -12 & -3 \\ 1 & 1 & -1 & 3 & 0 \\ 0 & -1 & 2 & -4 & -1 \\ -2 & 3 & -7 & 13 & 4 \\ -2 & 2 & -5 & 9 & 3 \end{Vmatrix}.$$

Далее проще всего добиваться новых нулей в третьей строке; мы вначале переставим ее со второй строкой, затем прибавим к третьему и четвертому столбцам второй со множителями 1 и -3:

$$A\sim\begin{Vmatrix} -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & 3 & 0 \\ 2 & -3 & 6 & -12 & -3 \\ 0 & -1 & 2 & -4 & -1 \\ -2 & 3 & -7 & 13 & 4 \\ -2 & 2 & -5 & 9 & 3 \end{Vmatrix}\sim\begin{Vmatrix} -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & -3 & 3 & -3 & -3 \\ 0 & -1 & 1 & -1 & -1 \\ -2 & 3 & -4 & 4 & 4 \\ -2 & 2 & -3 & 3 & 3 \end{Vmatrix}=A_1.$$

*) Знак $\sim$ между двумя матрицами означает в данном случае равенство их рангов.

Четвертый и пятый столбцы полученной матрицы A_1 пропорциональны третьему, и их можно зачеркнуть. Оставшаяся матрица имеет, очевидно, ранг 3; вместе с ней имеет ранг 3 и исходная матрица A.

Базисный минор матрицы A_1 расположен в ее первых трех строках и первых трех столбцах. Возвращаясь по цепочке преобразованных матриц к исходной матрице, мы легко можем проверить, что все произведенные преобразования не влияют на абсолютную величину этого минора. Следовательно, и в исходной матрице минор, стоящий в первых трех строках и первых трех столбцах, является базисным.

См. задачу 10 *).

ЗАДАЧИ

1. Доказать теорему: для того чтобы матрица m-го порядка $\| a_{ij} \|$ имела ранг $r \leqslant 1$, необходимо и достаточно, чтобы существовали числа $a_1, \ldots, a_m$ и $b_1, b_2, \ldots, b_m$ такие, что $a_{ij} = a_i b_j$ $(i, j = 1, 2, \ldots, m)$.

2. В n-мерном пространстве K_n взяты k линейно независимых векторов $x_1, x_2, \ldots, x_k$. Показать, что линейная оболочка $L = L(x_1, x_2, \ldots, x_k)$ определена однозначно, если известны значения всех миноров k-го порядка матрицы $A = \| a_i^{(j)} \|$ из координат векторов $x_1, x_2, \ldots, x_k$ в некотором базисе $e_1, e_2, \ldots, e_n$.

3. Показать, что система (2) при $k = n$ имеет решение

$$c_1 = A_{i1}, \quad c_2 = A_{i2}, \ \ldots, \ c_n = A_{in} \qquad (1 \leqslant i \leqslant n),$$

где A_{ik} — алгебраическое дополнение элемента a_{ik} (i фиксировано), если ранг матрицы A меньше n.

Примечание. Это обстоятельство позволяет легко строить ненулевые решения системы (2) в случае, когда ранг матрицы системы равен $n-1$.

4. Решить систему уравнений

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 7, \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 &= -2, \\ x_2 + 2x_3 + 2x_4 + 6x_5 &= 23, \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 &= 12. \end{aligned}$$

5. Исследовать решения системы

$$\begin{aligned} \lambda x + y + z &= 1, \\ x + \lambda y + z &= \lambda, \\ x + y + \lambda z &= \lambda^2 \end{aligned}$$

в зависимости от значения λ.

6. При каком условии три прямые $a_1x + b_1y + c_1 = 0$, $a_2x + b_2y + c_2 = 0$, $a_3x + b_3y + c_3 = 0$ проходят через одну точку?

7. При каком условии n прямых $a_1x + b_1y + c_1 = 0$, $a_2x + b_2y + c_2 = 0$, $\ldots$, $a_nx + b_ny + c_n = 0$ проходят через одну точку?

---

*) Далее в книге указания на относящиеся непосредственно к тексту задачи уже не делаются.

8. Написать нормальную фундаментальную систему решений для системы уравнений

$$\begin{aligned} x_1+x_2+x_3+x_4+x_5&=0,\\ 3x_1+2x_2+x_3+x_4-3x_5&=0,\\ x_2+2x_3+2x_4+6x_5&=0,\\ 5x_1+4x_2+3x_3+3x_4-x_5&=0. \end{aligned}$$

9. Написать общее решение системы задачи 4, используя нормальную фундаментальную систему решений соответствующей однородной системы, найденную в задаче 8.

10. Определить ранг и базисный минор следующих матриц:

$$A_1=\begin{Vmatrix} 1 & -2 & 3 & -1 & -1 & -2\\ 2 & -1 & 1 & 0 & -2 & -2\\ -2 & -5 & 8 & -4 & 3 & -1\\ 6 & 0 & -1 & 2 & -7 & -5\\ -1 & -1 & 1 & -1 & 2 & 1 \end{Vmatrix}, \qquad A_2=\begin{Vmatrix} 1&0&1&0&0\\ 1&1&0&0&0\\ 0&1&1&0&0\\ 0&0&1&1&0\\ 0&1&0&1&1 \end{Vmatrix}.$$

11. Пусть в матрице A имеется минор M r-го порядка, отличный от нуля, а всякий минор $(r+1)$-го порядка, включающий все элементы минора M, равен нулю. Доказать, что тогда ранг матрицы A равен числу r.

12. Построить матрицу

$$A=\begin{Vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{Vmatrix}$$

с заданными значениями миноров

$$\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}=P, \quad \begin{vmatrix} a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}=Q, \quad \begin{vmatrix} a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}=R.$$

13. Для квадратной системы уравнений

$$\sum_{k=1}^{n} a_{jk}x_k=b_j \qquad (j=1, \ldots, n) \tag{9}$$

доказать «альтернативу Фредгольма»: или система (9) имеет решение, притом единственное, для любых $b_1, \ldots, b_n$, или однородная система

$$\sum_{k=1}^{n} a_{jk}x_k=0 \qquad (j=1, \ldots, n)$$

имеет ненулевое решение.

14. Доказать, что система уравнений

$$\begin{aligned} a_{11}x_1+\ldots+a_{1n}x_n &= b_1,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\ldots\\ a_{n1}x_1+\ldots+a_{nn}x_n &= b_n,\\ a_{n+1,\,1}x_1+\ldots+a_{n+1,\,n}x_n &= b_{n+1} \end{aligned}$$

при условии

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \neq 0$$

разрешима тогда и только тогда, когда

$$\begin{vmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & b_n \\ a_{n+1,1} & \dots & a_{n+1,n} & b_{n+1} \end{vmatrix} = 0.$$

15. (Задача об исключении неизвестных.) Доказать, что содержащая параметры $y_1, \dots, y_k$ система

$$\begin{array}{l} a_{11}x_1 + \dots + a_{1n}x_n = b_{11}y_1 + \dots + b_{1k}y_k + c_1, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_{n1}y_1 + \dots + b_{nk}y_k + c_n, \\ a_{n+1,1}x_1 + \dots + a_{n+1,n}x_n = b_{n+1,1}y_1 + \dots + b_{n+1,k}y_k + c_{n+1} \end{array}$$

при условии

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \neq 0$$

разрешима тогда и только тогда, когда параметры $y_1, \dots, y_k$ удовлетворяют уравнению

$$y_1 \begin{vmatrix} a_{11} & \dots & a_{1n} & b_{11} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & b_{n1} \\ a_{n+1,1} & \dots & a_{n+1,n} & b_{n+1,1} \end{vmatrix} + \dots + y_k \begin{vmatrix} a_{11} & \dots & a_{1n} & b_{1k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & b_{nk} \\ a_{n+1,1} & \dots & a_{n+1,n} & b_{n+1,k} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{11} & \dots & a_{1n} & c_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & c_n \\ a_{n+1,1} & \dots & a_{n+1,n} & c_{n+1} \end{vmatrix} = 0.$$

ЛИНЕЙНЫЕ ФУНКЦИИ ВЕКТОРНОГО АРГУМЕНТА

В общем курсе математического анализа изучаются функции одного или нескольких вещественных переменных. В случае функции, например, трех вещественных переменных можно говорить о функции, аргументом которой является вектор пространства V_3. Мы будем здесь изучать функции, аргументом которых будет вектор произвольного линейного пространства. При этом мы ограничимся пока простейшими типами таких функций, а именно, линейными функциями. Мы будем рассматривать линейные *числовые* функции векторного аргумента, т. е. функции, значения которых суть числа, и линейные *векторные* функции векторного аргумента, значения которых суть векторы.

Векторные линейные функции, называемые иначе *линейными операторами*, имеют важное значение в линейной алгебре и ее приложениях.

§ 4.1. Линейные формы

4.11. Числовая функция $L(x)$ векторного аргумента x, определенная на линейном пространстве $\mathbf{K}$ над числовым полем K, называется *линейной формой*, если она удовлетворяет следующим условиям:

а) $L(x+y)=L(x)+L(y)$ *для любых* $x, y \in \mathbf{K}$;

б) $L(\alpha x)=\alpha L(x)$ *для любого* $x \in \mathbf{K}$ *и любого* $\alpha \in K$.

Иначе говоря, линейная форма $L(x)$ есть морфизм линейного пространства $\mathbf{K}$ в одномерное пространство $K_1=K$. (*2.71*).

Из условий а)—б) по индукции легко получить формулу

$$L(\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_k x_k)=$$
$$=\alpha_1 L(x_1)+\alpha_2 L(x_2)+\ldots+\alpha_k L(x_k), \quad (1)$$

справедливую для любых $x_1, x_2, \ldots, x_k \in \mathrm{K}$ и любых чисел $\alpha_1, \alpha_2, \ldots, \alpha_k$ из K.

4.12. Примеры.

а. Пусть в n-мерном пространстве K выбран базис, так что каждый вектор $x \in \mathrm{K}$ может быть задан своими координатами $\xi_1, \xi_2, \ldots, \xi_n$. Тогда $L(x) = \xi_1$ (первая координата) есть, очевидно, линейная форма от x.

б. Более общей линейной формой в том же пространстве является выражение

$$L(x) = \sum_{k=1}^{n} l_k \xi_k$$

с произвольно фиксированными коэффициентами $l_1, l_2, \ldots, l_n$.

в. В пространстве $K(a, b)$ (где K есть R или C) примером линейной формы является выражение

$$L(x) = x(t_0),$$

где t_0 — фиксированная точка отрезка $a \leqslant t \leqslant b$.

г. В том же пространстве можно рассматривать линейную форму вида

$$L(x) = \int_a^b l(t)\, x(t)\, dt,$$

где $l(t)$ — фиксированная непрерывная функция.

д. В пространстве V_3 скалярное произведение (x, x_0) вектора x с фиксированным вектором $x_0 \in V_3$ есть линейная форма от x.

Линейные формы, заданные в бесконечномерных пространствах, обычно называют *линейными функционалами*.

4.13. Найдем общий вид линейной формы $L(x)$ в n-мерном линейном пространстве K_n. Пусть $e_1, e_2, \ldots, e_n$ — произвольный базис пространства K_n. Обозначим число $L(e_k)$ через l_k $(k = 1, 2, \ldots, n)$. Тогда для любого $x = \sum_{k=1}^{n} \xi_k e_k$ в силу формулы (1)

$$f(x) = f\left(\sum_{k=1}^{n} \xi_k e_k\right) = \sum_{k=1}^{n} \xi_k f(e_k) = \sum_{k=1}^{n} l_k \xi_k,$$

т. е. *значения линейной формы $f(x)$ линейно выражаются через координаты вектора x с фиксированными коэффициентами $l_1, l_2, \ldots, l_n$.*

Таким образом, в примере *4.12б* мы имели самый общий вид линейной формы в n-мерном пространстве.

4.14. В комплексном линейном пространстве С рассматривают еще один вид линейной формы, называемый *линейной формой 2-го рода.* Линейная форма, определенная в *4.11*, называется в этом случае *формой 1-го рода.* Числовая функция $L(x)$ аргумента x, определенного на комплексном линейном пространстве С, называется линейной формой 2-го рода, если она удовлетворяет следующим условиям:

а) $L(x+y)=L(x)+L(y)$ для любых $x, y \in$ С;

б) $L(\alpha x)=\overline{\alpha}L(x)$ для любого $x \in$ С и любого комплексного числа $\alpha=\alpha_1+i\alpha_2$; число $\overline{\alpha}=\alpha_1-i\alpha_2$ есть здесь комплексно сопряженное к числу α.

Формула, аналогичная (1), в случае формы 2-го рода имеет вид

$$L(\alpha_1 x_1+\ldots+\alpha_k x_k)=\overline{\alpha}_1 L(x_1)+\ldots+\overline{\alpha}_k L(x_k) \qquad (2)$$

для любых $x_1, \ldots, x_k$ из С и любых комплексных $\alpha_1, \ldots, \alpha_k$.

4.15. Примером линейной формы 2-го рода в n-мерном комплексном пространстве C_n с базисом $e_1, \ldots, e_n$ служит функция

$$L(x)=\sum_{k=1}^{n} l_k \overline{\xi}_k,$$

где $l_1, \ldots, l_n$ — произвольно фиксированные комплексные числа, а $\xi_1, \ldots, \xi_n$ — координаты вектора x в базисе $e_1, \ldots, e_n$. Покажем, что эта формула дает общий вид линейной формы 2-го рода на пространстве C_n. Пусть $L(x)$ — произвольная линейная форма 2-го рода; положим $l_1=L(e_1), \ldots, l_n=L(e_n)$. Тогда для любого $x \in \mathrm{C}_n$ по формуле (2)

$$L(x)=L\left(\sum_{1}^{n}\xi_k e_k\right)=\sum_{1}^{n}\overline{\xi}_k L(e_k)=\sum_{1}^{n} l_k \overline{\xi}_k,$$

что и требуется.

§ 4.2. Линейные операторы и их матричная запись

4.21. Линейная форма $L(x)$, определенная в линейном пространстве K, как мы указывали, является морфизмом пространства K в одномерное пространство K_1.

Рассмотрим теперь морфизм $\mathrm{A}=\mathrm{A}(x)$ линейного пространства X в любое линейное пространство Y над тем же самым полем K. Будем писать иногда короче: $\mathrm{A}x$ вместо $\mathrm{A}(x)$. Согласно определению, функция $\mathrm{A}(x)$ удовлетворяет условиям:

а) $\mathrm{A}(x+y)=\mathrm{A}x+\mathrm{A}y$ для любых x и y из X;

б) $\mathrm{A}(\alpha x)=\alpha\mathrm{A}x$ для любого $x\in \mathrm{X}$ и любого числа α.

Как и для линейных форм, из условий а), б) следует более общая формула

в) $\mathrm{A}(\alpha_1x_1+\ldots+\alpha_kx_k)=\alpha_1\mathrm{A}x_1+\ldots+\alpha_k\mathrm{A}x_k$ для любых $x_1, \ldots, x_k$ из X и любых $\alpha_1, \ldots, \alpha_k$ из K.

Морфизм A называют также, как отмечено в *2.71*, *линейным оператором,* действующим из X в Y.

4.22. Примеры.

а. Оператор, который каждому вектору x пространства X ставит в соответствие нуль-вектор пространства Y, является, очевидно, линейным оператором. Он называется *нулевым оператором.*

б. Пусть имеется некоторый линейный оператор A, действующий из пространства X в пространство Y. Положим по определению $\mathrm{B}x=-\mathrm{A}x$. Полученный оператор B, как легко видеть, также является линейным оператором, переводящим X в Y; он называется оператором, *противоположным* оператору A.

в. Пусть векторам базиса $e_1, \ldots, e_n$ пространства X произвольно поставлены в соответствие векторы $f_1, \ldots, f_n$ пространства Y. *Тогда существует и единствен линейный оператор* A, *переводящий* X *в* Y *и при этом каждый вектор* e_k *переводящий в соответствующий вектор* f_k $(k=1, \ldots, n)$.

Действительно, если искомый оператор A существует, то для любого вектора $x=\sum_{k=1}^{n}\xi_ke_k\in \mathrm{X}$ выполняется равенство

$$\mathrm{A}x=\mathrm{A}\left(\sum_{k=1}^{n}\xi_ke_k\right)=\sum_{k=1}^{n}\xi_k\mathrm{A}e_k=\sum_{k=1}^{n}\xi_kf_k,$$

чем доказана единственность оператора A. С другой стороны, для любого $x = \sum_{k=1}^{n} \xi_k e_k \in X$ мы можем положить по определению

$$Ax = \sum_{k=1}^{n} \xi_k f_k.$$

Получающийся при этом оператор A, как легко проверить, является линейным оператором, переводящим X в Y и при этом переводящим каждый вектор e_k в соответствующий вектор f_k $(k = 1, \ldots, n)$, что нам и требуется.

г. Поставим в соответствие каждому вектору x пространства X этот же вектор x. Мы получим линейный оператор E, действующий из X в X. Этот оператор называется *тождественным*, или *единичным*, оператором.

4.23. Матричная запись линейных операторов. Пусть A есть линейный оператор, действующий из пространства X размерности n в пространство Y размерности m. Фиксируем в пространстве X базис $e_1, \ldots, e_n$ и в пространстве Y базис $f_1, \ldots, f_m$.

Вектор e_1 переводится оператором A в некоторый вектор Ae_1 пространства Y, который, как всякий вектор этого пространства, мы можем разложить по базисным векторам:

$$Ae_1 = a_1^{(1)} f_1 + a_2^{(1)} f_2 + \ldots + a_m^{(1)} f_m.$$

Аналогично оператор A действует на остальные базисные векторы:

$$\begin{gathered} Ae_2 = a_1^{(2)} f_1 + a_2^{(2)} f_2 + \ldots + a_m^{(2)} f_m, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ Ae_n = a_1^{(n)} f_1 + a_2^{(n)} f_2 + \ldots + a_m^{(n)} f_m. \end{gathered}$$

Эти формулы можно записать короче:

$$Ae_j = \sum_{i=1}^{m} a_i^{(j)} f_i \qquad (j = 1, 2, \ldots, n). \tag{3}$$

Коэффициенты $a_i^{(j)}$ $(i = 1, \ldots, m, \ j = 1, \ldots, n)$ определяют матрицу из m строк и n столбцов или, короче, $m\times n$-матрицу

$$A=A_{(e,\,f)}=\begin{Vmatrix} a_1^{(1)} & a_1^{(2)} & \dots & a_1^{(n)} \\ a_2^{(1)} & a_2^{(2)} & \dots & a_2^{(n)} \\ \cdot & \cdot & \cdot & \cdot \\ a_m^{(1)} & a_m^{(2)} & \dots & a_m^{(n)} \end{Vmatrix},$$

которая называется *матрицей оператора* A *в базисах* $\{e\}=\{e_1, \dots, e_n\}$ и $\{f\}=\{f_1, \dots, f_m\}$. Столбцами этой матрицы служат координаты векторов $\mathrm{A}e_1, \mathrm{A}e_2, \dots, \mathrm{A}e_n$ относительно базиса $f_1, \dots, f_m$.

Пусть теперь $x=\sum_{j=1}^{n}\xi_j e_j$ — произвольный вектор и $y=\mathrm{A}x=\sum_{i=1}^{m}\eta_i f_i$; выясним, как выражаются координаты η_i вектора y через координаты ξ_j вектора x. Мы имеем

$$y=\sum_{i=1}^{m}\eta_i f_i=\mathrm{A}x=\mathrm{A}\left(\sum_{j=1}^{n}\xi_j e_j\right)=\sum_{j=1}^{n}\xi_j \mathrm{A}e_j=$$
$$=\sum_{j=1}^{n}\xi_j\sum_{i=1}^{m}a_i^{(j)}f_i=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_i^{(j)}\xi_j\right)f_i.$$

Сравнивая коэффициенты при векторе f_i, находим

$$\eta_i=\sum_{j=1}^{n}a_i^{(j)}\xi_j \qquad (i=1, 2, \dots, m). \tag{4}$$

В раскрытом виде

$$\left.\begin{aligned} \eta_1 &= a_1^{(1)}\xi_1+a_1^{(2)}\xi_2+\dots+a_1^{(n)}\xi_n, \\ \eta_2 &= a_2^{(1)}\xi_1+a_2^{(2)}\xi_2+\dots+a_2^{(n)}\xi_n, \\ &\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ \eta_m &= a_m^{(1)}\xi_1+a_m^{(2)}\xi_2+\dots+a_m^{(n)}\xi_n. \end{aligned}\right\} \tag{5}$$

Следовательно, зная матрицу оператора A в базисе $e_1, e_2, \dots, e_n$, можно определить результат применения оператора к любому вектору $x=\sum_{k=1}^{n}\xi_k e_k$ пространства X: *координаты вектора* $y=\mathrm{A}x$ *линейно выражаются через координаты вектора* x по формулам (5). Заметим, что *матрица коэффициентов в формулах* (5) *совпадает с матрицей* $\mathrm{A}_{(e,\,f)}$.

Пусть теперь $\| a_i^{(j)} \|$ *) — произвольная $m \times n$-матрица. Если $x = \sum_{j=1}^{n} \xi_j e_j$, мы можем построить вектор $y = \sum_{i=1}^{m} \eta_i f_i$ по формулам (5).

Легко проверить, что оператор A, задающий этот переход от вектора x к вектору y, является линейным оператором. Построим матрицу оператора A в базисе $e_1, e_2, \ldots, e_n$. Вектор e_1 имеет координаты $\xi_1 = 1$, $\xi_2 = \ldots = \xi_n = 0$; в силу формул (5) координатами вектора $f_1 = \mathrm{A} e_1$ будут числа $a_1^{(1)}, a_2^{(1)}, \ldots, a_n^{(1)}$, так что

$$f_1 = \mathrm{A} e_1 = a_1^{(1)} e_1 + a_2^{(1)} e_2 + \ldots + a_n^{(1)} e_n.$$

Аналогично

$$f_j = \mathrm{A} e_j = a_1^{(j)} e_1 + a_2^{(j)} e_2 + \ldots + a_n^{(j)} e_n \qquad (j = 1, 2, \ldots, m).$$

Следовательно, матрица оператора A совпадает с исходной матрицей $\| a_i^{(j)} \|$.

Итак, *каждая $m \times n$-матрица является матрицей некоторого линейного оператора* A, *действующего из n-мерного пространства* X *в m-мерное пространство* Y, *с фиксированными базисами $e_1, \ldots, e_n$ в пространстве* X *и $f_1, \ldots, f_m$ в пространстве* Y.

Тем самым между линейными операторами, действующими из пространства X (с фиксированным базисом $e_1, \ldots, e_n$) в пространство Y (с фиксированным базисом $f_1, \ldots, f_m$), и $m \times n$-матрицами, заполненными числами из поля K, устанавливается *взаимно однозначное соответствие*, осуществляемое с помощью формул (3) или, что то же, (4).

Заметим, что оператор A можно было бы восстановить по матрице $A = \| a_i^{(j)} \|$ (и притом однозначно), исходя непосредственно из формул (5). В этих формулах j-й столбец матрицы A представляет собой набор координат вектора $f_j = \mathrm{A} e_j$.

4.24. Примеры.

а. Матрица нулевого оператора *(4.22а)* в любом базисе пространства X и любом базисе пространства Y, очевидно, состоит из одних нулей.

*) Верхний индекс — номер столбца, нижний — номер строки.

б. Если $\|a_i^{(j)}\|$ есть матрица оператора A, то матрицей противоположного оператора (*4.22б*) является, очевидно, матрица $\|-a_i^{(j)}\|$.

в. Пусть $m \geqslant n$ и оператор A переводит векторы базиса $e_1, \ldots, e_n$ пространства X в линейно независимые векторы $f_1, \ldots, f_n$ пространства Y. Дополним векторы $f_1, \ldots, f_n$ до полного базиса пространства Y векторами $f_{n+1}, \ldots, f_m$. Тогда матрица оператора A в базисах $e_1, \ldots, e_n$ и $f_1, \ldots, f_m$ примет, очевидно, следующий вид:

$$m\left\{\begin{array}{l} n\left\{\begin{array}{l} \\ \\ \\ \\ \end{array}\right. \\ \\ \\ \\ \end{array}\right. \overbrace{\left\|\begin{array}{cccc} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ . & .. & \ldots & .. \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & \ldots & 0 \\ . & .. & \ldots & .. \\ 0 & 0 & \ldots & 0 \end{array}\right\|}^{n}.$$

г. В частности, матрица тождественного оператора (*4.22г*) в базисе $e_1, \ldots, e_n$ пространства X, как области определения, и в том же базисе $e_1, \ldots, e_n$ пространства X, как области значений, имеет вид

$$\left\|\begin{array}{cccc} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ . & .. & \ldots & .. \\ 0 & 0 & \ldots & 1 \end{array}\right\|. \qquad (5')$$

Матрица E вида (5′) называется *тождественной* или *единичной* $n \times n$-матрицей.

§ 4.3. Действия над линейными операторами

Мы рассмотрим здесь действия сложения операторов, умножения оператора на число и умножения операторов друг на друга. Два оператора A и B, действующих из пространства X в пространство Y, будем считать *равными* (и писать $A = B$), если $Ax = Bx$ для каждого $x \in X$.

4.31. Сложение операторов. Пусть даны два линейных оператора A и B, отображающих пространство X в пространство Y. Оператор $C = A + B$ определяется по формуле

$$Cx \equiv (A + B)\, x = Ax + Bx. \tag{6}$$

Очевидно, что C также отображает пространство X в пространство Y. Покажем что C — снова линейный оператор. Пусть $x = \alpha_1 x_1 + \alpha_2 x_2$; тогда

$$\begin{aligned} C(\alpha_1 x_1 + \alpha_2 x_2) &= A(\alpha_1 x_1 + \alpha_2 x_2) + B(\alpha_1 x_1 + \alpha_2 x_2) = \\ &= \alpha_1 Ax_1 + \alpha_2 Ax_2 + \alpha_1 Bx_1 + \alpha_2 Bx_2 = \\ &= \alpha_1 (Ax_1 + Bx_1) + \alpha_2 (Ax_2 + Bx_2) = \alpha_1 Cx_1 + \alpha_2 Cx_2. \end{aligned}$$

Таким образом, оба условия *4.21 а)* — *б)* выполнены.

Линейный оператор C, определенный равенством (6), называется *суммой операторов* A и B.

Легко проверить следующие равенства:

$$\left.\begin{aligned} A + B &= B + A, \\ (A + B) + C &= A + (B + C), \\ A + 0 &= A, \\ A + (-A) &= 0. \end{aligned}\right\} \tag{7}$$

Здесь A, B, C — произвольные линейные операторы, 0 — нулевой оператор (*4.22 а*), — A — оператор, противоположный к оператору A, т. е. переводящий каждый вектор $x \in X$ в вектор $-Ax$ (*4.22 б*).

4.32. Умножение оператора на число. Если A — линейный оператор, действующий из пространства X в пространство Y, и λ — число из поля K, то оператор $B = \lambda A$, называемый *произведением оператора* A *на число* λ, определяется формулой

$$Bx \equiv (\lambda A)\, x = \lambda (Ax).$$

Легко проверить, так же как в *4.31*, что построенный оператор является линейным. При этом имеют место соотношения

$$\left.\begin{aligned} \lambda_1 (\lambda_2 A) &= (\lambda_1 \lambda_2)\, A, \\ 1 \cdot A &= A, \\ (\lambda_1 + \lambda_2)\, A &= \lambda_1 A + \lambda_2 A, \\ \lambda (A + B) &= \lambda A + \lambda B. \end{aligned}\right\} \tag{7'}$$

Соотношения (7) — (7′) показывают, что *совокупность всех линейных операторов, действующих из линейного пространства* X *в линейное пространство* Y, *образует новое линейное пространство.*

4.33. Умножение операторов. Если A — линейный оператор, действующий из пространства X в пространство Y, а B — линейный оператор, действующий из пространства Y в пространство Z (все пространства над одним и тем же числовым полем K), то мы определяем оператор $\mathrm{P} = \mathrm{BA}$, называемый *произведением* оператора B на оператор A, как оператор, действующий из пространства X в пространство Z по формуле

$$\mathrm{P}x \equiv (\mathrm{BA})\, x = \mathrm{B}(\mathrm{A}x)$$

(т. е. сначала на вектор x действует оператор A, а затем на результат, лежащий в пространстве Y, действует оператор B). Построенный оператор P является снова линейным, так как

$$\mathrm{P}(\alpha_1 x_1 + \alpha_2 x_2) = \mathrm{B}[\mathrm{A}(\alpha_1 x_1 + \alpha_2 x_2)] = \mathrm{B}(\alpha_1 \mathrm{A} x_1 + \alpha_2 \mathrm{A} x_2) =$$
$$= \alpha_1 \mathrm{BA} x_1 + \alpha_2 \mathrm{BA} x_2 = \alpha_1 \mathrm{P} x_1 + \alpha_2 \mathrm{P} x_2.$$

4.34. Легко проверяются следующие соотношения:

а) $\lambda(\mathrm{BA}) = (\lambda \mathrm{B})\mathrm{A}$ для любых операторов A и B с указанными свойствами и любого $\lambda \in K$;

б) $(\mathrm{A} + \mathrm{B})\mathrm{C} = \mathrm{AC} + \mathrm{BC}$ для любых операторов A и B, действующих из пространства Y в пространство Z, и любого C, действующего из X в Y;

в) $\mathrm{A}(\mathrm{B} + \mathrm{C}) = \mathrm{AB} + \mathrm{AC}$ для любых операторов B и C, действующих из X в Y, и любого A, действующего из Y в Z;

г) $(\mathrm{AB})\mathrm{C} = \mathrm{A}(\mathrm{BC})$ для любых операторов: C — из пространства X в пространство Y, B — из пространства Y в пространство Z, A — из пространства Z в пространство W.

Проверим, например, равенство г). В соответствии с принятым нами определением равенства операторов мы должны для любого вектора $x \in \mathrm{X}$ доказать равенство

$$[\mathrm{A}(\mathrm{BC})]\, x = [(\mathrm{AB})\,\mathrm{C}]\, x.$$

Но по самому определению произведения операторов

$$[\mathrm{A}(\mathrm{BC})]\,x = \mathrm{A}\,[(\mathrm{BC})\,x] = \mathrm{A}\,[\mathrm{B}\,(\mathrm{C}x)],$$
$$[(\mathrm{AB})\,\mathrm{C}]\,x = (\mathrm{AB})\,(\mathrm{C}x),$$

откуда и вытекает требуемое равенство.

Справедливость всех остальных равенств проверяется аналогично.

§ 4.4. Соответствующие действия над матрицами

Выясним, как отражаются действия над операторами, описанные в § 4.3, на матрицах этих операторов.

4.41. Сложение операторов. Пусть даны два линейных оператора A и B, отображающих пространство X с базисом $e_1, \ldots, e_n$ в пространство Y с базисом $f_1, \ldots, f_m$. Пусть, далее, оператору A в этих базисах соответствует матрица $A = \|a_i^{(j)}\|$, а оператору B — матрица $B = \|b_i^{(j)}\|$; следовательно,

$$\mathrm{A}e_j = \sum_{i=1}^{m} a_i^{(j)} f_i, \quad \mathrm{B}e_j = \sum_{i=1}^{m} b_i^{(j)} f_i \qquad (j = 1, 2, \ldots, n).$$

В таком случае

$$(\mathrm{A} + \mathrm{B})\,e_j = \mathrm{A}e_j + \mathrm{B}e_j = \sum_{i=1}^{m} (a_i^{(j)} + b_i^{(j)})\,f_i,$$

откуда следует, что оператору A + B соответствует матрица $\|a_i^{(j)} + b_i^{(j)}\|$. Такая матрица называется *суммой матриц* $\|a_i^{(j)}\|$ и $\|b_i^{(j)}\|$. Таким образом, сумма $A + B$ определена для всяких двух матриц A и B с одинаковым числом строк и с одинаковым числом столбцов.

4.42. Умножение оператора на число. При тех же условиях

$$(\lambda\mathrm{A})\,e_j = \lambda\,(\mathrm{A}e_j) = \sum_{i=1}^{m} \lambda a_i^{(j)} f_i.$$

Следовательно, оператору λA соответствует матрица $\|\lambda a_i^{(j)}\|$, получающаяся умножением всех элементов матрицы $\|a_i^{(j)}\|$

на число λ. Такая матрица называется *произведением матрицы* $\|a_i^{(j)}\|$ *на число* λ.

Поскольку $m\times n$-матрицы и линейные операторы, действующие из n-мерного пространства в m-мерное, соответствуют друг другу взаимно однозначно (*4.23*), действия с операторами соответствуют одноименным действиям с матрицами и действия с операторами подчиняются правилам (6)—(7), мы можем сделать вывод, что действия с матрицами также подчиняются правилам (6)—(7), что, впрочем, легко было бы проверить и непосредственно. Тем самым мы получаем, что *совокупность всех $m\times n$-матриц образует линейное пространство.* По самому построению оно изоморфно линейному пространству всех линейных операторов, действующих из n-мерного пространства X в m-мерное пространство Y.

4.43. Умножение операторов. Выберем в пространствах X, Y, Z базисы: в пространстве X — базис $e_1, \ldots, e_n$, в пространстве Y — базис $f_1, \ldots, f_m$, в пространстве Z — базис $g_1, \ldots, g_q$. Пусть оператор B, действующий из X в Y, имеет $m\times n$-матрицу $\|b_i^{(j)}\|$, так что

$$\mathrm{B}e_j = \sum_{i=1}^{m} b_i^{(j)} f_i \qquad (j=1, \ldots, n),$$

а оператор A, действующий из Y в Z, имеет $q\times m$-матрицу $\|a_k^{(i)}\|$, так что

$$\mathrm{A}f_i = \sum_{k=1}^{q} a_k^{(i)} g_k \qquad (i=1, \ldots, m).$$

Для произведения P = AB мы получаем

$$(\mathrm{AB})\, e_j = \mathrm{A}\,(\mathrm{B}e_j) = \mathrm{A} \sum_{i=1}^{m} b_i^{(j)} f_i =$$
$$= \sum_{i=1}^{m} b_i^{(j)} \mathrm{A} f_i = \sum_{i=1}^{m} b_i^{(j)} \sum_{k=1}^{q} a_k^{(i)} g_k = \sum_{k=1}^{q} \left(\sum_{i=1}^{m} a_k^{(i)} b_i^{(j)} \right) g_k.$$

Следовательно, элементы $p_k^{(j)}$ матрицы P оператора P = AB имеют вид

$$p_k^{(j)} = \sum_{i=1}^{m} a_k^{(i)} b_i^{(j)} \qquad (j=1, 2, \ldots, n;\ k=1, \ldots, q). \qquad (8)$$

Это и есть интересующий нас результат. Его можно выразить так: *элемент матрицы P, стоящий на пересечении ее k-й строки и j-го столбца, равен сумме произведений всех элементов k-й строки матрицы A на соответствующие элементы j-го столбца матрицы B.*

Матрица $P = \|p_k^{(j)}\|$, которая получается из матриц $\|a_k^{(i)}\|$ и $\|b_i^{(j)}\|$ по формуле (8), называется *произведением первой из этих матриц на вторую.*

Заметим, что число столбцов первой матрицы должно быть равно числу строк второй матрицы, иначе произведение матриц не будет иметь смысла. При выполнении этого условия матрица-произведение имеет столько строк, сколько их имеется у первой матрицы, и столько столбцов, сколько их имеется у второй матрицы.

Более выразительна $m \times n$-запись: произведение $q \times l$-матрицы A на $m \times n$-матрицу B определено, если $l = m$, и в этом случае произведение AB есть $q \times n$-матрица. Оба произведения AB и BA определены, если $l = m$, $q = n$; в этом случае AB есть квадратная $n \times n$-матрица, а BA есть также квадратная $m \times m$-матрица. Если, наконец, $n = m = p = q$, т. е. обе матрицы A и B суть квадратные $n \times n$-матрицы, то AB и BA — также квадратные $n \times n$-матрицы. Однако они не обязаны быть равными. Например,

$$\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix}, \quad \begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} = \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}.$$

Таким образом, умножение квадратных матриц, вообще говоря, некоммутативно. Что касается сочетательного и распределительного законов, то здесь положение более благополучное: умножение операторов, как мы видели в *4.34*, подчиняется сочетательному и распределительным законам; поскольку матрицы и операторы находятся во взаимно однозначном соответствии и умножение матриц соответствует умножению операторов, мы можем сделать вывод, что сочетательный и распределительные законы удовлетворяются также и для умножения матриц.

4.44. Примеры.

В следующих примерах индексы элементов матриц пишутся внизу, так что элемент a_{jk} матрицы $A = \|a_{jk}\|$ стоит на пересечении ее j-й строки и k-го столбца. Формула умножения $P=AB$ [см. *4.43* (8)] в этих обозначениях принимает вид

$$p_{kj}=\sum_{i=1}^{m} a_{ki} b_{ij} \qquad (j=1, \ldots, n, \ k=1, \ldots, q).$$

а. Умножим $m\times n$-матрицу $A=\|a_{jk}\|$ слева на $m\times m$-матрицу $B_{rs}=\|b_{jk}\|$, в которой все элементы b_{jk} равны 0, кроме одного элемента b_{rs}, равного 1. По общему правилу (8) мы получим $m\times n$-матрицу

$$B_{rs}A=\begin{array}{cc} & (s) \\ (r) & \begin{Vmatrix} & \cdot & \\ & \cdot & \\ & \cdot & \\ \ldots & 1 & \ldots \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ & \cdot & \end{Vmatrix} \end{array} \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & \ldots & a_{sn} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{Vmatrix} = \begin{array}{cc} (r) & \begin{Vmatrix} 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{s1} & a_{s2} & \ldots & a_{sn} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 \end{Vmatrix} \end{array},$$

так что в r-й строке матрицы $B_{rs}A$ стоят элементы s-й строки матрицы A, а остальные элементы матрицы $B_{rs}A$ равны 0.

б. Умножим $m\times n$-матрицу $A=\|a_{jk}\|$ справа на $n\times n$-матрицу $C_{qt}=\|c_{jk}\|$, в которой все элементы c_{jk} равны 0, кроме одного элемента c_{qt}, равного 1. По общему правилу (8) мы получим $m\times n$-матрицу

$$AC_{qt}=\begin{Vmatrix} a_{11} & \ldots & a_{1q} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2q} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & \ldots & a_{mq} & \ldots & a_{mn} \end{Vmatrix} \begin{array}{cc} & (t) \\ (q) & \begin{Vmatrix} & \cdot & \\ & \cdot & \\ \ldots & 1 & \ldots \\ & \cdot & \\ & \cdot & \end{Vmatrix} \end{array} = \begin{array}{c} (t) \\ \begin{Vmatrix} 0 & \ldots & a_{1q} & \ldots & 0 \\ 0 & \ldots & a_{2q} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & a_{mq} & \ldots & 0 \end{Vmatrix} \end{array},$$

так что в t-м столбце матрицы AC_{qt} стоят элементы q-го столбца матрицы A, а остальные элементы матрицы AC_{qt} равны 0.

в. При тех же B_{rs}, A, C_{qt} получаем

$$B_{rs}AC_{qt} = \begin{array}{c} \\ \\ (r) \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} & & (t) & & \end{array} \\ \left\| \begin{array}{ccccc} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & a_{sq} & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{array} \right\| \end{array},$$

так что операция $B_{rs}AC_{qt}$ приводит к $m\times n$-матрице, в которой все элементы равны нулю, кроме стоящего на пересечении r-й строки и t-го столбца, который равен a_{sq}.

г. На какую $m\times m$-матрицу D нужно умножить слева $m\times n$-матрицу A, чтобы матрица DA совпала с матрицей A, в которой переставлены местами r-я и s-я строки?

Пример *а* показывает, как получить матрицу, в которой r-я строка есть s-я строка матрицы A, — умножением слева на $m\times m$-матрицу B_{rs}. Но остальные строки в результате равны 0. Теперь ясно: чтобы получить требуемую матрицу, нужно матрицу A умножить слева на $m\times m$-матрицу

$$D = B_{rs} + B_{sr} + \sum_{\substack{j\neq r\\ j\neq s}} B_{jj} = \begin{array}{c||ccccccc||} & & & (r) & & (s) & & \\ & 1 & & \vdots & & \vdots & & \\ & & \ddots & \vdots & & \vdots & & \\ & & 1 & & & & & \\ (r) & & \dots & 0 & \dots & 1 & \dots & \\ & & & & 1 & & & \\ & & & \vdots & \ddots & \vdots & & \\ & & & & 1 & & & \\ (s) & & \dots & 1 & \dots & 0 & \dots & \\ & & & & & & 1 & \\ & & & \vdots & & \vdots & \ddots & \\ & & & & & & & 1 \end{array}.$$

д. На какую $n\times n$-матрицу G нужно умножить справа $m\times n$-матрицу A, чтобы матрица AG совпала с матрицей A, в которой переставлены местами q-й и t-й столбцы?

Аналогично *г* мы получаем

$$G=C_{qt}+C_{tq}+\sum_{\substack{k\neq q\\ k\neq t}} C_{kk}.$$

е. На какую $m\times m$-матрицу F нужно умножить слева $m\times n$-матрицу A, чтобы получившаяся матрица совпала с матрицей A, в которой к r-й строке прибавлена s-я строка с коэффициентом λ?

После сказанного в *а* ответ очевиден: $F=E+\lambda B_{rs}$.

ж. На какую $n\times n$-матрицу H нужно умножить справа $m\times n$-матрицу A, чтобы получившаяся матрица совпала с матрицей A, у которой к t-му столбцу прибавлен q-й столбец с коэффициентом μ?

Ответ очевиден: $H=E+\mu C_{qt}$.

§ 4.5. Дальнейшие свойства умножения матриц

4.51. Умножение матриц, разбитых на блоки. В некоторых случаях бывает удобно разбивать перемножаемые матрицы на блоки и действовать далее с блоками.

Предположим, что нам дана $m\times n$-матрица A и $n\times p$-матрица B, которые разбиты на блоки:

$$A=(m)\left\{\overbrace{\left\|\begin{array}{c|c|c} A_{11} & A_{12} & \dots \\ \hline A_{21} & A_{22} & \dots \\ \hline \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{array}\right\|}^{n}\right., \quad B=(n)\left\{\overbrace{\left\|\begin{array}{c|c|c} B_{11} & B_{12} & \dots \\ \hline B_{21} & B_{22} & \dots \\ \hline \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{array}\right\|}^{p}\right..$$

Предположим, что в каждой блок-строке матрицы A столько же блоков, сколько в блок-столбцах матрицы B, и при этом ширина любого блока A_{jk} матрицы A совпадает с высотой любого блока B_{ks} матрицы B. Тогда все произведения $A_{jk}B_{ks}$ имеют смысл; это — прямоугольные матрицы, размеры которых зависят от индексов j и s, но не зависят от индекса k.

Интересующее нас правило умножения матриц состоит в следующем: *матрицу AB можно составить из блоков, построенных из блоков матриц A и B так же, как элементы матрицы AB составляются из элементов матриц A и B*:

$$AB = \left\| \begin{array}{c|c|c} A_{11}B_{11} + A_{12}B_{21} + \ldots & A_{11}B_{12} + \ldots & \ldots \\ \hline A_{21}B_{11} + \ldots & \ldots & \ldots \\ \hline \cdots\cdots\cdots\cdots & \cdots\cdots & \cdots \\ \cdots\cdots\cdots\cdots & \cdots\cdots & \cdots \\ \cdots\cdots\cdots\cdots & \cdots\cdots & \cdots \end{array} \right\|.$$

Действительно, пусть i — номер блок-строки матрицы A, содержащей k-ю строку самой матрицы A, и j — номер блок-столбца матрицы B, содержащего q-й столбец матрицы B. По общему правилу *4.43* элементы матрицы $P = AB$ имеют вид

$$p_{kq} = a_{k1}b_{1q} + \ldots + a_{kn}b_{nq} =$$
$$= (a_{k1}b_{1q} + \ldots + a_{kp}b_{pq}) + \ldots + (a_{kr}b_{rq} + \ldots + a_{kn}b_{nq}),$$

где скобки расставлены в соответствии с шириной блоков матрицы A (и высотой блоков матрицы B). Будем нумеровать строки и столбцы блоков теми же номерами, что и в самой матрице A. В первой скобке стоит элемент, стоящий на пересечении k-й строки и q-го столбца блока $A_{i1}B_{1j}$, во второй скобке — элемент, стоящий на пересечении k-й строки и q-го столбца блока $A_{i2}B_{2j}$, и т. д.; в результате получается элемент, стоящий на пересечении k-й строки и q-го столбца блока $A_{i1}B_{1j} + \ldots + A_{ir}B_{rj}$, что и утверждалось.

4.52. Умножение квазидиагональных матриц. Матрица называется *квазидиагональной*, если она имеет вид

$$A = \begin{Vmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{ss} \end{Vmatrix},$$

причем необозначенные блоки состоят из нулей. Предположим, что блок A_{kk} представляет собой $m_k \times n_k$-матрицу ($k = 1, \ldots, s$).

Рассмотрим квазидиагональную матрицу

$$B = \begin{Vmatrix} B_{11} & & & \\ & B_{22} & & \\ & & \ddots & \\ & & & B_{ss} \end{Vmatrix},$$

у которой блок B_{kk} представляет собой $n_k \times p_k$-матрицу ($k = 1, \ldots, s$). Матрицы A и B можно перемножить по правилу *4.51*, которое в данном случае немедленно приводит к результату

$$AB = \begin{Vmatrix} A_{11}B_{11} & & & \\ & A_{22}B_{22} & & \\ & & \ddots & \\ & & & A_{ss}B_{ss} \end{Vmatrix}. \qquad (9)$$

Таким образом, в данном случае матрица AB снова есть квазидиагональная матрица, причем блок $A_{kk}B_{kk}$ имеет m_k строк и p_k столбцов.

4.53. Произведение транспонированных матриц. Пусть дана $m\times n$-матрица $A=\|a_{jk}\|$. *Транспонированной по отношению к* A называется $n\times m$-матрица $A'=\|a'_{pq}\|$, для которой

$$a'_{pq}=a_{qp} \qquad (p=1, \ldots, n; \quad q=1, \ldots, m).$$

Пусть A есть $m\times n$-матрица и B есть $n\times p$-матрица, так что произведение $P=AB$ определено и является $m\times p$-матрицей. Определено и произведение транспонированных матриц $B'A'$, которое представляет собой $p\times m$-матрицу. Покажем, что справедлива формула

$$B'A'=(AB)'. \tag{10}$$

Для доказательства обозначим элементы матриц A, B, $P=AB$, A', B', P' соответственно через a_{ij}, b_{ij}, p_{ij}, $a'_{ij}=a_{ji}$, $b'_{ij}=b_{ji}$, $p'_{ij}=p_{ji}$. Равенство (8), определяющее элементы p_{ik}, мы можем переписать в виде

$$p_{ik}=p'_{ki}=\sum_{j=1}^{n} a_{ij}b_{jk}=\sum_{j=1}^{n} a'_{ji}b'_{kj}=\sum_{j=1}^{n} b'_{kj}a'_{ji}.$$

Суммирование проводится по индексу j при фиксированных индексах i и k. Фиксированные индексы указывают, что для образования элемента p'_{ki} в матрице B' используется k-я строка, а в матрице A' — i-й столбец. В результате образования суммы произведений соответствующих элементов получается элемент p'_{ki}, лежащий на пересечении k-й строки и i-го столбца матрицы P'. Но по определению произведения матриц это и означает, что матрица P' есть произведение матрицы B' на матрицу A'. Тем самым равенство (10) доказано.

4.54. Миноры произведения двух матриц. Пусть даны $m\times n$-матрица $A=\|a_{jk}\|$ и $n\times p$-матрица $B=\|b_{jk}\|$; построим $m\times p$-матрицу $P=AB=\|p_{jk}\|$. Фиксируем в матрице P строки с номерами $\alpha_1\leqslant \ldots \leqslant\alpha_k$ и столбцы с номерами $\beta_1\leqslant \ldots \leqslant\beta_k$ ($k\leqslant m$; $k\leqslant p$) и поставим себе целью вычислить минор $M = M_{\beta_1 \ldots \beta_k}^{\alpha_1 \ldots \alpha_k}(AB)$, построенный на фиксированных строках и столбцах:

$$M_{\beta_1, \ldots, \beta_k}^{\alpha_1, \ldots, \alpha_k}(AB) =$$
$$= \begin{vmatrix} a_{\alpha_1 1} b_{1\beta_1} + \ldots + a_{\alpha_1 n} b_{n\beta_1} & \ldots & a_{\alpha_1 1} b_{1\beta_k} + \ldots + a_{\alpha_1 n} b_{n\beta_k} \\ a_{\alpha_2 1} b_{1\beta_1} + \ldots + a_{\alpha_2 n} b_{n\beta_1} & \ldots & a_{\alpha_2 1} b_{1\beta_k} + \ldots + a_{\alpha_2 n} b_{n\beta_k} \\ \cdot & \cdot & \cdot \\ a_{\alpha_k 1} b_{1\beta_1} + \ldots + a_{\alpha_k n} b_{n\beta_1} & \ldots & a_{\alpha_k 1} b_{1\beta_k} + \ldots + a_{\alpha_k n} b_{n\beta_k} \end{vmatrix}. \quad (11)$$

Для вычисления используем линейное свойство определителя *(1.44)*. Столбец минора M с номером ν является суммой k «элементарных» столбцов с элементами вида $a_{\alpha_j i} b_{i\beta_\nu}$ (где индексы столбца i и ν фиксированы, индекс строки j изменяется от 1 до k). Поэтому весь минор M равен сумме k^k «элементарных» определителей, составленных только из «элементарных» столбцов. В каждом из элементарных столбцов множитель $b_{j\beta_\nu}$ вдоль столбца не меняется и его можно вывести за знак «элементарного» определителя. После этого каждый из «элементарных» определителей принимает следующий вид:

$$b_{i_1\beta_1} b_{i_2\beta_2} \ldots b_{i_k\beta_k} \begin{vmatrix} a_{\alpha_1 i_1} & a_{\alpha_1 i_2} & \ldots & a_{\alpha_1 i_k} \\ a_{\alpha_2 i_1} & a_{\alpha_2 i_2} & \ldots & a_{\alpha_2 i_k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{\alpha_k i_1} & a_{\alpha_k i_2} & \ldots & a_{\alpha_k i_k} \end{vmatrix}, \quad (12)$$

где $i_1, \ldots, i_k$ — некоторые числа от 1 до n. Если среди этих чисел есть одинаковые, то соответствующий элементарный определитель равен 0. Между прочим, так всегда будет, если $n < k$. Поэтому *если в матрице AB есть вообще миноры порядка $k > n$, то все они равны* 0.

Возвращаясь к случаю $k \leqslant n$, мы видим, что следует рассматривать лишь такие элементарные определители, для которых все индексы $i_1, \ldots, i_k$ различны. В этом случае определитель

$$M_{i_1, \ldots, i_n}^{\alpha_1, \ldots, \alpha_n} = \begin{vmatrix} a_{\alpha_1 i_1} & a_{\alpha_1 i_2} & \ldots & a_{\alpha_1 i_k} \\ a_{\alpha_2 i_1} & a_{\alpha_2 i_2} & \ldots & a_{\alpha_2 i_k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{\alpha_k i_1} & a_{\alpha_k i_2} & \ldots & a_{\alpha_k i_k} \end{vmatrix}$$

с точностью до знака совпадает с минором $M_{j_1, j_2, \ldots, j_k}^{\alpha_1, \alpha_2, \ldots, \alpha_k}$, где $j_1 \leqslant j_2 \leqslant \ldots \leqslant j_k$ суть индексы $i_1, \ldots, i_k$, переставленные в порядке возрастания. Выясним, какой именно знак нужно поставить перед минором $M_{i_1, \ldots, i_n}^{\alpha_1, \ldots, \alpha_n}$, чтобы получить в точности минор $M_{j_1, \ldots, j_k}^{\alpha_1, \ldots, \alpha_k}$. Для этого будем последовательно переставлять в миноре $M_{i_1, i_2, \ldots, i_k}^{\alpha_1, \alpha_2, \ldots, \alpha_k}$ соседние столбцы с тем, чтобы в результате получить их нормальное расположение (т. е. совпадающее с расположением столбцов в самой матрице A). При каждой перестановке двух соседних столбцов минор $M_{i_1, i_2, \ldots, i_n}^{\alpha_1, \alpha_2, \ldots, \alpha_n}$ изменит знак; с другой стороны, при этом число «беспорядков» в перестановке индексов $i_1, i_2, \ldots, i_n$ изменится на единицу. Так как в окончательном расположении столбцов нижние индексы идут в натуральном порядке, без «беспорядков», то *число последовательных перемен знака равно числу «беспорядков» в перестановке индексов* $i_1, i_2, \ldots, i_n$ *). Обозначим это число через $N(i)$. Тогда выражение (12) приобретает следующий вид:

$$(-1)^{N(i)} b_{i_1\beta_1} b_{i_2\beta_2} \ldots b_{i_k\beta_k} M_{i_1, \ldots, i_k}^{\alpha_1, \ldots, \alpha_k}(A). \qquad (13)$$

Чтобы получить величину M, мы должны сложить все выражения вида (13).

Будем сначала складывать выражения с одним и тем же набором $i_1 < \ldots < i_k$, так что числа $i_1, \ldots, i_k$ представляют некоторую перестановку в этом наборе. При этом общий множитель $M_{i_1, \ldots, i_k}^{\alpha_1, \ldots, \alpha_k}(A)$ можно вынести за скобки. В скобках остается величина

$$\sum (-1)^{N(i)} b_{i_1\beta_1} b_{i_2\beta_2} \ldots b_{i_k\beta_k},$$

которая, очевидно, есть минор $M_{\beta_1, \ldots, \beta_k}^{i_1, \ldots, i_k}(B)$. Окончательно мы получаем формулу

$$M_{\beta_1, \ldots, \beta_k}^{\alpha_1, \ldots, \alpha_k}(AB) = \sum M_{i_1, \ldots, i_k}^{\alpha_1, \ldots, \alpha_k}(A) M_{\beta_1, \ldots, \beta_k}^{i_1, \ldots, i_k}(B). \qquad (14)$$

Суммирование производится здесь по всем наборам номеров $i_1 < i_2 < \ldots < i_k$, причем сами эти числа могут изменяться

*) Предполагается, что при каждой перестановке индексов меньший индекс становится впереди большего и тем самым число беспорядков уменьшается ровно на единицу.

от 1 до n. Общее число слагаемых в этой сумме равно $C_n^k = \frac{n!}{k!\,(n-k)!}$.

Сформулируем результат в виде теоремы.

Теорема. *Каждый минор k-го порядка матрицы AB может быть выражен через миноры того же порядка матриц A и B по формуле* (14).

§ 4.6. Область значений и нуль-многообразие линейного оператора

4.61. Пусть дан линейный оператор A, действующий из линейного пространства X в линейное пространство Y. Мы используем здесь обозначение *2.71* A: X $\to$ Y.

Пусть n — размерность пространства X, а m — размерность пространства Y. Выберем произвольно базис $e_1, \ldots, e_n$ в пространстве X и базис $f_1, \ldots, f_m$ в пространстве Y. Тогда оператору A можно поставить в соответствие по правилу *4.23* $m \times n$-матрицу $A = \| a_i^{(j)} \|$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Обозначим через T(A) область значений оператора A, т. е. совокупность всех векторов $y = \mathrm{A}x$, $x \in$ X. Поставим задачу: вычислить по матрице A размерность подпространства T(A). Полагая

$$x = \sum_{k=1}^{n} \xi_k e_k,$$

мы получаем

$$y = \mathrm{A}x = \sum_{k=1}^{n} \xi_k \mathrm{A} e_k;$$

следовательно, область значений оператора A совпадает с линейной оболочкой векторов $\mathrm{A}e_1, \mathrm{A}e_2, \ldots, \mathrm{A}e_n$. Размерность линейной оболочки $\mathrm{L}(\mathrm{A}e_1, \mathrm{A}e_2, \ldots, \mathrm{A}e_n)$ согласно *2.54в* равна максимальному количеству линейно независимых векторов в системе $\mathrm{A}e_1, \mathrm{A}e_2, \ldots, \mathrm{A}e_n$. Мы знаем, что в столбцах матрицы оператора A выписаны координаты векторов $\mathrm{A}e_i$ относительно базиса $\{f\}$; таким образом, вопрос о максимальном количестве линейно независимых векторов в системе $\mathrm{A}e_j$ $(j = 1, 2, \ldots, n)$ немедленно сводится к вопросу о максимальном количестве линейно независимых столбцов у матрицы оператора A. Но это последнее в силу *3.12в* равно рангу матрицы оператора A. Итак, *размерность области значений линейного оператора* A, *действующего из n-мерного пространства* X *в m-мерное пространство* Y, *равна рангу матрицы оператора* A *в любом базисе* $\{e\}$ *пространства* X *и любом базисе* $\{f\}$ *пространства* Y.

Заметим, что выбор базисов в данном случае безразличен: следовательно, *ранг матрицы оператора* A *не зависит от выбора базисов, т. е. зависит только от самого оператора* A. В дальнейшем мы будем ранг матрицы оператора A (в любых базисах) называть просто *рангом оператора* A *и обозначать через* r_{A}.

4.62. Теперь обозначим через N(A) нуль-многообразие оператора A, т. е. совокупность всех тех векторов $x \in \mathrm{X}$, для которых $\mathrm{A}x = 0$. Поставим задачу: по матрице $A = \| a_i^{(j)} \|$ оператора A вычислить размерность подпространства N(A).

Пусть $x = \sum_{i=1}^{n} \xi_i e_i \in \mathrm{N}(\mathrm{A})$. Тогда система (5) принимает вид

$$\left. \begin{array}{l} a_1^{(1)}\xi_1 + a_1^{(2)}\xi_2 + \ldots + a_1^{(n)}\xi_n = 0, \\ a_2^{(1)}\xi_1 + a_2^{(2)}\xi_2 + \ldots + a_2^{(n)}\xi_n = 0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_m^{(1)}\xi_1 + a_m^{(2)}\xi_2 + \ldots + a_m^{(n)}\xi_n = 0. \end{array} \right\} \qquad (15)$$

Очевидно, что и обратно, всякий вектор $x \in \mathrm{X}$, координаты которого удовлетворяют системе (15), входит в нуль-многообразие оператора A. Таким образом, вопрос о размерности нуль-многообразия оператора A равнозначен вопросу о размерности подпространства решений системы (15). В силу *3.51* размерность n_{A} этого подпространства равна числу $n - r$, где r — ранг матрицы из коэффициентов системы, или, что то же самое, ранг оператора A; таким образом, $n_{\mathrm{A}} = n - r_{\mathrm{A}}$.

Итак, *размерность нуль-многообразия оператора* A *равна дополнению ранга оператора* A *до размерности пространства* X, *из которого действует оператор* A.

4.63. В частности, если морфизм $\mathrm{A}\colon \mathrm{X} \to \mathrm{Y}$ есть эпиморфизм, то $\mathrm{T}(\mathrm{A}) = \mathrm{Y}$ и, следовательно, $r_{\mathrm{A}} = m$. Если морфизм $\mathrm{A}: \mathrm{X} \longrightarrow \mathrm{Y}$ есть мономорфизм, то $\mathrm{N}(\mathrm{A})=0$ и, следовательно, $r_{\mathrm{A}}=n$. Верны и обратные утверждения. Именно, если ранг матрицы A равен числу m ее строк, то размерность $\mathrm{T}(\mathrm{A})$ совпадает с размерностью всего Y, откуда следует, что $\mathrm{T}(\mathrm{A})=\mathrm{Y}$. Поэтому *морфизм* A *есть эпиморфизм тогда и только тогда, когда* $r_{\mathrm{A}}=m$. Если ранг матрицы A равен числу n ее столбцов, то векторы $f_1=\mathrm{A}e_1, \ldots, f_n=\mathrm{A}e_n$ являются линейно независимыми и, следовательно, оператор A является мономорфизмом *(2.73в)*. Поэтому *морфизм* A *есть мономорфизм тогда и только тогда, когда* $r_{\mathrm{A}}=n$.

4.64. Следующее предложение есть обращение результатов *4.61—4.62:*

Теорема. *Пусть* X *есть n-мерное пространство*, Y — *произвольное пространство. Каковы бы ни были подпространства* $\mathrm{N} \subset \mathrm{X}$ *и* $\mathrm{T} \subset \mathrm{Y}$, *сумма размерностей которых равна n, существует линейный оператор* $\mathrm{A}: \mathrm{X} \longrightarrow \mathrm{Y}$, *для которого* $\mathrm{N}(\mathrm{A})=\mathrm{N}$, $\mathrm{T}(\mathrm{A})=\mathrm{T}$.

Доказательство. Обозначим размерности подпространств N и T соответственно через k и $m=n-k$. В подпространстве T выберем m линейно независимых векторов $f_1, f_2, \ldots, f_m$. Выберем, далее, произвольный базис $e_1, e_2, \ldots, e_n$ в пространстве X так, чтобы первые k векторов базиса лежали в подпространстве N *(2.43)*.

Определим оператор A условиями

$$\left.\begin{aligned} \mathrm{A}e_i &= 0 \quad (i=1, 2, \ldots, k), \\ \mathrm{A}e_{i+k} &= f_i \quad (i=1, 2, \ldots, m). \end{aligned}\right\} \tag{16}$$

Покажем, что оператор A удовлетворяет поставленным требованиям. Прежде всего, очевидно, что $\mathrm{T}(\mathrm{A})$ есть линейная оболочка векторов $f_i\ (i=1, 2, \ldots, m)$ и, следовательно, совпадает с подпространством T. Затем всякий вектор подпространства N по условию принадлежит к $\mathrm{N}(\mathrm{A})$; нам остается показать, что любой вектор пространства $\mathrm{N}(\mathrm{A})$ входит в N. Допустим, что для некоторого

$$x=\sum_{i=1}^{n} \xi_i e_i$$

будет $\mathrm{A}x=0$. Используя условия (16), мы получаем

$$0=\mathrm{A}x=\mathrm{A}(\xi_1 e_1+\ldots+\xi_n e_n)=\xi_{k+1}f_1+\ldots+\xi_n f_m.$$

Поскольку $f_1, f_2, \ldots, f_m$ линейно независимы, мы имеем $\xi_{k+1}=\ldots=\xi_n=0$. Но тогда $x=\xi_1 e_1+\ldots+\xi_k e_k \in \mathrm{N}$, что и утверждалось.

4.65. Следующая теорема о ранге произведения двух матриц вытекает из свойств только что введенных геометрических характеристик.

Теорема. *Ранг произведения AB матриц A и B не превосходит ранга каждого из сомножителей.*

Доказательство. Естественно, мы предполагаем, что число столбцов матрицы A совпадает с числом строк матрицы B, иначе их нельзя было бы перемножить. Пусть A есть $m\times n$-матрица, а B есть $n\times p$-матрица. Введем в рассмотрение линейные пространства X, Y, Z с размерностями, соответственно, n, m и p. В пространстве X выберем базис $e_1, \ldots, e_n$, в пространстве Y — базис $f_1, \ldots, f_m$, в пространстве Z — базис $g_1, \ldots, g_p$. Используя их, можно матрице A поставить в соответствие линейный оператор A: $\mathrm{X}\longrightarrow\mathrm{Y}$, а матрице B — линейный оператор B: $\mathrm{Z}\longrightarrow\mathrm{X}$. Произведению AB матриц A и B отвечает линейный оператор AB: $\mathrm{Z}\longrightarrow\mathrm{Y}$. Область значений оператора AB в силу самого его определения содержится в области значений оператора A. Поскольку согласно *4.61* размерность области значений любого оператора равна рангу соответствующей матрицы, мы получаем, что *ранг произведения двух матриц не превосходит ранга первого множителя.* Чтобы доказать, что он не превосходит также и ранга второго множителя, перейдем к транспонированным матрицам; используя *4.53*, мы получим ранг $AB=$ рангу $(AB)'=$ рангу $B'A'\leqslant$ ранга $B'=$ рангу B, что и требуется.

4.66. Ранг произведения двух матриц может быть и меньше, чем ранг каждого из сомножителей. Например, матрицы

$$A=\begin{Vmatrix}1 & 0\\ 0 & 0\end{Vmatrix}, \quad B=\begin{Vmatrix}0 & 1\\ 0 & 0\end{Vmatrix}$$

обе имеют ранг, равный единице, а их произведение

$$AB=\begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}$$

имеет нулевой ранг. Поэтому представляет интерес следующая теорема, дающая оценку ранга произведения двух матриц не сверху, а снизу.

Теорема. *Пусть A есть $m\times n$-матрица ранга r_A и B есть $n\times p$-матрица ранга r_B. Тогда ранг $m\times p$-матрицы AB не меньше*

$$r_A+r_B-n.$$

Доказательство. Покажем сначала, что любой оператор $\mathrm{A}\colon \mathrm{X}\to\mathrm{Y}$ ранга r переводит всякое k-мерное подпространство $\mathrm{X}'\subset\mathrm{X}$ в подпространство $\mathrm{Y}'\subset\mathrm{Y}$, размерность которого не ниже $r-(n-k)$. Выберем базис $e_1, e_2, \ldots, e_n$ в пространстве X так, чтобы первые k векторов базиса лежали в подпространстве X′ (*2.43*). Координаты векторов $\mathrm{A}e_1, \mathrm{A}e_2, \ldots, \mathrm{A}e_k$, порождающих подпространство Y′, в матрице оператора A занимают k первых столбцов. По условию в матрице оператора A имеется r линейно независимых столбцов. Эти столбцы разобьем на две группы: в первую группу отнесем те, которые имеют номера от 1 до k, а во вторую группу — те, которые имеют номера от $k+1$ до n. Численность второй группы не больше $n-k$; следовательно, численность первой группы не меньше $r-(n-k)$. Таким образом, подпространство Y′ имеет не менее $r-(n-k)$ линейно независимых векторов, что и утверждалось.

Пусть теперь $\mathrm{A}\colon \mathrm{X}\to\mathrm{Y}$ и $\mathrm{B}\colon \mathrm{Z}\to\mathrm{X}$ — линейные операторы, соответствующие перемножаемым матрицам. Оценка ранга матрицы оператора AB согласно *4.61* есть оценка размерности области значений этого оператора. Оператор B переводит все пространство Z в подпространство $\mathrm{T}(\mathrm{B})\subset\mathrm{X}$ размерности r_B. По доказанному оператор A переводит подпространство T(B) в подпространство, размерность которого не ниже $r_A-(n-r_B)=r_A+r_B-n$. Таким образом, область значений оператора AB, а с ней и ранг матрицы AB, имеют величину не ниже r_A+r_B-n, что и требовалось.

4.67. Следствие. *Если одна из перемножаемых матриц, т. е. $m\times n$-матрица A или $n\times p$-матрица B, имеет ранг, равный n, то ранг произведения равен рангу второй матрицы.*

Действительно, в этом случае оценки ранга произведения сверху и снизу, полученные в теоремах *4.64* и *4.65*, дают одинаковый результат, равный рангу второй матрицы.

4.68. Пусть дан линейный оператор A, переводящий линейное пространство X в линейное пространство Y. Линейный оператор B, переводящий линейное пространство Y в линейное пространство X, называется *левым обратным* к оператору A, если

$$BA = E$$

есть единичный оператор в пространстве X. Оператор A в этом случае называется *правым обратным* к оператору B. В каком случае оператор A (B) имеет левый (правый) обратный? На этот вопрос отвечает теорема:

Теорема. *Оператор* A: $X \longrightarrow Y$ *имеет левый обратный тогда и только тогда, когда* A *есть мономорфизм. Оператор* B: $Y \longrightarrow X$ *имеет правый обратный тогда и только тогда, когда* B *есть эпиморфизм.*

Доказательство. Пусть A есть мономорфизм и $T(A) \subset Y$ — его область значений. Каждому $y \in T(A)$ отвечает $x \in X$, для которого $Ax = y$, причем x определяется однозначно по y в силу предположенной мономорфности A. Пусть $Q \subset Y$ есть подпространство, дающее в прямой сумме с T(A) все пространство Y (*2.46*). Определим оператор B: $Y \longrightarrow X$ по следующему правилу. Для $y \in T(A)$ положим By равным тому (единственному) x, для которого $Ax = y$; для $y \in Q$ положим $By = 0$; для $y = y_1 + y_2$, где $y_1 \in T(A)$, $y_2 \in Q$, положим $By = By_1$. Оператор B, как легко видеть, линейный, и для каждого $x \in X$ мы имеем $BAx = x$, так что B есть левый обратный для A. Если A не есть мономорфизм, то существует вектор $x \in X$, отличный от 0 и такой, что $Ax = 0$. Тогда для любого B: $Y \longrightarrow X$ мы имеем $(BA)x = B(Ax) = B(0) = 0$, так что левого обратного для оператора A заведомо не существует.

Пусть B: $Y \longrightarrow X$ есть эпиморфизм и пусть $N(B) \subset Y$ есть нуль-многообразие оператора B, а $Q \subset Y$ в прямой сумме с N(B) дает все пространство Y. Так как $X = B(Y) = B(N(B) + Q) = B(Q)$, то отображение B: $Q \rightarrow X$ есть также эпиморфизм и, более того, изоморфизм, так как никакой элемент $y \in Q$, отличный от 0, не переходит при воздействии оператора B в нуль. Определим оператор A: $X \to Y$ по следующему правилу: для любого $x \in X$ вектор Ax есть тот (единственный) вектор $y \in Q$, для которого $By = x$. Оператор A, очевидно, линеен, и для каждого $x \in X$ мы имеем $BAx = x$, так что A есть правый обратный для B. Если B: $Y \to X$ не есть эпиморфизм, то для вектора $x \in X$, не входящего в $T(B)$, и любого оператора A: $X \to Y$ мы имеем $BAx \neq x$, так что B не имеет правого обратного. Теорема доказана.

4.69a. Мы знаем *(4.43)*, что результатом умножения $n \times m$-матрицы P на $m \times n$-матрицу A является квадратная $n \times n$-матрица

$$S = PA.$$

Если S — единичная $n \times n$-матрица (*4.24г*), то матрица P называется *левой обратной для матрицы A*. Аналогично результатом умножения $m \times n$-матрицы A на $n \times m$-матрицу Q является квадратная $m \times m$-матрица

$$T = AQ,$$

и если T — единичная $m \times m$-матрица, то Q называется *правой обратной для матрицы A*.

б. Используя результаты *4.63*, можно сформулировать теорему *4.68* в терминах ранга матрицы.

Теорема. *Тогда и только тогда некоторая $m \times n$-матрица A имеет левую обратную, когда ее ранг равен числу n; тогда и только тогда она имеет правую обратную, когда ее ранг равен числу m.*

§ 4.7. Линейные операторы, переводящие пространство K_n в себя

4.71. Рассмотрим линейный оператор A, переводящий пространство X в себя (так что в *4.21* следует положить $Y = X$). Будем называть такой оператор A *действующим в пространстве* X.

Пусть оператор A действует в n-мерном пространстве $X = K_n$. Выберем в пространстве X базис $e_1, \ldots, e_n$ и *этот же базис в* X *используем для построения матрицы операто-*

ра А. В соответствии с *4.22*, матрица A оператора А строится по формулам

$$\mathrm{A}e_k = \sum_{i=1}^{n} a_i^{(j)} e_i, \tag{17}$$

так что коэффициенты $a_i^{(j)}$ образуют на этот раз квадратную $n \times n$-матрицу; она называется *матрицей оператора* А *в базисе* $\{e\} = \{e_1, \ldots, e_n\}$. Мы будем ее иногда обозначать через $A(e)$. Соответствующая формула для координат вектора $y = \mathrm{A}x$, $y = \sum_{j=1}^{n} \eta_j e_j$, $x = \sum_{j=1}^{n} \xi_j e_j$, имеет вид (*4.23*):

$$\eta_i = \sum_{j=1}^{n} a_i^{(j)} \xi_j. \tag{18}$$

При фиксированном базисе $(e_1, \ldots, e_n) = \{e\}$ получается взаимно однозначное соответствие между всеми линейными операторами, действующими в пространстве K_n, и всеми квадратными $n \times n$-матрицами, заполненными элементами из поля K.

4.72. Примеры.

а. Оператор, который каждому вектору пространства X ставит в соответствие нуль-вектор, очевидно, является линейным. Он называется *нулевым оператором* (ср. *4.22а*).

Матрица нулевого оператора в любом базисе, очевидно, состоит из одних нулей.

б. Единичный или тождественный оператор Е, ставящий в соответствие каждому вектору $x \in \mathrm{X}$ сам вектор x, мы рассмотрели в *4.22г*.

Матрица единичного оператора имеет вид (*4.24г*)

$$E = \begin{Vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & 1 \end{Vmatrix}.$$

Такая матрица называется *единичной*.

в. Оператор A, который переводит каждый вектор x в λx, где λ — фиксированное число из поля K, очевидно, линеен; он называется *оператором подобия* (с коэффициентом подобия λ).

Аналогично предыдущему матрица оператора подобия в любом базисе имеет вид

$$\begin{Vmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \lambda \end{Vmatrix}.$$

г. На евклидовой плоскости V_2 векторы можно определять полярными координатами: $x=\{\varphi, \rho\}$. Оператор A, переводящий вектор $x=\{\varphi, \rho\}$ в $\mathrm{A}x=\{\varphi+\varphi_0, \rho\}$ (φ_0 — фиксированный угол), является, как легко проверить с помощью чертежа, линейным. Этот оператор называется *оператором поворота на угол* φ_0.

Для построения матрицы оператора поворота выберем в плоскости V_2 базис из двух единичных взаимно ортогональных векторов e_1, e_2. Построив чертеж, легко проверим, что вектор e_1 после поворота на угол φ_0 перейдет в вектор $\cos\varphi_0\, e_1+\sin\varphi_0\, e_2$, а вектор e_2 — в вектор $-\sin\varphi_0\, e_1+\cos\varphi_0\, e_2$. Следовательно, матрица оператора поворота в любом из указанных базисов имеет вид

$$\begin{Vmatrix} \cos\varphi_0 & -\sin\varphi_0 \\ \sin\varphi_0 & \cos\varphi_0 \end{Vmatrix}.$$

д. Пусть $e_1, e_2, \dots, e_n$ — некоторый базис в n-мерном пространстве K_n. Поставим в соответствие вектору $x=\sum_{k=1}^{n}\xi_k e_k$ вектор $\mathrm{P}x=\sum_{k=1}^{m}\xi_k e_k$, где $m<n$. Оператор P — линейный оператор; он называется *оператором проектирования* на подпространство K_m, порожденное векторами $e_1, e_2, \dots, e_m$.

Для построения матрицы оператора проектирования заметим, что под его воздействием векторы $e_1, e_2, \dots, e_m$ переходят в себя, а векторы $e_{m+1}, \dots, e_n$ — в нуль. Поэтому матрица оператора проектирования в базисе $e_1, e_2, \dots, e_n$

имеет вид

$$\begin{array}{r} \\ \\ \\ m\text{-я строка} \\ \\ \\ \\ \end{array}\left\|\begin{array}{cccccccc} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{array}\right\|.$$

е. Пусть $e_1, e_2, \ldots, e_n$ — базис в n-мерном пространстве K_n и даны n фиксированных чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$. Определим оператор A для векторов базиса условиями $\mathrm{A}e_1 = \lambda_1 e_1$, $\mathrm{A}e_2 = \lambda_2 e_2, \ldots, \mathrm{A}e_n = \lambda_n e_n$ и для любого другого вектора $x = \sum_{k=1}^{n} \xi_k e_k$, естественно, по линейности условием $\mathrm{A}x = \sum_{k=1}^{n} \lambda_k \xi_k e_k$. Полученный оператор называется *диагональным* относительно базиса $e_1, \ldots, e_n$ или *диагонализируемым* оператором.

Матрица диагонального относительно базиса $e_1, \ldots, e_n$ оператора в этом же самом базисе $e_1, e_2, \ldots, e_n$ имеет вид

$$\left\|\begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & \lambda_n \end{array}\right\|.$$

Элементы, отличные от нуля, могут находиться в этой матрице только на главной диагонали. Такая матрица называется *диагональной*; отсюда и название оператора. Заметим, что в другом базисе $f_1, \ldots, f_n$ матрица оператора, диагонального относительно базиса $e_1, \ldots, e_n$, уже не будет, вообще говоря, диагональной.

4.73а. Линейные операторы, действующие в пространстве X, можно по общим правилам *4.31—4.32* складывать друг с другом и умножать на числа, причем снова получаются линейные операторы, действующие в X.

Равенства (7) и (7′) показывают, что при введенных там операциях сложения и умножения на числа совокупность всех линейных операторов, действующих в пространстве X, сама становится линейным пространством над тем же полем K. Кроме того, для операторов, действующих в пространстве X, всегда определена и операция умножения *4.33*, в результате которой также получается оператор, действующий в пространстве X. В частности, если B — любой оператор, то

$$(\mathrm{BE})\,x = \mathrm{B}\,(\mathrm{E}x) = \mathrm{B}x = \mathrm{E}\,(\mathrm{B}x),$$

так что

$$\mathrm{BE} = \mathrm{EB} = \mathrm{B}.$$

Определим *степени* данного оператора A по правилам

$$\begin{aligned} &\mathrm{A}^1 = \mathrm{A},\\ &\mathrm{A}^2 = \mathrm{AA},\\ &\mathrm{A}^3 = \mathrm{A}^2\mathrm{A} = (\mathrm{AA})\,\mathrm{A} = \mathrm{A}\,(\mathrm{AA}) = \mathrm{AA}^2,\\ &\ldots\ldots\ldots\ldots\ldots\\ &\mathrm{A}^n = \mathrm{A}^{n-1}\mathrm{A} = \mathrm{AA}^{n-1}. \end{aligned}$$

При этом имеет место формула

$$\mathrm{A}^{m+n} = \mathrm{A}^m\mathrm{A}^n \qquad (m,\ n = 1,\ 2,\ \ldots), \tag{19}$$

которую легко доказать по индукции.

Еще положим по определению

$$\mathrm{A}^0 = \mathrm{E} \qquad \text{(тождественный оператор)}.$$

Очевидно формула (19) остается справедливой и в том случае, когда один из показателей обращается в нуль.

б. Пусть пространство X конечномерно, $\mathrm{X} = \mathrm{K}_n$. Фиксируем в пространстве X произвольно базис $e_1, \ldots, e_n$. Тогда каждому линейному оператору A, действующему в пространстве X, можно поставить в соответствие матрицу оператора A в этом базисе. В соответствии с правилами *4.41—4.43* вместе с операторами складываются, умножаются и возводятся в степени и соответствующие им матрицы. В этом случае легко можно найти размерность линейного пространства всех матриц n-го порядка. Именно, матрицы E_{jk}, имеющие один-единственный элемент, отличный от нуля, для определенности 1, на пересечении j-й строки и k-го столбца, очевидно, линейно независимы; с другой стороны, каждая матрица n-го порядка есть линейная комбинация указанных матриц E_{jk}. Таким образом, матрицы E_{jk} образуют базис в пространстве всех матриц n-го порядка. Поскольку число матриц E_{jk} равно n^2, это число n^2 и есть размерность пространства всех матриц n-го порядка (*2.35*).

Ту же размерность n^2 имеет, очевидно, и пространство всех линейных операторов, действующих в пространстве K_n.

4.74. Примеры.

а. Умножение на комплексное число $\omega = \alpha + i\beta$ есть линейное преобразование на плоскости $z = x + iy$, которое можно записать с помощью вещественной матрицы 2-го порядка. Из формулы умножения $(\alpha + i\beta)(x + iy) = (\alpha x - \beta y) + i(\beta x + \alpha y)$ следует, что в базисе 1, i соответствующая матрица имеет вид

$$\tilde{\omega} = \begin{Vmatrix} \alpha & -\beta \\ \beta & \alpha \end{Vmatrix}.$$

Таким образом, комплексным числам $\omega = \alpha + i\beta$ взаимно однозначно ставятся в соответствие вещественные матрицы $\tilde{\omega}$ 2-го порядка; нетрудно видеть, что при этом сумме и произведению чисел отвечают сумма и произведение соответствующих матриц. Говорят, что вещественные матрицы $\tilde{\omega}$ образуют *точное представление поля комплексных чисел*.

б. Обозначим через B_k $(k \geqslant 0)$ оператор «сдвига на k шагов по индексу»; по определению, он переводит каждый базисный вектор e_m в базисный вектор e_{m-k}, если $m - k > 0$, и в 0, если $m - k \leqslant 0$. Очевидно, $\mathrm{B}_0 = \mathrm{E}$, $\mathrm{B}_k \cdot \mathrm{B}_r = \mathrm{B}_{k+r}$; в частности, $\mathrm{B}_1^k = \mathrm{B}_k$. Матрица оператора B_1 имеет вид

$$\begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{Vmatrix}.$$

Матрица оператора B_k имеет вид $(k < n)$

$$\begin{array}{r} \\ \\ \\ k+1 \\ \\ \\ \end{array}\begin{Vmatrix} 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 1 & \ldots & 0 \\ . & . & . & . & . & . \\ 0 & \ldots & 0 & 0 & \ldots & 1 \\ . & . & . & . & . & . \\ 0 & \ldots & 0 & 0 & \ldots & 0 \end{Vmatrix}.$$

4.75. Определитель произведения двух матриц. Пусть $A=\|a_{jk}\|$, $B=\|b_{jk}\|$ — две произвольные $n\times n$-матрицы, $C=AB$ — их произведение. В силу теоремы *4.54*, примененной к минору $M^{1,\ldots,n}_{1,\ldots,n}(AB)$, т. е. к самому определителю матрицы AB, мы получим

$$\det AB=\det A\cdot\det B.$$

Таким образом, справедлива теорема:

Теорема. *Определитель произведения двух $n\times n$-матриц равен произведению определителей этих матриц.*

Существуют и прямые (т. е. не опирающиеся на предложения типа *4.54*) доказательства этой теоремы. Вот одно из них. Рассмотрим определитель порядка $2n$

$$D=\begin{vmatrix} b_{11} & \ldots & b_{1n} & -1 & 0 & \ldots & 0\\ b_{21} & \ldots & b_{2n} & 0 & -1 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ b_{n1} & \ldots & b_{nn} & 0 & 0 & \ldots & -1\\ 0 & \ldots & 0 & a_{11} & a_{12} & \ldots & a_{1n}\\ 0 & \ldots & 0 & a_{21} & a_{22} & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \ldots & 0 & a_{n1} & a_{n2} & \ldots & a_{nn}\end{vmatrix}.$$

В силу *1.82* определитель D равен произведению определителей матриц

$$A=\begin{Vmatrix} a_{11} & \ldots & a_{1n}\\ \cdot & \cdot & \cdot\\ a_{n1} & \ldots & a_{nn}\end{Vmatrix}\quad\text{и}\quad B=\begin{Vmatrix} b_{11} & \ldots & b_{1n}\\ \cdot & \cdot & \cdot\\ b_{n1} & \ldots & b_{nn}\end{Vmatrix}.$$

Но можно и другим путем получить величину определителя D. Используя числа -1, стоящие в первых n строках и в последних n столбцах определителя D, можно обратить в нуль все элементы, лежащие в последних n строках и в последних n столбцах определителя D. Для этого нужно к $(n+1)$-й строке определителя D прибавить первую, умноженную на a_{11}; затем вторую, умноженную на a_{12}; ... ...; n-ю, умноженную на a_{1n}, далее к $(n+2)$-й строке прибавить первую, умноженную на a_{21}, вторую, умноженную на a_{22}; ...; $2n$-ю, умноженную на a_{2n} и т. д., пока не дойдем до последней, $2n$-й, строки. В результате мы получим

$$D=\begin{vmatrix} b_{11} & \ldots & b_{1n} & -1 & 0 & \ldots & 0\\ b_{21} & \ldots & b_{2n} & 0 & -1 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ b_{n1} & \ldots & b_{nn} & 0 & 0 & \ldots & -1\\ b_{11}a_{11}+b_{21}a_{12}+\ldots+b_{n1}a_{1n} & \ldots & b_{1n}a_{11}+\ldots+b_{nn}a_{1n} & 0 & 0 & \ldots & 0\\ b_{11}a_{21}+b_{21}a_{22}+\ldots+b_{n1}a_{2n} & \ldots & b_{1n}a_{21}+\ldots+b_{nn}a_{2n} & 0 & 0 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ b_{11}a_{n1}+b_{21}a_{n2}+\ldots+b_{n1}a_{nn} & \ldots & b_{1n}a_{n1}+\ldots+b_{nn}a_{nn} & 0 & 0 & \ldots & 0\end{vmatrix},$$

откуда по теореме Лапласа *1.81*, разлагая определитель D по последним n строкам, мы будем иметь

$$D=(-1)^{1+\ldots+n}\begin{vmatrix} -1 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & -1 \end{vmatrix}\begin{vmatrix} b_{11}a_{11}+\ldots+b_{n1}a_{1n} & \ldots & b_{1n}a_{11}+\ldots+b_{nn}a_{1n} \\ b_{11}a_{21}+\ldots+b_{n1}a_{2n} & \ldots & b_{1n}a_{21}+\ldots+b_{nn}a_{2n} \\ \cdot & \cdot & \cdot \\ b_{11}a_{n1}+\ldots+b_{n1}a_{nn} & \ldots & b_{1n}a_{n1}+\ldots+b_{nn}a_{nn} \end{vmatrix}=$$

$$=\begin{vmatrix} a_{11}b_{11}+\ldots+a_{1n}b_{n1} & \ldots & a_{11}b_{1n}+\ldots+a_{1n}b_{nn} \\ a_{21}b_{11}+\ldots+a_{2n}b_{n1} & \ldots & a_{21}b_{1n}+\ldots+a_{2n}b_{nn} \\ \cdot & \cdot & \cdot \\ a_{n1}b_{11}+\ldots+a_{nn}b_{n1} & \ldots & a_{n1}b_{1n}+\ldots+a_{nn}b_{nn} \end{vmatrix}=\det(A,\ B).$$

Сравнивая этот результат с полученным выше значением определителя D, убеждаемся в справедливости теоремы.

В частности, отметим, что если обе перемножаемые матрицы A и B невырождены (т. е. $\det A \neq 0$, $\det B \neq 0$), то и матрица AB невырождена; если хотя бы одна из матриц, например A, вырождена, $\det A = 0$, то и $\det AB = 0$. Впрочем, это утверждение можно вывести и из *4.67*.

4.76. Обратный оператор. В соответствии с определением *4.68* оператор B, действующий в пространстве X, называется *левым обратным* к оператору A, действующему в том же пространстве X, если

$$\mathrm{BA} = \mathrm{E}.$$

Оператор A в этом случае называется *правым обратным* к B.

а. Возможно, что оператор A имеет много левых обратных и ни одного правого, или наоборот, много правых обратных и ни одного левого (см. задачу 25). Допустим, что оператор A обладает левым обратным P и правым обратным Q; тогда справедливо равенство

$$\mathrm{P} = \mathrm{PE} = \mathrm{P}\,(\mathrm{AQ}) = (\mathrm{PA})\,\mathrm{Q} = \mathrm{EQ} = \mathrm{Q}. \qquad (20)$$

Фиксируем Q; мы видим, что любой левый обратный оператор P совпадает с Q и тем самым определен единственным образом. Точно так же в рассматриваемом случае правый обратный Q также определен единственным образом. Этот оператор $\mathrm{P} = \mathrm{Q}$, определенный единственным образом и одновременно и правый, и левый обратный к оператору A, называется *обратным оператором к оператору* A и обозначается через A^{-1}. Сам оператор A, обладающий обратным, называется *обратимым оператором*.

б. Рассмотрим случай оператора A в n-мерном пространстве $X = K_n$. Пусть A — матрица оператора A в некотором фиксированном базисе $e_1, \ldots, e_n$. Возможно одно из двух: или $\det A \neq 0$, или $\det A = 0$. В первом случае ранг матрицы A равен n и по *4.69б* матрица A обладает и левой обратной матрицей, и правой обратной матрицей. Соответственно и оператор A обладает и левым обратным, и правым обратным. По *а* оператор A является обратимым оператором.

Если же $\det A = 0$, то по тому же *4.69б* матрица A не имеет ни левой, ни правой обратной; так же и соответствующий оператор A, действующий в пространстве K_n, не имеет ни левого, ни правого обратного.

4.77. Матрица обратного оператора. Пусть A — обратимый оператор в n-мерном пространстве X и $B = A^{-1}$ — его обратный оператор. Выберем базис $e_1, \ldots, e_n$ и обозначим через $\|a_i^{(j)}\|$ и $\|b_i^{(j)}\|$ матрицы операторов A и B в этом базисе.

Найдем явное выражение элементов $b_i^{(j)}$ через элементы $a_i^{(j)}$. Фиксируя номер строки i, последовательно выписываем выражения элементов i-й строки матрицы $E = AB$ по формулам (8):

$$\begin{gathered}
b_i^{(1)} a_1^{(1)} + b_i^{(2)} a_2^{(1)} + \ldots + b_i^{(n)} a_n^{(1)} = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
b_i^{(1)} a_1^{(i)} + b_i^{(2)} a_2^{(i)} + \ldots + b_i^{(n)} a_n^{(i)} = 1, \\
\cdots\cdots\cdots\cdots\cdots\cdots \\
b_i^{(1)} a_1^{(n)} + b_i^{(2)} a_2^{(n)} + \ldots + b_i^{(n)} a_n^{(n)} = 0.
\end{gathered}$$

Неизвестные $b_i^{(1)}, \ldots, b_i^{(n)}$ определяются из этой системы уравнений по правилу Крамера *1.73*, так как по условию $\det A \neq 0$. Разлагая определитель в числителе по j-му столбцу, получаем

$$b_i^{(j)} = \frac{A_j^{(i)}}{\det A}, \tag{21}$$

где $A_j^{(i)}$ — алгебраическое дополнение элемента $a_j^{(i)}$ в матрице A. Итак, *элемент $b_i^{(j)}$ обратной матрицы равен отношению алгебраического дополнения элемента $a_j^{(i)}$ исходной матрицы к ее определителю.*

Мы получаем теорему:

Теорема. *Для всякой невырожденной матрицы* $A = \|a_i^{(j)}\|$ *существует и единственна обратная матрица* $B = \|b_i^{(j)}\|$, *для которой*

$$AB = BA = E.$$

Элементы матрицы B вычисляются по формулам (21).

4.78. Оператор, обратный к оператору A, мы обозначили через A^{-1}. Далее, $(A^{-1})^k$ обозначается через A^{-k}. Легко доказать по индукции, что формула (19) распространяется также и на все отрицательные показатели.

Аналогичные обозначения применяются для степеней обратной матрицы. Распространение формулы (19) на отрицательные показатели для матриц вытекает непосредственно из справедливости этого распространения для операторов.

§ 4.8. Инвариантные подпространства

4.81. Пусть в линейном пространстве K задан линейный оператор A. Введем следующее определение. Подпространство K′ пространства K называется *инвариантным* относительно оператора A, если из $x \in K'$ следует $Ax \in K'$.

В частности, тривиальные подпространства — нулевое и все пространство — являются инвариантными для всякого линейного оператора; нас будут интересовать, естественно, нетривиальные инвариантные подпространства.

4.82. Рассмотрим с этой точки зрения примеры линейных операторов, указанные в *4.72*.

а — в. Для операторов в примерах *4.72а — в* (нулевой оператор, тождественный оператор и оператор подобия) *каждое подпространство является инвариантным.*

г. Оператор поворота (*4.72г*) на угол $\varphi_0 \neq m\pi$, m целое — *не имеет нетривиальных инвариантных подпространств.*

д. Оператор проектирования (*4.72д*) имеет, например, следующие инвариантные подпространства: подпространство K′ из векторов $x = \sum_{k=1}^{m} \xi_k e_k$, которые не изменяются, и подпространство K″ из векторов $y=\sum_{k=m+1}^{n}\xi_k e_k$, которые переходят в нуль.

е. Каждое подпространство, порожденное некоторыми из базисных векторов $e_1, e_2, \ldots, e_n$, является инвариантным для диагонального оператора (*4.72е*).

4.83. Пусть оператор A имеет в n-мерном пространстве K_n инвариантное m-мерное подпространство K_m. Выберем базис $e_1, \ldots, e_n$ пространства K_n так, чтобы его первые m векторов $e_1, \ldots, e_m$ лежали в подпространстве K_m. Тогда мы будем иметь

$$\begin{gathered} Ae_1 = a_1^{(1)}e_1 + \ldots + a_1^{(m)}e_m, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ Ae_m = a_m^{(1)}e_1 + \ldots + a_m^{(m)}e_m, \end{gathered}$$

и, следовательно, матрица оператора A в указанном базисе будет иметь вид

$$A = \begin{Vmatrix} a_1^{(1)} & \ldots & a_m^{(1)} & a_{m+1}^{(1)} & \ldots & a_n^{(1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_1^{(m)} & \ldots & a_m^{(m)} & a_{m+1}^{(m)} & \ldots & a_n^{(m)} \\ 0 & \ldots & 0 & a_{m+1}^{(m+1)} & \ldots & a_n^{(m+1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & a_{m+1}^{(n)} & \ldots & a_n^{(n)} \end{Vmatrix}.$$

В первых m столбцах матрицы все элементы $(m+1)$-й и следующих строк равны 0. Обратно, если матрица оператора A имеет такой вид, то подпространство, определяемое векторами $e_1, \ldots, e_m$, является инвариантным для оператора A.

4.84. Предположим, что пространство K_n можно представить в виде прямой суммы инвариантных подпространств E, F, ..., H (*2.45*). Выберем базис пространства K_n так, чтобы векторы

$$\begin{array}{lll} e_1, \ldots, e_r & \text{лежали в} & \mathrm{E}, \\ f_1, \ldots, f_s & \text{»} & \text{»}\ \mathrm{F}, \\ \ldots\ldots & & \\ h_1, \ldots, h_t & \text{»} & \text{»}\ \mathrm{H}. \end{array}$$

Тогда матрица оператора А примет квазидиагональный вид

$$A = \left\| \begin{matrix} A_e & & & \\ & \boxed{A_f} & & \\ & & \ddots & \\ & & & A_h \end{matrix} \right\| .$$

Диагональные квадраты матрицы A заполняются элементами $a_k^{(j)}, b_k^{(j)}, \ldots, c_k^{(j)}$ в соответствии с формулами

$$\mathrm{A} e_j = \sum_{k=1}^{r} a_k^{(j)} e_k,$$

$$\mathrm{A} f_j = \sum_{k=1}^{s} a_k^{(j)} f_k,$$

$$\cdots\cdots\cdots$$

$$\mathrm{A} h_j = \sum_{k=1}^{t} a_k^{(j)} h_k;$$

вне диагональных квадратов в матрице A всюду стоят нули. Обратно, если матрица оператора А в некотором базисе имеет квазидиагональную структуру, то пространство K_n разлагается в прямую сумму подпространств, порожденных соответствующими группами базисных элементов.

§ 4.9. Собственные векторы и собственные значения

4.91. Особую роль играют одномерные инвариантные подпространства оператора А; они называются иначе инвариантными, или собственными, направлениями. Всякий (ненулевой) вектор, принадлежащий к одномерному инвариантному направлению оператора А, называется *собственным вектором* оператора А; иначе говоря, вектор $x \neq 0$ называется *собственным вектором* оператора А, если оператор А переводит вектор x в коллинеарный ему вектор:

$$\mathrm{A} x = \lambda x. \tag{22}$$

Число λ, фигурирующее в этом равенстве, называется *собственным значением* (*собственным числом*) *оператора* A, *соответствующим собственному вектору* x.

4.92. Обратимся опять к примерам *4.72*.

а — в. В примерах *4.72а—в* каждый ненулевой вектор пространства есть собственный вектор соответствующего оператора с собственными значениями соответственно 0, 1, λ.

г. Оператор поворота (*4.72г*) на угол, не равный $m\pi$ с целым m, не имеет собственных векторов.

д. Оператор проектирования (*4.72д*) имеет собственные векторы вида $x=\sum_{k=1}^{m}\xi_k e_k$ и $y=\sum_{k=m+1}^{n}\xi_k e_k$ с собственными значениями соответственно 1 и 0. Можно проверить, что иных собственных векторов у оператора проектирования нет.

е. Диагональный оператор (*4.72е*) по самому определению имеет собственные векторы $e_1, e_2, \ldots, e_n$ с собственными значениями соответственно $\lambda_1, \lambda_2, \ldots, \lambda_n$.

4.93. Укажем два простых свойства собственных векторов.

а. Лемма. *Собственные векторы $x_1, x_2, \ldots, x_m$ оператора* A *с попарно различными собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_m$ линейно независимы.*

Это утверждение мы докажем индукцией по числу m. Очевидно, что для $m=1$ лемма верна. Допустим, что лемма верна для всяких $m-1$ собственных векторов оператора A; покажем, что она продолжает оставаться верной и для всяких m собственных векторов оператора A. Предполагая противное, допустим, что между m собственными векторами оператора A имеется линейная зависимость

$$\alpha_1 x_1+\alpha_2 x_2+\ldots+\alpha_m x_m=0,$$

где, например, $\alpha_1\neq 0$. Применяя к этому равенству оператор A, получаем

$$\alpha_1\lambda_1 x_1+\alpha_2\lambda_2 x_2+\ldots+\alpha_m\lambda_m x_m=0.$$

Умножим первое равенство на λ_m и вычтем из второго; мы получим

$$\alpha_1(\lambda_1-\lambda_m)x_1+\alpha_2(\lambda_2-\lambda_m)x_2+\ldots$$
$$\ldots+\alpha_{m-1}(\lambda_{m-1}-\lambda_m)x_{m-1}=0,$$

откуда по индуктивному предположению все коэффициенты должны быть равны нулю. В частности, $\alpha_1(\lambda_1-\lambda_m)=0$, что противоречит условиям $\alpha_1\neq 0$, $\lambda_1\neq\lambda_m$. Следовательно, наше предположение неверно и векторы $x_1, x_2, \ldots, x_m$ линейно независимы.

В частности, *в n-мерном пространстве линейный оператор* A *не может иметь более n собственных векторов с различными собственными значениями.*

б. Лемма. *Все собственные векторы линейного оператора* A, *отвечающие данному собственному значению* λ, *образуют подпространство* $K^{(\lambda)}\subset K$.

В самом деле, если $Ax_1=\lambda x_1$ и $Ax_2=\lambda x_2$, то

$$A(\alpha x_1+\beta x_2)=\alpha A x_1+\beta A x_2=\alpha\lambda x_1+\beta\lambda x_2=\lambda(\alpha x_1+\beta x_2),$$

чем утверждение леммы доказано.

Подпространство $K^{(\lambda)}$ называется *собственным подпространством оператора* A, *отвечающим собственному значению* λ.

4.94. Мы укажем здесь, как можно вычислить координаты собственных векторов оператора A, заданного своей матрицей в некотором базисе $e_1, \ldots, e_n$ пространства K_n. Допустим, что вектор $x=\sum_{k=1}^{n}\xi_k e_k$ есть собственный вектор оператора A, так что $Ax=\lambda x$ с некоторым λ. Используя формулы *4.23* (5), мы можем это равенство переписать в координатной форме:

$$\begin{aligned}
\lambda\xi_1&=a_1^{(1)}\xi_1+a_1^{(2)}\xi_2+\ldots+a_1^{(n)}\xi_n,\\
\lambda\xi_2&=a_2^{(1)}\xi_1+a_2^{(2)}\xi_2+\ldots+a_2^{(n)}\xi_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\lambda\xi_n&=a_n^{(1)}\xi_1+a_n^{(2)}\xi_2+\ldots+a_n^{(n)}\xi_n,
\end{aligned}$$

или

$$\left.\begin{aligned}
(a_1^{(1)}-\lambda)\,\xi_1+a_1^{(2)}\xi_2+\ldots+a_1^{(n)}\xi_n&=0,\\
a_2^{(1)}\xi_1+(a_2^{(2)}-\lambda)\,\xi_2+\ldots+a_2^{(n)}\xi_n&=0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots&\\
a_n^{(1)}\xi_1+a_n^{(2)}\xi_2+\ldots+(a_n^{(n)}-\lambda)\,\xi_n&=0.
\end{aligned}\right\}\quad(23)$$

Эта однородная система уравнений относительно величин $\xi_1, \xi_2, \ldots, \xi_n$ допускает ненулевое решение в том и только в том случае, когда ее определитель равен нулю (*3.21*):

$$\Delta(\lambda) \equiv \begin{vmatrix} a_1^{(1)}-\lambda & a_1^{(2)} & \dots & a_1^{(n)} \\ a_2^{(1)} & a_2^{(2)}-\lambda & \dots & a_2^{(n)} \\ \cdot & \cdot & \cdot & \cdot \\ a_n^{(1)} & a_n^{(2)} & \dots & a_n^{(n)}-\lambda \end{vmatrix} = 0. \qquad (24)$$

Многочлен n-й степени относительно λ, стоящий в левой части этого уравнения, называется *характеристическим многочленом* матрицы A. Всякому его корню $\lambda_0 \in K$ отвечает собственный вектор, который определяется после подстановки в (23) вместо λ величины λ_0 путем решения получившейся совместной системы относительно величин $\xi_1, \xi_2, \dots, \xi_n$.

Полученный результат показывает, между прочим, что хотя матрица оператора A зависит от выбора базиса $e_1, \dots, e_n$, но корни характеристического многочлена этой матрицы уже не зависят от выбора базиса. Мы еще вернемся к этому вопросу в *5.53*.

4.95. Разберем некоторые возможности, которые могут представиться при решении характеристического уравнения (24).

а. Случай отсутствия корней в поле K. Если уравнение $\Delta(\lambda)$ в поле K вовсе не имеет корней, то линейный оператор A не имеет в пространстве K_n собственных векторов.

Например, оператор поворота на угол $\varphi_0 \neq m\pi$ ($m=0, \pm 1, \pm 2, \dots$) на плоскости V_2, как мы уже отметили, не имеет собственных векторов. Этот факт, геометрически очевидный, легко устанавливается алгебраически. В самом деле, уравнение (24) для оператора поворота имеет вид

$$\begin{vmatrix} \cos\varphi_0-\lambda & -\sin\varphi_0 \\ \sin\varphi_0 & \cos\varphi_0-\lambda \end{vmatrix} = 0$$

или после раскрытия определителя

$$1-2\lambda\cos\varphi_0+\lambda^2=0,$$

и если $\varphi_0 \neq m\pi$ ($m=0, \pm 1, \pm 2, \dots$), то это уравнение не имеет вещественных корней.

б. Если $K=C$ есть поле комплексных чисел, то в силу основной теоремы алгебры уравнение (24) *всегда* имеет корень

$\lambda_0 \in K$. Таким образом, *в пространстве* C_n *всякий линейный оператор имеет (хотя бы один) собственный вектор.*

в. Случай наличия n различных корней. Если все n корней $\lambda_1, \lambda_2, \ldots, \lambda_n$ уравнения $\Delta(\lambda) = 0$ лежат в поле K и различны, то мы сможем в пространстве K_n найти n различных собственных векторов оператора A, решая систему (23) последовательно при $\lambda = \lambda_1, \lambda_2, \ldots, \lambda_n$. В силу леммы *4.93а* собственные векторы $f_1, f_2, \ldots, f_n$ будут линейно независимы. Примем их за новый базис и построим матрицу оператора A в этом новом базисе. Поскольку

$$\begin{aligned} Af_1 &= \lambda_1 f_1, \\ Af_2 &= \qquad \lambda_2 f_2, \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \\ Af_n &= \qquad\qquad \lambda_n f_n, \end{aligned}$$

матрица $A_{(f)}$ имеет вид

$$\left\| \begin{matrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{matrix} \right\|. \tag{25}$$

Используя определение диагонализируемого оператора (*4.72е*), мы можем сформулировать полученный результат следующим образом: *в пространстве* K_n *всякий линейный оператор, характеристический многочлен матрицы которого (в каком-либо базисе) имеет* n *различных корней в поле* K, *является диагонализируемым; матрица этого оператора, построенная в базисе из его собственных векторов, диагональна, и ее диагональные элементы суть собственные значения оператора.*

г. С другой стороны, если оператор A в некотором базисе $f_1, \ldots, f_n$ пространства K_n имеет диагональную матрицу (25) с произвольными, необязательно различными числами $\lambda_1, \ldots, \lambda_n$ на диагонали, то векторы $f_1, \ldots, f_n$ — собственные, а числа $\lambda_1, \ldots, \lambda_n$ — соответствующие собственные значения.

Покажем, что у оператора A в этом случае нет иных собственных значений (отличных от чисел $\lambda_1, \ldots, \lambda_n$). Действительно, если λ — собственное значение, отвечающее собственному вектору $f=\sum_{j=1}^{n}\beta_j f_j$, то, выделяя в равенстве

$$Af=A\left(\sum_{1}^{n}\beta_j f_j\right)=\sum_{1}^{n}\beta_j A f_j=\sum_{1}^{n}\beta_j\lambda_j f_j=$$

$$=\lambda f=\lambda\sum_{1}^{n}\beta_j f_j=\sum_{1}^{n}\lambda\beta_j f_j$$

слагаемое по вектору f_j, мы получаем

$$\lambda\beta_j=\lambda_j\beta_j \qquad (j=1, \ldots, n). \tag{26}$$

Среди чисел $\beta_1, \ldots, \beta_n$ есть хотя бы одно отличное от нуля; пусть, например, $\beta_1\neq 0$. Тогда равенство (26) при $j=1$ дает $\lambda=\lambda_1$, что и требуется.

д. Случай кратного корня. Пусть $\lambda=\lambda_0$ — некоторый корень уравнения (24) кратности $r\geqslant 1$. Возникает следующий вопрос: какова размерность соответствующего собственного подпространства $K^{(\lambda_0)}$, или, иными словами, сколько линейно независимых решений допускает система (23) при $\lambda=\lambda_0$? Зная ранг матрицы системы, мы бы могли точно ответить на этот вопрос (*3.51*). Но мы желаем связать этот ответ только с кратностью r корня λ_0.

В примерах *4.72а—в* и *е*, как легко убедиться, размерность каждого собственного подпространства $K^{(\lambda_0)}$ совпадает с кратностью соответствующего собственного значения λ_0 как корня характеристического многочлена оператора A. Однако в общем случае это не так. Рассмотрим оператор A в R_2, заданный матрицей

$$A=\begin{Vmatrix}\lambda_0 & 0\\ \mu & \lambda_0\end{Vmatrix}$$

с произвольным $\mu\neq 0$. Характеристический многочлен имеет вид $(\lambda_0-\lambda)^2$; он имеет двойной корень $\lambda=\lambda_0$. Система (23) в данном случае принимает вид

$$0\cdot\xi_1+0\cdot\xi_2=0,$$
$$\mu\cdot\xi_1+0\cdot\xi_2=0$$

и имеет единственное (с точностью до числового множителя) решение

$$\xi_1=0, \qquad \xi_2=1.$$

Таким образом, собственное подпространство оператора А, соответствующее собственному значению $\lambda = \lambda_0$, имеет размерность 1, меньшую, чем кратность корня λ_0. Можно доказать, что в общем случае размерность собственного подпространства $K^{(\lambda_0)}$ не превышает кратности корня λ_0 (см. задачу 11 к гл. 5). Полный ответ на вопрос о размерности пространства K^{λ} для случая $K = C$ мы дадим в гл. 6 в результате определения канонической формы матрицы оператора А.

ЗАДАЧИ

1. Определив естественным образом сложение линейных форм и умножение линейной формы на вещественное число, построить из линейных форм, определенных на линейном пространстве K, новое линейное пространство K*. Какова размерность пространства K*, если размерность пространства K равна n?

2. Выяснить, какие из следующих векторных функций в пространстве V_3 являются линейными операторами:

а) $Ax = x + a$ (a — фиксированный ненулевой вектор);

б) $Ax = a$;

в) $Ax = (a, x)\,a$ *);

г) $Ax = (a, x)\,x$ *);

д) $Ax = (\xi_1^2, \xi_2 + \xi_3, \xi_3^2)$, где $x = (\xi_1, \xi_2, \xi_3)$;

е) $Ax = (\sin \xi_1, \cos \xi_2, 0)$;

ж) $Ax = (2\xi_1 - \xi_2, \xi_2 + \xi_3, \xi_1)$.

3. Будут ли линейными операторами в пространстве всех многочленов от t

а) умножение на t,

б) умножение на t^2,

в) дифференцирование?

4. Составить матрицу оператора А в V_3, переводящего векторы

$$x_1 = (0, 0, 1) \text{ в } y_1 = (2, 3, 5),$$
$$x_2 = (0, 1, 1) \text{ в } y_2 = (1, 0, 0),$$
$$x_3 = (1, 1, 1) \text{ в } y_3 = (0, 1, -1)$$

в базисе

а) $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$;

б) x_1, x_2, x_3.

5. В трехмерном пространстве обозначим через А оператор поворота на 90° вокруг оси OX (от OY к OZ), через B — оператор поворота на 90° вокруг оси OY (от OZ к OX), через C — на 90° вокруг

*) Здесь (a, x) означает обычные скалярные произведения векторов a и x, т. е. число, равное произведению их длин и косинуса угла между ними.

оси OZ (от OX к OY). Показать, что $\mathrm{A}^4=\mathrm{B}^4=\mathrm{C}^4=\mathrm{E}$, $\mathrm{AB}\neq\mathrm{BA}$, $\mathrm{A}^2\mathrm{B}^2=\mathrm{B}^2\mathrm{A}^2$. Имеет ли место равенство $\mathrm{ABAB}=\mathrm{A}^2\mathrm{B}^2$?

6. В пространстве всех многочленов от t обозначим через A оператор дифференцирования, а через B — оператор умножения на независимое переменное:

$$\mathrm{A}P(t)=P'(t),\quad \mathrm{B}P(t)=tP(t).$$

Имеет ли место равенство $\mathrm{AB}=\mathrm{BA}$? Найти оператор $\mathrm{AB}-\mathrm{BA}$.

7. В предположении, что $\mathrm{AB}=\mathrm{BA}$, доказать формулы

$$(\mathrm{A}+\mathrm{B})^2=\mathrm{A}^2+2\mathrm{AB}+\mathrm{B}^2,$$
$$(\mathrm{A}+\mathrm{B})^3=\mathrm{A}^3+3\mathrm{A}^2\mathrm{B}+3\mathrm{AB}^2+\mathrm{B}^3.$$

Как следует изменить эти формулы, когда $\mathrm{AB}\neq\mathrm{BA}$?

8. В предположении, что $\mathrm{AB}-\mathrm{BA}=\mathrm{E}$, доказать формулу

$$\mathrm{A}^m\mathrm{B}-\mathrm{BA}^m=m\mathrm{A}^{m-1}\qquad(m=1,\ 2,\ \ldots).$$

9. Найти размерность линейного пространства $\mathrm{L}(\mathrm{K}_n,\ \mathrm{K}_m)$ всех линейных операторов, действующих из n-мерного пространства K_n в m-мерное пространство K_m, и построить базис пространства $\mathrm{L}(\mathrm{K}_n,\ \mathrm{K}_m)$.

10. Найти произведение матрицы A на матрицу B:

$$A=\begin{Vmatrix}1&2&3\\2&4&6\\3&6&9\end{Vmatrix},\qquad B=\begin{Vmatrix}-1&-2&-4\\-1&-2&-4\\1&2&4\end{Vmatrix}.$$

11. Выполнить действие возведения в n-ю степень матриц

$$A=\begin{Vmatrix}1&1\\0&1\end{Vmatrix}\ \text{и}\ B=\begin{Vmatrix}\cos\varphi&-\sin\varphi\\\sin\varphi&\cos\varphi\end{Vmatrix}.$$

12. Найти все матрицы A 2-го порядка, удовлетворяющие условию

$$A^2=\begin{Vmatrix}0&0\\0&0\end{Vmatrix}.$$

13. Вычислить $AB-BA$, где

а) $$A=\begin{Vmatrix}1&2&2\\2&1&2\\1&2&3\end{Vmatrix},\qquad B=\begin{Vmatrix}4&1&1\\-4&2&0\\1&2&1\end{Vmatrix},$$

б) $$A=\begin{Vmatrix}2&1&0\\1&1&2\\-1&2&1\end{Vmatrix},\qquad B=\begin{Vmatrix}3&1&2\\3&-2&4\\-3&5&-1\end{Vmatrix}.$$

14. Сумма диагональных элементов $a_{11}+\ldots+a_{nn}$ матрицы $A=\|a_{jk}\|$ называется *следом* матрицы A и обозначается $\mathrm{sp}\,A$. Проверить формулы $\mathrm{sp}\,(A+B)=\mathrm{sp}\,A+\mathrm{sp}\,B$, $\mathrm{sp}\,(AB)=\mathrm{sp}\,(BA)$.

15. Используя задачу 14, доказать, что для операторов A и B, действующих в пространстве K_n, равенство $\mathrm{AB}-\mathrm{BA}=\mathrm{E}$ невозможно.

Примечание. Результат задачи 6 показывает, что в рассматриваемом случае предположение о конечномерности пространства K_n существенно.

16. Для данной квадратной матрицы C 2-го порядка с $\operatorname{sp} C = 0$ (см. задачу 14) найти представление в форме

$$C = AB - BA,$$

где A и B — (неизвестные) матрицы 2-го порядка.

17. Пусть в n-мерном пространстве даны m линейно независимых векторов

$$x_j = \sum_{i=1}^{n} \xi_i^{(j)} e_i \qquad (j = 1, 2, \ldots, m)$$

и оператор A действует в линейной оболочке $L(x_1, \ldots, x_m)$ по формулам

$$y_j = \mathrm{A} x_j = \sum_{k=1}^{m} a_k^{(j)} x_k \qquad (j = 1, 2, \ldots, m).$$

Показать, что каждый минор m-го порядка в матрице из координат векторов y_j (относительно базиса $e_1, e_2, \ldots, e_n$) равен произведению соответствующего минора в матрице из координат векторов x_j на $\det \| a_k^{(j)} \|$.

18. Доказать, что если в матрице A ранга r базисный минор расположен в левом верхнем углу, то отношение любого минора M r-го порядка к минору, находящемуся в тех же столбцах, что и минор M, но в первых r строках, зависит только от номеров столбцов минора M.

19. Доказать, что если A — матрица ранга r, то любой определитель 2-го порядка, составленный из миноров r-го порядка матрицы A, вида

$$\begin{vmatrix} M_{i_1, i_2, \ldots, i_r}^{i_1, i_2, \ldots, i_r} & M_{k_1, k_2, \ldots, k_r}^{i_1, i_2, \ldots, i_r} \\ M_{i_1, i_2, \ldots, i_r}^{k_1, k_2, \ldots, k_r} & M_{k_1, k_2, \ldots, k_r}^{k_1, k_2, \ldots, k_r} \end{vmatrix}$$

равен нулю.

20. Доказать, что каждый минор k-го порядка матрицы ABC равен сумме произведений некоторых миноров k-го порядка матриц A, B и C.

21. Для матриц

$$A = \begin{Vmatrix} 1 & 2 \\ 2 & 5 \end{Vmatrix}, \quad B = \begin{Vmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{Vmatrix},$$

$$C = \begin{Vmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{Vmatrix}$$

найти обратные матрицы.

22. Для любой невырожденнной матрицы A доказать равенство

$$(A')^{-1}=(A^{-1})'.$$

23. Найти все решения уравнения $XA=0$, где A — данная матрица 2-го порядка, X — искомая матрица 2-го порядка, 0 — нулевая матрица (матрица, все элементы которой равны нулю).

24. Пусть $A=\|a_i^{(j)}\|$ — любая квадратная матрица n-го порядка и $A_i^{(j)}$ — алгебраическое дополнение элемента $a_i^{(j)}$ в определителе матрицы A. Матрица $\tilde{A}=\|A_j^{(i)}\|$ называется *присоединенной* для матрицы A. Показать, что $\tilde{A}A=A\tilde{A}=(\det A)\cdot E$.

25. Пусть P — линейное пространство всех многочленов от аргумента t. Рассмотрим операторы A и B, определенные условиями

$$\mathrm{A}[a_0+a_1t+\ldots+a_nt^n]=a_1+a_2t+\ldots+a_nt^{n-1},$$
$$\mathrm{B}[a_0+a_1t+\ldots+a_nt^n]=a_0t+a_1t^2+\ldots+a_nt^{n+1}.$$

Показать, что A и B — линейные операторы и что

$$\mathrm{AB}=\mathrm{E},$$
$$\mathrm{BA}\neq\mathrm{E}.$$

Обладает ли оператор A обратным?

26. Показать, что оператор B в задаче 25 имеет бесконечное число левых обратных.

27. Если A — невырожденный линейный оператор в n-мерном линейном пространстве, то всякое подпространство, инвариантное относительно A, будет инвариантным и относительно A^{-1}.

28. Если линейные операторы A и B перестановочны (т. е. $\mathrm{AB}=\mathrm{BA}$), то всякое собственное подпространство оператора B является инвариантным подпространством для оператора A.

29. Если прямая сумма (*2.45*) собственных подпространств оператора A совпадает со всем пространством K и каждое собственное подпространство оператора A является инвариантным для оператора B, то A и B перестановочны.

30. Если x и y — собственные векторы оператора A с различными собственными значениями, то $\alpha x+\beta y$ ($\alpha\neq0$, $\beta\neq0$) заведомо не является собственным вектором оператора A.

31. Если каждый вектор пространства K является собственным вектором для оператора A, то $\mathrm{A}=\lambda\mathrm{E}$ ($\lambda\in K$).

32. Если линейный оператор A перестановочен со всеми линейными операторами, действующими в данном пространстве, то $\mathrm{A}=\lambda\mathrm{E}$.

33. Если линейный оператор A имеет собственный вектор e_0 с собственным значением λ_0, то для оператора A^2 вектор e_0 также является собственным с собственным значением, равным λ_0^2.

34. Если линейный оператор A не имеет собственных векторов, то оператор A^2 может их иметь (пример: оператор поворота на 90° в плоскости). Показать, что если оператор A^2 в пространстве R_n имеет собственный вектор с неотрицательным собственным значением $\lambda=\mu^2$, то оператор A также имеет собственный вектор.

35. Найти собственные значения и собственные векторы операторов, заданных следующими матрицами:

$$\text{а)}\ \begin{Vmatrix} 2 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 2 & 1 \end{Vmatrix};\quad \text{б)}\ \begin{Vmatrix} -1 & -2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix};$$

$$\text{в)}\ \begin{Vmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{Vmatrix};\quad \text{г)}\ \begin{Vmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

36. Проверить выполнение следующих фактов:

а) соотношение $N(A) \supset T(A)$ необходимо и достаточно для того, чтобы имело место равенство $A^2 = 0$;

б) для любого оператора A имеем $N(A) \subset N(A^2) \subset N(A^3) \subset \ldots$;

в) для любого оператора A имеем $T(A) \supset T(A^2) \supset T(A^3) \supset \ldots$;

г) если $T(A^k) \subset N(A^m)$, то

$$T(A) \subset N(A^{m+k-1}),\quad T(A^{m+k-1}) \subset N(A).$$

37. Показать, что каждый линейный оператор A ранга r может быть представлен в виде суммы r линейных операторов ранга 1.

38. Найти все инвариантные подпространства диагонального оператора с различными диагональными элементами и доказать, что их число равно 2^n.

ПРЕОБРАЗОВАНИЯ КООРДИНАТ

Известно, какую большую роль при решении геометрических задач средствами аналитической геометрии играет правильный выбор системы координат. В значительно более обширном круге вопросов, которые открываются в связи с геометрией n-мерного линейного пространства, роль правильного выбора системы координат будет также весьма велика. Эта глава посвящена правилам, по которым совершается преобразование координат в n-мерном линейном пространстве. Результаты, полученные здесь, послужат, в частности, основой для классификации квадратичных форм, которая будет проведена в седьмой главе.

§ 5.1. Формулы перехода к новому базису

5.11. Пусть $\{e\}=\{e_1, e_2, \ldots, e_n\}$ — некоторый базис в n-мерном пространстве K_n и $\{f\}=\{f_1, f_2, \ldots, f_n\}$ — некоторый другой базис в том же пространстве. Векторы системы $\{f\}$ однозначно определяются своими разложениями по векторам исходного базиса:

$$\left.\begin{aligned} f_1 &= p_1^{(1)}e_1+p_2^{(1)}e_2+\ldots+p_n^{(1)}e_n, \\ f_2 &= p_1^{(2)}e_1+p_2^{(2)}e_2+\ldots+p_n^{(2)}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n &= p_1^{(n)}e_1+p_2^{(n)}e_2+\ldots+p_n^{(n)}e_n \end{aligned}\right\} \qquad (1)$$

или, короче,

$$f_j=\sum_{i=1}^{n} p_i^{(j)}e_i \quad (j=1, 2, \ldots, n). \qquad (2)$$

В формулах (1) или (2) коэффициенты $p_i^{(j)}$ $(i, j=1, 2, \ldots, n)$ определяют матрицу

$$P=\|p_i^{(j)}\|=\begin{Vmatrix} p_1^{(1)} & p_1^{(2)} & \ldots & p_1^{(n)} \\ p_2^{(1)} & p_2^{(2)} & \ldots & p_2^{(n)} \\ \cdot & \cdot & \cdot & \cdot \\ p_n^{(1)} & p_n^{(2)} & \ldots & p_n^{(n)} \end{Vmatrix},$$

которая называется *матрицей перехода от базиса* $\{e\}$ *к базису* $\{f\}$. Как и ранее в аналогичных случаях (§ 4.2 и далее), мы выписываем координаты векторов f_j (относительно базиса $\{e\}$) в столбцах матрицы P.

Определитель D матрицы P отличен от нуля; действительно, в противном случае ее столбцы, а с ними и векторы $f_1, f_2, \ldots, f_n$ были бы линейно зависимыми (*3.12a*). Матрицу с определителем, отличным от нуля, мы уже назвали ранее невырожденной. Таким образом, *переход от одного базиса n-мерного пространства* K_n *к другому базису всегда осуществляется с помощью некоторой невырожденной матрицы.*

Формулы (1) вместе с матрицей P задают и соответствующий линейный оператор P, определяемый из соотношений $f_i = \mathrm{P}e_i$ $(i = 1, 2, \ldots, n)$. Он также называется *оператором перехода от базиса* $\{e\}$ *к базису* $\{f\}$.

5.12. Обратно, пусть $\{e\} = \{e_1, e_2, \ldots, e_n\}$ — заданный базис n-мерного пространства K_n и $P = \|p_i^{(j)}\|$ — невырожденная матрица порядка n. Построим по формулам (1) систему векторов $f_1, f_2, \ldots, f_n$. Очевидно, что эти векторы линейно независимы, поскольку столбцы всякой невырожденной матрицы линейно независимы (*3.12a*). Следовательно, векторы $f_1, f_2, \ldots, f_n$ образуют новый базис пространства K_n. Итак, *всякая невырожденная матрица* $P = \|p_i^{(j)}\|$ *определяет по формулам* (1) *переход от одного базиса n-мерного пространства* K_n *к другому базису.*

5.13. Отметим один частный случай перехода к новому базису: именно, тот, когда каждый из векторов f_k совпадает с соответствующим вектором e_k, умноженным на некоторое число $\lambda_k \neq 0$ $(k = 1, 2, \ldots, n)$. Формулы (1) принимают вид

$$\begin{aligned}
f_1 &= \lambda_1 e_1, \\
f_2 &= \qquad \lambda_2 e_2, \\
&\cdots\cdots\cdots \\
f_n &= \qquad\qquad \lambda_n e_n,
\end{aligned}$$

и матрица P имеет диагональную форму

$$P=\begin{Vmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_n \end{Vmatrix}.$$

В частности, при $\lambda_1=\lambda_2=\ldots=\lambda_n=1$ получаем матрицу *тождественного преобразования* — единичную матрицу

$$E=\begin{Vmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{Vmatrix}; \tag{3}$$

при тождественном преобразовании исходный базис не изменяется.

§ 5.2. Последовательные преобразования

5.21. Пусть $P=\|p_i^{(j)}\|$ — матрица перехода от базиса $\{e\}=\{e_1, e_2, \ldots, e_n\}$ к базису $\{f\}=\{f_1, f_2, \ldots, f_n\}$ и $Q=\|q_j^{(k)}\|$ — матрица перехода от базиса $\{f\}$ к базису $\{g\}=\{g_1, g_2, \ldots, g_n\}$. Определим матрицу перехода от базиса $\{e\}$ непосредственно к базису $\{g\}$. Формула перехода от базиса $\{e\}$ к базису $\{f\}$ имеет вид (2)

$$f_j=\sum_{i=1}^{n} p_i^{(j)}e_i \qquad (j=1, 2, \ldots, n) \tag{4}$$

и от базиса $\{f\}$ к базису $\{g\}$ соответственно

$$g_k=\sum_{j=1}^{n} q_j^{(k)}f_j \qquad (k=1, 2, \ldots, n). \tag{5}$$

Подставляя (4) в (5), получаем

$$g_k=\sum_{j=1}^{n} q_j^{(k)} \sum_{i=1}^{n} p_i^{(j)}e_i=\sum_{i=1}^{n}\left(\sum_{j=1}^{n} p_i^{(j)}q_j^{(k)}\right)e_i \tag{6}$$
$$(k=1, 2, \ldots, n).$$

С другой стороны, если $T=\| t_i^{(k)} \|$ означает искомую матрицу перехода от базиса $\{e\}$ к базису $\{g\}$, то мы можем написать

$$g_k=\sum_{i=1}^{n} t_i^{(k)} e_i \qquad (k=1,\ 2,\ \ldots,\ n). \tag{7}$$

Сравнивая (6) и (7), мы получаем

$$t_i^{(k)}=\sum_{j=1}^{n} p_i^{(j)} q_j^{(k)} \qquad (i,\ k=1,\ 2,\ \ldots,\ n). \tag{8}$$

Полученная формула (8) только обозначениями индексов (но не их ролью) отличается от формулы *4.43* (8). Следовательно, *искомая матрица T есть произведение матрицы P на матрицу Q.*

5.22. Рассмотрим следующий частный случай последовательных преобразований. Систему уравнений (1), поскольку матрица P не выраждена, можно разрешить относительно векторов $e_1,\ \ldots,\ e_n$; получающаяся при этом система равенств вида

$$\left.\begin{aligned} e_1&=q_1^{(1)}f_1+q_2^{(1)}f_2+\ldots+q_n^{(1)}f_n,\\ e_2&=q_1^{(2)}f_1+q_2^{(2)}f_2+\ldots+q_n^{(2)}f_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ e_n&=q_1^{(n)}f_1+q_2^{(n)}f_2+\ldots+q_n^{(n)}f_n \end{aligned}\right\} \tag{9}$$

определяет, очевидно, переход от базиса $\{f\}$ к базису $\{e\}$.

Последовательный переход от базиса $\{e\}$ к базису $\{f\}$ с помощью матрицы P и затем от базиса $\{f\}$ к базису $\{e\}$ с помощью матрицы $Q=\| q_k^{(j)} \|$ есть переход от базиса $\{e\}$ к нему же самому, т. е. в конечном счете тождественное преобразование с единичной матрицей; поэтому здесь $PQ=E$.

§ 5.3. Преобразование координат вектора при изменении базиса

5.31. Пусть $\{e\}=\{e_1,\ e_2,\ \ldots,\ e_n\}$ и $\{f\}=\{f_1,\ f_2,\ \ldots,\ f_n\}$ — два базиса в n-мерном линейном пространстве K_n. Для любого вектора $x\in\mathrm{K}_n$ имеют место разложения

$$x=\xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n=\eta_1f_1+\eta_2f_2+\ldots+\eta_nf_n, \tag{10}$$

где $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x относительно базиса $\{e\}$ и $\eta_1, \eta_2, \ldots, \eta_n$ — его координаты относительно базиса $\{f\}$. Поставим задачу вычислить координаты вектора x относительно базиса $\{f\}$ по известным его координатам относительно базиса $\{e\}$. Пусть нам дана матрица $P = \| p_i^{(j)} \|$ перехода от базиса $\{e\}$ к базису $\{f\}$. Тогда векторы $\{e\}$ выражаются через векторы $\{f\}$ по формулам (9) или, короче,

$$e_j = \sum_{k=1}^{n} q_k^{(j)} f_k \qquad (j = 1, 2, \ldots, n), \tag{11}$$

где $Q = \| q_k^{(j)} \|$ — матрица, обратная к матрице P. Подставляя формулы (11) в разложение (10), мы получаем

$$x = \sum_{j=1}^{n} \xi_j e_j = \sum_{k=1}^{n} \eta_k f_k =$$
$$= \sum_{j=1}^{n} \xi_j \left(\sum_{k=1}^{n} q_k^{(j)} f_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} q_k^{(j)} \xi_j \right) f_k,$$

откуда в силу единственности разложения вектора x по базису $\{f\}$

$$\eta_k = \sum_{j=1}^{n} q_k^{(j)} \xi_j \qquad (k = 1, 2, \ldots, n). \tag{12}$$

В раскрытом виде получается система равенств

$$\begin{aligned} \eta_1 &= q_1^{(1)} \xi_1 + q_1^{(2)} \xi_2 + \ldots + q_1^{(n)} \xi_n, \\ \eta_2 &= q_2^{(1)} \xi_1 + q_2^{(2)} \xi_2 + \ldots + q_2^{(n)} \xi_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \eta_n &= q_n^{(1)} \xi_1 + q_n^{(2)} \xi_2 + \ldots + q_n^{(n)} \xi_n. \end{aligned}$$

Таким образом, *координаты вектора x относительно базиса $\{f\}$ линейно выражаются через координаты вектора x относительно базиса $\{e\}$; коэффициенты этих линейных выражений образуют матрицу, транспонированную по отношению к матрице перехода от базиса $\{f\}$ к базису $\{e\}$* (т. е. *транспонированную по отношению к матрице, обратной к матрице P*).

Используя обозначение обратной матрицы P^{-1} и транспонированной P', можно записать матрицу S, определяемую соотношениями (12), формулой $S = (P^{-1})'$.

5.32. Имеет место и обратная теорема:

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — *координаты произвольного вектора* x *относительно базиса* $\{e\} = \{e_1, e_2, \ldots, e_n\}$ *n-мерного пространства* K_n *и величины* $\eta_1, \eta_2, \ldots, \eta_n$ *определены посредством равенств*

$$\begin{aligned} \eta_1 &= s_{11}\xi_1 + s_{12}\xi_2 + \ldots + s_{1n}\xi_n, \\ \eta_2 &= s_{21}\xi_1 + s_{22}\xi_2 + \ldots + s_{2n}\xi_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \eta_n &= s_{n1}\xi_1 + s_{n2}\xi_2 + \ldots + s_{nn}\xi_n, \end{aligned}$$

где $\det \| s_{jk} \| \neq 0$. *Тогда в пространстве* K_n *можно найти новый базис* $\{f\} = \{f_1, f_2, \ldots, f_n\}$ *таким образом, чтобы числа* $\eta_1, \eta_2, \ldots, \eta_n$ *стали координатами вектора* x *относительно базиса* $\{f\}$.

Доказательство. Введем матрицу $S = \| s_{jk} \|$ и матрицу $P = (S')^{-1}$, элементы которой обозначим через $p_i^{(j)}$. С помощью матрицы P построим новый базис по формулам (1). Утверждается, что этот базис и есть искомый. Действительно, составим формулы перехода (12) к координатам вектора x относительно нового базиса. Как мы видели, эти формулы записываются с помощью матрицы $(P^{-1})'$. В данном случае эта матрица совпадает с матрицей S, так как

$$(P^{-1})' = ([(S')^{-1}]^{-1})' = (S')' = S.$$

Следовательно, величины $\eta_1, \eta_2, \ldots, \eta_n$ и координаты вектора x относительно базиса $\{f\}$ — одни и те же для любого вектора x, что и требовалось.

5.33. Аналогично *5.21* можно построить матрицу последовательного преобразования координат. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x относительно базиса $\{e\}$ и величины $\eta_1, \eta_2, \ldots, \eta_n$ и $\tau_1, \tau_2, \ldots, \tau_n$ определены посредством равенств

$$\eta_j = \sum_{i=1}^{n} p_{ji}\xi_i \qquad (j = 1, 2, \ldots, n),$$

$$\tau_k = \sum_{j=1}^{n} q_{kj}\eta_j \qquad (k = 1, 2, \ldots, n),$$

соответственно с невырожденными матрицами $P=\|p_{ji}\|$ и $Q=\|q_{kj}\|$. Тогда можно выразить величины $\{\tau\}$ непосредственно через величины $\{\xi\}$ по формулам

$$\tau_k=\sum_{i=1}^{n}\sum_{j=1}^{n} q_{kj}p_{ji}\,\xi_i=\sum_{i=1}^{n} t_{ki}\xi_i \qquad (k=1,\ 2,\ \ldots,\ n),$$

где величины t_{ki} $(i,\ k=1,\ 2,\ \ldots,\ n)$ образуют матрицу T, равную произведению матрицы Q на матрицу P.

§ 5.4. Преобразование коэффициентов линейной формы

Пусть в пространстве K_n задана линейная форма $L(x)$. Как мы видели в *4.13*, когда в пространстве K_n выбран базис $\{e\}=\{e_1,\ e_2,\ \ldots,\ e_n\}$, значения формы $L(x)$ можно вычислять по формуле

$$L(x)=\sum_{k=1}^{n} l_k\xi_k,$$

где ξ_k $(k=1,\ 2,\ \ldots,\ n)$ — координаты вектора x относительно базиса $\{e\}$, а $l_k=L(e_k)$ $(n=1,\ 2,\ \ldots,\ n)$. Коэффициенты l_k, очевидно, зависят от выбора базиса $\{e\}$. Мы выведем здесь правило, по которому совершается преобразование коэффициентов линейной формы при переходе к новому базису.

Допустим, что формулы

$$f_j=\sum_{i=1}^{n} p_i^{(j)}e_i \qquad (j=1,\ 2,\ \ldots,\ n) \tag{13}$$

определяют переход от базиса $\{e\}$ к новому базису $\{f\}$. Найдем коэффициенты линейной формы $L(x)$ в базисе $\{f\}$. Эти коэффициенты суть числа $\lambda_j=L(f_j)$; вычисляя эти числа с помощью формулы (13), находим

$$\lambda_j=L(f_j)=\sum_{i=1}^{n} p_i^{(j)}L(e_i)=\sum_{i=1}^{n} p_i^{(j)}l_i.$$

Таким образом, *коэффициенты линейной формы преобразуются так же, как преобразуются базисные векторы.*

§ 5.5. Преобразование матрицы линейного оператора

5.51. Пусть дан линейный оператор A в n-мерном пространстве K_n. Обозначим через $A_{(e)}=\|a_i^{(j)}\|$ матрицу оператора A относительно базиса $\{e\}=\{e_1, e_2, \ldots, e_n\}$ и через $A_{(f)}=\|\alpha_q^{(p)}\|$ его матрицу относительно базиса $\{f\}=\{f_1, f_2, \ldots, f_n\}$. Предположим, далее, что формулы перехода от базиса $\{e\}$ к базису $\{f\}$ имеют вид

$$f_k=\sum_{j=1}^{n} p_j^{(k)} e_j \qquad (k=1, 2, \ldots, n). \tag{14}$$

Матрицу перехода $\|p_j^{(k)}\|$ обозначим через P. Установим связь между матрицами $A_{(e)}$, $A_{(f)}$ и P. Матрица $A_{(e)}=\|a_i^{(j)}\|$ определяется из системы равенств

$$\mathrm{A}e_j=\sum_{i=1}^{n} a_i^{(j)} e_i \qquad (j=1, 2, \ldots, n), \tag{15}$$

а матрица $A_{(f)}=\|\alpha_k^{(m)}\|$ — из системы равенств

$$\mathrm{A}f_m=\sum_{k=1}^{n} \alpha_k^{(m)} f_k \qquad (m=1, 2, \ldots, n).$$

Заменим в последней формуле векторы f_k их выражениями через векторы e_j по формулам (14), изменив индекс суммирования j на i:

$$\mathrm{A}f_m=\sum_{k=1}^{n} \alpha_k^{(m)} \sum_{i=1}^{n} p_i^{(k)} e_i=\sum_{i=1}^{n}\left(\sum_{k=1}^{n} p_i^{(k)} \alpha_k^{(m)}\right) e_i.$$

Теперь применим оператор A к обеим частям формулы (14) и используем выражение векторов $\mathrm{A}e_j$ из (15):

$$\mathrm{A}f_m=\mathrm{A}\sum_{j=1}^{n} p_j^{(m)} e_j=\sum_{j=1}^{n} p_j^{(m)} \mathrm{A}e_j=\sum_{j=1}^{n} p_j^{(m)} \sum_{i=1}^{n} a_i^{(j)} e_i=$$
$$=\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_i^{(j)} p_j^{(m)}\right) e_i.$$

Сравнивая коэффициенты при e_i в двух последних разложениях, находим

$$\sum_{k=1}^{n} p_i^{(k)}\alpha_k^{(m)} = \sum_{j=1}^{n} a_i^{(j)}p_j^{(m)}$$

или в матричной форме

$$PA_{(f)} = A_{(e)}P. \tag{16}$$

Это и есть искомая связь между матрицами $A_{(e)}$, $A_{(f)}$ и P. Умножая обе части слева на матрицу P^{-1}, получаем выражение матрицы $A_{(f)}$:

$$A_{(f)} = P^{-1}A_{(e)}P.$$

5.52. Используя теорему о произведении определителей (*4.75*), получаем из (16) следующее соотношение:

$$\det P \det A_{(f)} = \det A_{(e)} \det P$$

или, так как $\det P \neq 0$,

$$\det A_{(e)} = \det A_{(f)}.$$

Итак, *определитель матрицы оператора не зависит от выбора базиса в пространстве*. Поэтому мы можем говорить об определителе оператора, подразумевая под ним определитель матрицы этого оператора в произвольном базисе.

5.53. Кроме определителя, существуют и другие функции от элементов матрицы оператора, остающиеся неизменными при переходе к новому базису. Чтобы построить такие функции, рассмотрим оператор $\mathrm{A}-\lambda\mathrm{E}$, где λ — параметр, взятый из поля K. Матрицей этого оператора в базисе $\{e\}$ является, очевидно, матрица $A_{(e)}-\lambda E$, а в базисе $\{f\}$ — матрица $A_{(f)}-\lambda E$. По доказанному, при любом λ

$$\det(A_{(e)}-\lambda E) = \det(A_{(f)}-\lambda E).$$

Справа и слева стоят многочлены n-й степени от λ. Так как эти многочлены совпадают тождественно, то у них коэффициенты при любой степени λ должны быть одинаковыми. Эти коэффициенты суть некоторые функции от элементов матрицы оператора, которые, следовательно, остаются неизменными при изменении базиса. Выясним вид этих функций. Определитель матрицы $A_{(e)} - \lambda E$ имеет вид

$$\begin{vmatrix} a_1^{(1)} - \lambda & a_1^{(2)} & \ldots & a_1^{(n)} \\ a_2^{(1)} & a_2^{(2)} - \lambda & \ldots & a_2^{(n)} \\ \cdot & \cdot & \cdot & \cdot \\ a_n^{(1)} & a_n^{(2)} & \ldots & a_n^{(n)} - \lambda \end{vmatrix} =$$

$$= (-1)^n \lambda^n + \Delta_1 \lambda^{n-1} + \ldots + \Delta_{n-1} \lambda + \Delta_n.$$

Коэффициент Δ_1 при λ^{n-1}, как легко вывести из определения определителя, равен (со знаком $(-1)^{n-1}$) сумме диагональных элементов $a_1^{(1)} + a_2^{(2)} + \ldots + a_n^{(n)}$; это число называется *следом* оператора A. Коэффициент Δ_2 при λ^{n-2} есть сумма всех диагональных миноров 2-го порядка*), взятая со знаком $(-1)^{n-2}$; аналогично коэффициент Δ_k при λ^{n-k} есть сумма всех диагональных миноров k-го порядка, взятая со знаком $(-1)^{n-k}$. Наконец, коэффициент Δ_n при λ^0 — свободный член — равен, очевидно, самому определителю оператора. Многочлен $\det(A_{(e)} - \lambda E)$, который, как мы доказали, не зависит от выбора базиса в пространстве, называется *характеристическим многочленом оператора* A.

§ 5.6*. Тензоры

5.61. Координаты вектора, коэффициенты линейной формы, элементы матрицы линейного оператора являются примерами геометрических величин, называемых *тензорами*.

Прежде чем перейти к соответствующему определению, несколько рационализируем нашу систему обозначений.

Векторы базиса n-мерного пространства K_n будем обозначать, как и раньше, символами $e_1, e_2, \ldots, e_n$ (с индексами внизу).

Координаты векторов $x, y, \ldots$ будем обозначать соответственно символами $\xi^1, \xi^2, \ldots, \xi^n$; $\eta^1, \eta^2, \ldots, \eta^n, \ldots$ (с индексами вверху).

Коэффициенты линейной формы $L(x)$ обозначим $l_1, l_2, \ldots, l_n$ (с индексами внизу).

Элементы матрицы линейного оператора обозначим через a_i^j; при этом верхний индекс обозначает номер строки,

*) Минор $M_{i_1, i_2, \ldots, i_k}^{j_1, j_2, \ldots, j_k}$ называется *диагональным*, если $i_1 = j_1$, $i_2 = j_2, \ldots, i_k = j_k$.

нижний — номер столбца (в отличие от обозначений, принятых в *4.23*).

Целесообразность такого расположения индексов определяется следующим соглашением о суммировании: если имеется сумма n одночленных выражений, причем индекс суммирования i встречается в общем члене суммы два раза — один раз наверху, другой раз внизу, — знак суммы мы будем опускать.

Например, разложение вектора x по базису $\{e_1, e_2, \ldots, e_n\}$ после нашего соглашения приобретает вид

$$x = \xi^i e_i$$

(знак суммирования по i опущен, но подразумевается). Выражение линейной формы $f(x)$ через координаты вектора и коэффициенты формы принимает такой вид:

$$f(x) = l_i \xi^i.$$

Результат применения оператора A к базисному вектору e_i принимает следующий вид:

$$\mathrm{A} e_i = a_i^j e_j$$

(подразумевается суммирование по индексу j). Координаты η^j вектора $\mathrm{A}x$ выражаются через координаты вектора x следующим образом:

$$\eta^j = a_i^j \xi^i$$

(подразумевается суммирование по индексу i).

Величины, относящиеся к новой системе координат, мы будем обозначать теми же символами, но со штрихами при индексах. Так, новые базисные векторы мы будем обозначать через $e_{1'}, e_{2'}, \ldots, e_{n'}$, новые координаты вектора x — через $\xi^{1'}, \xi^{2'}, \ldots, \xi^{n'}$ и т. д.

Элементы матрицы перехода от базиса e_i к базису $e_{i'}$ обозначим через $p_{i'}^i$, так что

$$e_{i'} = p_{i'}^i e_i \tag{17}$$

(суммирование по индексу i).

Коэффициенты матрицы обратного перехода обозначим через $q_i^{i'}$:

$$e_i = q_i^{i'} e_{i'} \tag{18}$$

(суммирование по i'). Матрица $q_i^{i'}$ обратна к матрице $p_{i'}^i$, что можно записать в виде равенства

$$p_{i'}^i q_j^{i'} = \begin{cases} 0 & \text{при } i \neq j, \\ 1 & \text{при } i = j \end{cases} \tag{19}$$

или равенства

$$p_{i'}^i q_i^{j'} = \begin{cases} 0 & \text{при } i' \neq j', \\ 1 & \text{при } i' = j'. \end{cases} \tag{20}$$

Для сокращения записи величину, зависящую от индексов i и j так, что она равна нулю при различных значениях индексов и единице при совпадающих значениях индексов, обозначают δ_j^i; таким образом, равенство (19) можно записать в виде

$$p_{i'}^i q_j^{i'} = \delta_j^i, \tag{21}$$

а равенство (20) — в виде

$$p_{i'}^i q_i^{j'} = \delta_{i'}^{j'}. \tag{22}$$

5.62. Чтобы показать преимущества пользования новыми обозначениями, выведем заново формулы преобразования координат вектора, коэффициентов линейной формы и элементов матрицы оператора при переходе к новому базису.

Пусть $x = \xi^i e_i = \xi^{i'} e_{i'}$. Подставляя вместо величин e_i равные им величины $q_i^{i'} e_{i'}$ (18), получаем

$$x = \xi^i q_i^{i'} e_{i'} = \xi^{i'} e_{i'}.$$

Поскольку $e_{i'}$ — базис,

$$\xi^{i'} = q_i^{i'} \xi^i. \tag{23}$$

Это и есть формула преобразования координат вектора.

Пусть дана линейная форма $L(x)$. Числа $l_{i'}$ определяются, как обычно, равенствами $l_{i'} = L(e_{i'})$. Подставляя вместо $e_{i'}$ выражение $p_{i'}^i e_i$ (17), получаем

$$l_{i'} = L(p_{i'}^i e_i) = p_{i'}^i L(e_i) = p_{i'}^i l_i.$$

Итак,

$$l_{i'} = p_{i'}^i l_i; \tag{24}$$

это и есть интересующая нас формула.

Наконец, пусть дан оператор A. Элементы его матрицы в новом базисе определяются из равенств

$$\mathrm{A}e_{i'} = a_{i'}^{j'} e_{j'}.$$

Подставив сюда вместо $e_{i'}$ $(e_{j'})$ равные им величины $p_{i'}^{i} e_i$ $(p_{j'}^{j} e_j)$ (17), получим

$$p_{i'}^{i} \mathrm{A} e_i = a_{i'}^{j'} p_{j'}^{j} e_j.$$

Но $\mathrm{A}e_i = a_i^j e_j$, так что в результате

$$p_{i'}^{i} a_i^j e_j = a_{i'}^{j'} p_{j'}^{j} e_j.$$

Поскольку e_j — базисные векторы,

$$p_{i'}^{i} a_i^j = a_{i'}^{j'} p_{j'}^{j}.$$

Чтобы получить отсюда $a_{i'}^{j'}$, умножим обе части равенства на $q_j^{k'}$ и по индексу j произведем суммирование. В силу формулы (22) мы получим

$$p_{i'}^{i} a_i^j q_j^{k'} = a_{i'}^{j'} p_{j'}^{j} q_j^{k'} = a_{i'}^{j'} \delta_{j'}^{k'}.$$

Согласно определению величины $\delta_{j'}^{k'}$ при суммировании по j' нужно учесть только одно слагаемое, отвечающее значению $j' = k'$. При этом $\delta_{k'}^{k'} = 1$ и, следовательно,

$$a_{i'}^{k'} = p_{i'}^{i} q_j^{k'} a_i^j. \tag{25}$$

Это и есть искомая формула.

Нетрудно проверить, что все три полученные нами теперь формулы преобразования совпадают с формулами, полученными нами ранее обычным путем (§§ 5.3, 5.4, 5.5). Формулы (23,) (24) и (25) имеют много общего. Прежде всего, эти формулы линейны относительно преобразуемых величин. Далее, коэффициентами этих формул служат или элементы матрицы перехода от старого базиса к новому, или элементы матрицы обратного перехода, или, наконец, и те и другие.

5.63. Теперь можно перейти к определению тензора. Тензоры разделяются на *ковариантные*, *контравариантные* и *смешанные*. Кроме того, каждый тензор имеет определенный *ранг*.

Начнем с определения ковариантного тензора, для определенности, третьего ранга.

Пусть имеется правило, позволяющее в каждой системе координат n-мерного пространства K_n построить n^3 чисел T_{ijk} (составляющих), каждое из которых определяется при придании индексам i, j, k определенных значений от 1 до n. Эти числа T_{ijk} образуют по определению *ковариантный тензор третьего ранга*, если преобразование величин T_{ijk} при переходе к новому базису производится по формуле

$$T_{i'j'k'} = p_{i'}^{i} p_{j'}^{j} p_{k'}^{k} T_{ijk}.$$

Аналогично определяется ковариантный тензор любого другого ранга: тензор m-го ранга имеет не n^3, а n^m составляющих, и в формуле преобразования стоит не три множителя вида $p_{i'}^{i}$, а m таких множителей.

Коэффициенты линейной формы, которые преобразуются, как мы видели, по формуле (24), дают пример ковариантного тензора первого ранга.

Определим теперь *контравариантный тензор* третьего ранга.

Пусть имеется правило, позволяющее в каждой системе координат построить n^3 чисел T^{ijk}, каждое из которых определяется при придании индексам i, j, k числовых значений от 1 до n. Эти числа T^{ijk} образуют по определению *контравариантный тензор третьего ранга*, если преобразование величин T^{ijk} при переходе к новому базису производится по формуле

$$T^{i'j'k'} = q_{i}^{i'} q_{j}^{j'} q_{k}^{k'} T^{ijk},$$

Аналогично определяется контравариантный тензор любого другого ранга. В частности, координаты вектора x образуют контравариантный тензор первого ранга.

Введенные нами термины «ковариантный» и «контравариантный» объясняются очень простым образом. «Ковариантный» означает «изменяющийся так же», как изменяются базисные векторы, т. е. с использованием коэффициентов $p_{i'}^{i}$. «Контравариантный» означает «изменяющийся в обратном направлении», т. е. с использованием коэффициентов $q_{i}^{i'}$.

Существуют еще и смешанные тензоры. Например, n^3 чисел T_{ij}^{k}, заданные в каждой системе координат, образуют *смешанный тензор третьего ранга, два раза ковариантный и один раз контравариантный*, если преобразование этих величин при переходе к новому базису производится по формуле

$$T_{i'j'}^{k'}=p_{i'}^{i}p_{j'}^{j}q_{k}^{k'}T_{ij}^{k}.$$

Аналогично определяется смешанный тензор, l раз ковариантный и m раз контравариантный.

Например, элементы матрицы линейного оператора образуют смешанный тензор второго ранга, один раз ковариантный и один раз контравариантный. Отметим, что целесообразная расстановка индексов предназначена прямо указывать характер того или иного тензора.

5.64. Действия с тензорами. Можно определить операцию *сложения* двух тензоров одинакового строения, например двух тензоров T_{ij}^{k} и S_{ij}^{k} (два раза ковариантных и один раз контравариантных). Это будет тензор Q_{ij}^{k} того же строения; в каждой системе координат при фиксированных i, j, k его составляющая по определению равна сумме соответствующих составляющих слагаемых. То, что величины Q_{ij}^{k} действительно образуют тензор, притом того же строения, что и у слагаемых, вытекает из следующих равенств:

$$\begin{aligned}Q_{i'j'}^{k'}=T_{i'j'}^{k'}+S_{i'j'}^{k'}=p_{i'}^{i}p_{j'}^{j}q_{k}^{k'}T_{ij}^{k}+p_{i'}^{i}p_{j'}^{j}q_{k}^{k'}S_{ij}^{k}=\\=p_{i'}^{i}p_{j'}^{j}q_{k}^{k'}(T_{ij}^{k}+S_{ij}^{k})=p_{i'}^{i}p_{j'}^{j}q_{k}^{k'}Q_{ij}^{k}.\end{aligned}$$

Операция *умножения* применима к тензорам любого строения. Например, умножим тензор T_{ij} на тензор S_{k}^{l}. Это будет тензор четвертого ранга Q_{ijk}^{l}; в каждой системе координат при фиксированных i, j, k, l его составляющая по определению равна произведению соответствующих составляющих множителей. Тензорный характер Q_{ijk}^{l} проверяется следующим образом:

$$\begin{aligned}Q_{i'j'k'}^{l'}=T_{i'j'}S_{k'}^{l'}=p_{i'}^{i}p_{j'}^{j}T_{ij}p_{k'}^{k}q_{l}^{l'}S_{k}^{l}=\\=p_{i'}^{i}p_{j'}^{j}p_{k'}^{k}q_{l}^{l'}T_{ij}S_{k}^{l}=p_{i'}^{i}p_{j'}^{j}p_{k'}^{k}q_{l}^{l'}Q_{ijk}^{l}.\end{aligned}$$

Рассмотрим еще операцию *свертывания*. Она применяется к тензорам, у которых имеется по меньшей мере один ковариантный и один контравариантный индекс. Пусть, например, дан тензор T_{ij}^{k}. Свернуть его по первому нижнему и верхнему индексам означает в каждой системе координат составить величины

$$T_{ij}^{i}.$$

Здесь по индексу i подразумевается суммирование; в итоге величины $T_j = T^i_{ij}$ зависят только от индекса j. *В результате свертывания снова получается тензор, ранг которого на две единицы меньше исходного.* Проверим это на рассмотренном примере. Мы имеем

$$T_{j'} = T^{i'}_{i'j'} = p^i_{i'} p^j_{j'} q^{i'}_k T^k_{ij} = (p^i_{i'} q^{i'}_k)\, p^j_{j'} T^k_{ij} = \delta^i_k p^j_{j'} T^k_{ij}.$$

Здесь при суммировании по индексу k достаточно ограничиться только значением $k = i$; поскольку $\delta^i_i = 1$, мы получаем

$$T_{j'} = p^j_{j'} T^i_{ij} = p^j_{j'} T_j,$$

что и требовалось.

Что получится, если свернуть смешанный тензор второго ранга T^j_i по его двум индексам? Величина $T = T^i_i$ уже не имеет ни одного индекса, т. е. в каждой системе координат она образует только одно число. Это число — одно и то же в любой системе координат; действительно,

$$T' = T^{i'}_{i'} = p^i_{i'} q^{i'}_j T^j_i = \delta^i_j T^j_i = T^i_i = T.$$

Такая числовая величина, не зависящая от системы координат, называется *инвариантом*. Следовательно, операция свертывания дает возможность получать инварианты тензоров.

Например, если тензор a^j_i, соответствующий линейному оператору A, свернуть по его индексам, то полученный инвариант a^i_i будет следом — суммой диагональных элементов матрицы оператора A. Инвариантность этой величины была уже доказана нами другим способом (*5.53*). Еще пример: матрица c^j_i произведения двух операторов с матрицами соответственно a^k_i и b^j_l есть смешанный тензор второго ранга, который получается при свертывании тензора четвертого ранга $a^k_i b^j_l$ по индексам k и l.

ЗАДАЧИ

1. Вектор $x \in K_n$ имеет относительно базиса $e_1, e_2, \ldots, e_n$ координаты $\xi_1, \xi_2, \ldots, \xi_n$. Как построить базис в K_n, чтобы координаты вектора x относительно этого нового базиса стали равными $1, 0, 0, \ldots \ldots, 0$?

2. В n-мерном пространстве K_n выбран базис $e_1, e_2, \ldots, e_n$. Доказать, что каждое подпространство $K' \subset K_n$ может быть задано как совокупность всех векторов $x \in K_n$, координаты которых (относительно базиса $e_1, e_2, \ldots, e_n$) удовлетворяют системе уравнений вида

$$\sum_{j=1}^{n} a_{ij}\xi_j=0 \qquad (i=1,\ 2,\ \ldots,\ k).$$

3. (Продолжение.) Доказать, что каждая гиперплоскость $\mathrm{H}\subset \mathrm{K}_n$ может быть задана как совокупность всех векторов $x\in \mathrm{K}_n$, координаты которых (относительно базиса $e_1, e_2, \ldots, e_n$) удовлетворяют системе уравнений вида

$$\sum_{j=1}^{n} a_{ij}\xi_j=b_i \qquad (i=1,\ 2,\ \ldots,\ k).$$

4. На плоскости выбрано три базиса; координаты вектора x относительно каждого из них равны соответственно ξ_1, ξ_2; η_1, η_2; τ_1, τ_2. Дано

$$\begin{gathered}\eta_1=a_{11}\xi_1+a_{12}\xi_2; \quad \tau_1=b_{11}\xi_1+b_{12}\xi_2;\\ \eta_2=a_{21}\xi_1+a_{22}\xi_2; \quad \tau_2=b_{21}\xi_1+b_{22}\xi_2;\\ A=\|a_{ij}\|, \quad B=\|b_{ij}\|.\end{gathered}$$

Выразить координаты τ_1, τ_2 через координаты η_1, η_2.

5. Для данной линейной формы $f(x)\not\equiv 0$ в пространстве K_n выбрать базис $g_1, g_2, \ldots, g_n$ так, чтобы для всякого вектора $x=\sum_{k=1}^{n}\eta_k g_k$ имело место равенство

$$f(x)=\eta_1.$$

6. Пусть оператор A, действующий в n-мерном пространстве R, имеет k-мерное инвариантное подпространство R′. Тогда, считая временно, что оператор A определен только в подпространстве R′, мы можем построить для него характеристический многочлен k-й степени. Показать, что этот многочлен является делителем характеристического многочлена оператора A, действующего во всем пространстве R.

7. Пусть $\lambda=\lambda_0$ есть r-кратный корень уравнения $\det\|A_{(e)}-\lambda E\|=0$. Показать, что размерность m собственного подпространства $R^{(\lambda_0)}$ оператора A, отвечающего корню λ_0, не превышает r.

8. Показать, что величины δ_i^j образуют тензор второго ранга, один раз ковариантный и один раз контравариантный.

9. Система величин S_{ij} определяется в каждой системе координат путем решения системы уравнений

$$T^{ik}S_{ij}=\delta_j^k,$$

где T^{ik} — ковариантный тензор второго ранга, причем $\det\|T^{ik}\|\neq 0$. Показать, что S_{ij} — ковариантный тензор второго ранга.

10. Если l_i и ζ^i имеют тот же смысл, что и в тексте, какой геометрический смысл имеет свертка тензора

$$l_i\zeta^j$$

по обоим его индексам?

КАНОНИЧЕСКАЯ ФОРМА МАТРИЦЫ ЛИНЕЙНОГО ОПЕРАТОРА

Два оператора А и В, действующие в n-мерном пространстве K_n, называются *эквивалентными*, если существуют такие два базиса в K_n, что матрица оператора А в первом базисе совпадает с матрицей оператора В во втором базисе. Очевидно, что эквивалентные операторы определяют в пространстве K_n одинаковые по своим свойствам линейные преобразования. Но как узнать по матрицам операторов А и В в одном и том же базисе, являются ли они эквивалентными?

В этой главе для данного линейного оператора А в n-мерном пространстве (комплексном или вещественном) мы укажем базис, в котором матрица A оператора А примет «канонический» вид — в некотором смысле наиболее простой из всех возможных.

При этом канонический вид может быть получен непосредственно по элементам матрицы оператора А в произвольном базисе. И оказывается, что если операторы А и В эквивалентны, то канонический вид их матриц совпадает. Таким образом, необходимым и достаточным условием эквивалентности операторов является совпадение их канонических матриц.

Мы начинаем построение с одного частного класса операторов (§ 6.1); общий случай будет рассмотрен в § 6.3.

§ 6.1. Каноническая форма матрицы нильпотентного оператора

6.11. Линейный оператор В, действующий в n-мерном пространстве K_n, называется *нильпотентным*, если при некотором натуральном r выполняется равенство $B^r = 0$, иначе говоря, если $B^r x = 0$ при любом $x \in K_n$. Предположим, что В — нильпотентный оператор и $B^r = 0$; будем считать при этом, что $B^{r-1} \neq 0$, т. е. имеются векторы $x \in K_n$, для которых $B^{r-1}x \neq 0$. Назовем *высотой* вектора $x \in K_n$ наименьшее из чисел m, при которых $B^m x = 0$. Все векторы $x \in K_n$, по нашему предположению, имеют высоту $\leqslant r$, причем имеются векторы с высотой, равной r. Для любого $k \leqslant r$ обозначим через H_k совокупность всех векторов высоты $\leqslant k$. Очевидно, H_k есть подпространство в K_n: если $x \in H_k$ и $y \in H_k$, то $B^k x = 0$, $B^k y = 0$, откуда при любых $\alpha \in K$ и $\beta \in K$ также $B^k(\alpha x + \beta y) = 0$, так что высота вектора $\alpha x + \beta y$ не превосходит k и $\alpha x + \beta y \in H_k$. Очевидно, далее, что $H_r = K_n$ и $0 = H_0 \subseteq H_1 \subseteq \ldots \subseteq H_{r-1} \subset H_r = K_n$. Размерности этих пространств обозначим соответственно через $m_0 = 0 \leqslant \leqslant m_1 \leqslant \ldots \leqslant m_r = n$. Будем строить базис в пространстве K_n следующим образом.

Как мы видели, H_{r-1} не совпадает со всем $K_n = H_r$. Поэтому можно найти векторы $f_1, \ldots, f_{p_1}$, лежащие в H_r и линейно независимые над H_{r-1} (*2.44*) ($p_1 = m_r - m_{r-1}$). Векторы $Bf_1, \ldots, Bf_{p_1}$ лежат в H_{r-1} и линейно независимы над H_{r-2}: действительно, если бы мы имели

$$c_1 Bf_1 + \ldots + c_{p_1} Bf_{p_1} = g \in H_{r-2},$$

то, применяя B^{r-2}, мы получили бы, что

$$c_1 B^{r-1} f_1 + \ldots + c_{p_1} B^{r-1} f_{p_1} = 0,$$

или, что то же,

$$c_1 f_1 + \ldots + c_{p_1} f_{p_1} \in H^{r-1},$$

что по построению не имеет места. Отсюда видно, что размерность $m_{r-1} - m_{r-2}$ пространства H_{r-1} над H_{r-2} равна или больше размерности $m_r - m_{r-1}$ пространства H_r над H_{r-1}. Дополним векторы $Bf_1, \ldots, Bf_{p_1}$ векторами $f_{p_1+1}, \ldots, f_{p_2}$ в H_{r-1} до максимальной системы, линейно независимой над H_{r-2} ($p_2 = m_{r-1} - m_{r-2}$). Применяя ко всем этим векторам оператор B, получаем векторы в H_{r-2}

$$B^2 f_1, \ \ldots, \ B^2 f_{p_1}, \quad Bf_{p_1+1}, \ \ldots, \ Bf_{p_2},$$

линейно независимые над H_{r-3}, что доказывается аналогично предыдущему. Отсюда $m_{r-2} - m_{r-3} \geqslant m_{r-1} - m_{r-2}$, и можно построить в пространстве H_{r-2} векторы $f_{p_2+1}, \ldots, f_{p_3}$, образующие вместе с предыдущими полную систему, линейно независимую над H_{r-3}.

Переходя таким же образом в подпространства $H_{r-3}, \ldots$ $\ldots, H_0=(0)$, мы и получаем, в конце концов, полную систему n линейно независимых векторов.

6.12. Полученную систему векторов можно записать в таблицу

$$
\begin{array}{l}
f_1, \ldots, f_{p_1}, \\
Bf_1, \ldots, B\ f_{p_1}, \qquad f_{p_1+1}, \ldots, \qquad f_{p_2}, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
B^{r-1}f_1, \ldots, B^{r-1}f_{p_1}, B^{r-2}f_{p_1+1}, \ldots, B^{r-2}f_{p_2}, \ldots \\
\qquad\qquad\qquad\qquad \ldots, f_{p_{r-1}+1}, \ldots, f_{p_r}.
\end{array}
$$

Векторы, стоящие в первой строке таблицы, имеют высоту r, векторы следующей строки — высоту $r-1$ и т. д.; векторы последней строки имеют высоту 1, т. е. оператором B переводятся в 0. Каждый столбец таблицы определяет инвариантное подпространство оператора B; первые p_1 инвариантных подпространств имеют размерность r (каждое), следующие p_2-p_1 подпространств — размерность $r-1$ (каждое) и т. д. Последние (одноэлементные) столбцы определяют одномерные инвариантные подпространства.

Все пространство K_n есть прямая сумма p_r указанных подпространств.

6.13. Напишем матрицу оператора B в подпространстве, определяемом векторами первого столбца. В качестве базиса возьмем векторы $B^{r-1}f_1$, $B^{r-2}f_1$, $\ldots$, Bf_1, f_1; в таком порядке они располагаются по возрастанию высоты. В этой записи первый вектор базиса оператором B переводится в 0, второй — в первый, $\ldots$, r-й переводится в $(r-1)$-й: поэтому матрица оператора B согласно *4.23* содержит по r строк и столбцов и имеет вид

$$
\left\|\begin{matrix}
0 & 1 & 0 & \ldots & 0 & 0 \\
0 & 0 & 1 & \ldots & 0 & 0 \\
. & . & . & . & . & . \\
0 & 0 & 0 & \ldots & 0 & 1 \\
0 & 0 & 0 & \ldots & 0 & 0
\end{matrix}\right\| \qquad (1)
$$

с отличными от нуля элементами, именно, равными 1, над главной диагональю.

Аналогичный вид имеет матрица оператора B в остальных инвариантных подпространствах, соответствующих остальным столбцам таблицы; отличие от матрицы (1) может состоять лишь в числе строк и столбцов.

6.14. Матрица оператора B во всем пространстве K_n будет квазидиагональной (*4.52*) с указанными блоками на диагонали:

$$\mathbf{B}=\left\|\begin{array}{ccccc} \boxed{\begin{array}{cccccc} 0 & 1 & 0 & \ldots & 0 & 0\\ 0 & 0 & 1 & \ldots & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \ldots & 0 & 1\\ 0 & 0 & 0 & \ldots & 0 & 0 \end{array}} & & & & \\ & \boxed{\begin{array}{cccccc} 0 & 1 & 0 & \ldots & 0 & 0\\ 0 & 0 & 1 & \ldots & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \ldots & 0 & 1\\ 0 & 0 & 0 & \ldots & 0 & 0 \end{array}} & & & \\ & & \ddots & & \\ & & \boxed{\begin{array}{cc} 0 & 1\\ 0 & 0 \end{array}} & & \\ & & & \boxed{0} & \\ & & & & \ddots \\ & & & & \boxed{0} \end{array}\right\|. \quad (2)$$

Число блоков размера r равно p_1, число блоков размера $r-1$ равно p_2-p_1, ..., число блоков размера 2 равно $p_{r-1}-p_{r-2}$, число блоков размера 1 равно p_r-p_{r-1}. Разумеется, если для некоторого j мы имеем $p_{r-j+1}=p_{r-j}$, то в матрице (2) блоков размера j не будет вовсе.

§ 6.2. Алгебры; алгебра многочленов от одного переменного

6.21. Приведем некоторые определения из теории алгебр. Линейное пространство K над числовым полем K называется *алгеброй* (точнее, алгеброй над K), если в K установлена операция *умножения*, приводящая в соответствие каждой паре элементов x, y из K элемент $z \in \mathrm{K}$ (обозначаемый $x \cdot y$ или xy) и удовлетворяющая следующим условиям:

1) $\alpha(xy) = (\alpha x) y = x(\alpha y)$ для любых x и y из K и $\alpha \in K$;
2) $(xy) z = x(yz)$ для любых x, y, z из K (ассоциативный закон);
3) $(x+y) z = xz + yz$ для любых x, y, z из K (дистрибутивный закон).

Вообще говоря, умножение может не быть коммутативным, так что $xy \neq yx$. Если умножение коммутативно, т. е. выполнено условие

4) $xy = yx$ для любых x и y из K,

то алгебра K называется *коммутативной*.

Пусть 0 есть нуль-вектор пространства K. Тогда для любого $x \in \mathrm{K}$

$$0 \cdot x = (0+0) x = 0 \cdot x + 0 \cdot x,$$

откуда следует, что $0 \cdot x = 0$.

Элемент $e \in \mathrm{K}$ называется *левой единицей*, если $ex = x$ для любого $x \in \mathrm{K}$; *правой единицей*, если $xe = x$ для любого $x \in \mathrm{K}$; двусторонней единицей, или просто *единицей* в K, если $ex = xe = x$ для любого $x \in \mathrm{K}$.

Элемент $x \in \mathrm{K}$ называется *левым обратным к элементу* $y \in \mathrm{K}$, если xy есть единица алгебры K; в этом случае y называется *правым обратным* к x. Если элемент z обладает и левым и правым обратными, то они могут быть лишь единственными и совпадают друг с другом (ср. 4.76а). Элемент z называется в этом случае *обратимым*, ему обратный обозначается через z^{-1}.

Произведение zu обратимого элемента z и обратимого элемента u есть обратимый элемент с обратным $u^{-1}z^{-1}$. Если элемент u обратим, то уравнение $ux = v$ имеет решение $x = u^{-1}v$; поскольку оно получается из самого уравнения умножением на u^{-1}, оно единственно. В коммутативном случае употребляют запись $x=\frac{v}{u}$ или $x=v:u$; этот элемент называют *частным* элементов v и u. Для частных справедливы обычные арифметические правила действий:

$$\frac{v_1}{u_1}+\frac{v_2}{u_2}=\frac{v_1u_2+u_1v_2}{u_1u_2} \quad \text{(если } u_1 \text{ и } u_2 \text{ обратимы);}$$

$$\frac{v_1}{u_1}\cdot\frac{v_2}{u_2}=\frac{v_1v_2}{u_1u_2} \quad \text{(если } u_1 \text{ и } u_2 \text{ обратимы);}$$

$$\frac{v_1}{u_1}:\frac{v_2}{u_2}=\frac{v_1u_2}{u_1v_2} \quad \text{(если } u_1,\ u_2,\ v_2 \text{ обратимы),}$$

доказательства которых мы предоставляем читателю.

Алгебра K по определению *имеет размерность* n, если эту размерность n имеет соответствующее линейное пространство K.

6.22. Примеры.

а. В любом линейном пространстве K положим $x\cdot y=0$ для любых $x\in$ K, $y\in$ K. Мы получим алгебру. Такая алгебра называется *тривиальной*.

б. Примером нетривиальной коммутативной алгебры над полем K является совокупность П всех многочленов

$$P(\lambda)=\sum_{k=0}^{p}a_k\lambda^k$$

с коэффициентами из K, с обычными операциями сложения и умножения. В этой алгебре есть единица, именно, многочлен $e(\lambda)$, у которого $a_0=1$, а остальные коэффициенты равны 0.

в. Линейное пространство $L(\mathrm{K}_n)$ из всех матриц n-го порядка с элементами из K с обычным матричным умножением дает пример конечномерной некоммутативной алгебры размерности n^2 (*4.73б*). Эта алгебра обладает единицей, которой является единичная матрица E.

г. Более общим примером некоммутативной алгебры с единицей является линейное пространство B(K) всех линейных операторов, действующих в линейном пространстве K, с обычным для операторов действием умножения (*4.33*).

6.23а. Пространство L ⊂ K называется *подалгеброй* алгебры K, если из $x\in$ L, $y\in$ L следует $xy\in$ L. Подпространство $L \subset K$ называется *правым идеалом в* K, если из $x \in L$, $y \in K$ следует $xy \in L$, и *левым идеалом* в K, если из $x \in L$, $y \in K$ следует $yx \in L$. Идеал, одновременно левый и правый, называется *двусторонним идеалом*. В коммутативной алгебре нет различия между левыми, правыми и двусторонними идеалами. Во всякой алгебре K имеются два очевидных двусторонних идеала: один, обозначаемый (0),— состоящий из единственного элемента 0, второй — состоящий из всех элементов $x \in K$. Все остальные идеалы, односторонние и двусторонние, называются *собственными* идеалами. Всякий идеал есть подалгебра; обратное, вообще говоря, несправедливо. Так, совокупность всех многочленов $P(\lambda)$, удовлетворяющих условию $P(0) = P(1)$, есть подалгебра алгебры Π, не являющаяся идеалом. Совокупность всех многочленов $P(\lambda)$, удовлетворяющих условию $P(0) = 0$, есть собственный идеал в алгебре Π.

б. Пусть $L \subset K$ есть подпространство в алгебре K. Рассмотрим фактор-пространство K/L (*2.48*), т. е. линейное пространство из классов X элементов $x \in K$, взаимно сравнимых относительно подпространства L. Если L — двусторонний идеал в K, то для классов $X \in K/L$, кроме линейных операций, можно ввести операцию умножения. Именно, имея классы X и Y, выберем произвольно элементы $x \in X$, $y \in Y$ и под произведением $X \cdot Y$ будем понимать класс, содержащий произведение xy. Проверим однозначность этого определения. Если в классе X мы возьмем элемент x', а в классе Y элемент y', то мы будем иметь

$$x'y' - xy = x'(y' - y) + (x' - x)y;$$

этот элемент лежит в идеале L вместе с $y' - y$ и $x' - x$. Из условий *6.21, 1)—3)*, выполненных в алгебре K, немедленно следует выполнение аналогичных условий для классов $X \in K/L$; поэтому фактор-пространство K/L с введенным в нем умножением является также алгеброй. Алгебра K/L называется *фактор-алгеброй алгебры* K *по двустороннему идеалу* L. Очевидно, она коммутативна, если коммутативна алгебра K.

6.24. Пусть имеются две алгебры K' и K'' над полем K. Морфизм ω линейного пространства K' в пространство K'' (*2.71*) называется *морфизмом алгебры* K' *в алгебру* K'', если наряду с условиями морфизма пространств (*2.71*)

а) $\omega(x'+y')=\omega(x')+\omega(y')$ для любых x', $y' \in \mathrm{K}'$,

б) $\omega(\alpha x')=\alpha\omega(x')$ для любого $x' \in \mathrm{K}'$ и любого $\alpha \in K$

выполняется также условие

в) $\omega(x' \cdot y')=\omega(x')\cdot\omega(y')$ для любых x', $y' \in \mathrm{K}'$.

Если морфизм ω есть эпиморфизм или мономорфизм, или изоморфизм пространства K' в пространство K'' (*2.71*), то при выполнении условия в) он соответственно называется *эпиморфизмом, мономорфизмом, изоморфизмом алгебры* K' *в алгебру* K''.

6.25. Примеры.

а. Пусть L — подалгебра алгебры K. Отображение ω, которое каждому вектору $x \in \mathrm{L}$ ставит в соответствие этот же вектор x в алгебре K, есть морфизм алгебры L в алгебру K, и именно мономорфизм. Как и в аналогичном случае в *2.72*, этот мономорфизм называется *вложением* L в K.

б. Пусть L — двусторонний идеал алгебры K и K/L — соответствующая фактор-алгебра (*6.23б*). Отображение ω, которое каждому вектору $x \in \mathrm{K}$ ставит в соответствие класс $\mathrm{X} \in \mathrm{K}/\mathrm{L}$, содержащий элемент x, есть морфизм алгебры K в алгебру K/L, и именно эпиморфизм. Как и в аналогичном случае в *2.72б*, этот эпиморфизм называется *каноническим* отображением K на K/L.

в. Пусть имеется мономорфизм ω алгебры K' в алгебру K''. Совокупность всех векторов $\omega(x') \in \mathrm{K}''$ есть подалгебра $\mathrm{L}'' \subset \mathrm{K}''$, и мономорфизм ω есть изоморфизм алгебры K' в алгебру L''.

г. Пусть имеется морфизм ω алгебры K' в алгебру K''. Покажем, что совокупность L' всех векторов $x' \in \mathrm{K}'$, для которых $\omega(x')=0$, есть двусторонний идеал алгебры K'. Действительно, L' есть подпространство в K' (*2.76б*); далее, если $x' \in \mathrm{L}'$, $y' \in \mathrm{K}'$, то $\omega(x' \cdot y')=\omega(x')\cdot\omega(y')=0$, так что $x' \cdot y' \in \mathrm{L}'$ и аналогично $y' \cdot x' \in \mathrm{L}'$, что и требовалось. Далее, мономорфизм Ω пространства K'/L' в пространство K'', ставящий в соответствие каждому классу $\mathrm{X}' \in \mathrm{K}'/\mathrm{L}'$ элемент $\omega(x')$, $x' \in \mathrm{X}'$ (*2.76б*), в данном случае есть мономорфизм алгебры K'/L' в алгебру K''. Действительно, выбирая $x' \in \mathrm{X}'$, $y' \in \mathrm{Y}'$, мы имеем $x'y' \in \mathrm{X}'\mathrm{Y}'$ и $\Omega(\mathrm{X}'\mathrm{Y}')=\omega(x'y')=\omega(x')\times$

$\times\omega(y')=\Omega(X')\,\Omega(Y')$. В частности, если морфизм ω есть эпиморфизм алгебры K' в алгебру K'', то морфизм Ω есть изоморфизм алгебры K'/L' с алгеброй K''.

д. Пусть A — линейный оператор, действующий в пространстве K. Так как для линейных операторов в пространстве K определены операции сложения и умножения, то каждому многочлену $P(\lambda)=\sum_{k=0}^{m} a_k\lambda^k \in \Pi$ мы можем поставить в соответствие оператор

$$P(A)=\sum_{k=0}^{m} a_k A^k,$$

действующий в том же пространстве K. При этом соответствии сложению и умножению многочленов отвечает сложение и умножение соответствующих операторов в смысле § 4.3. В самом деле, пусть

$$P(\lambda)=P_1(\lambda)+P_2(\lambda)=\sum_{k=0}^{m} a_k\lambda^k+\sum_{k=0}^{m} b_k\lambda^k=\sum_{k=0}^{m}(a_k+b_k)\,\lambda^k;$$

тогда

$$P(A)=\sum_{k=0}^{m}(a_k+b_k)\,A^k=\sum_{k=0}^{m} a_k A^k+\sum_{k=0}^{m} b_k A^k=P_1(A)+P_2(A).$$

Аналогично пусть

$$Q(\lambda)=P_1(\lambda)\,P_2(\lambda)=\sum_{k=0}^{m} a_k\lambda^k\sum_{j=0}^{m} b_j\lambda^j=\sum_{k=0}^{m}\sum_{j=0}^{m} a_k b_j\lambda^{k+j};$$

тогда по распределительному закону для операторов (*4.33*)

$$Q(A)=\sum_{k=0}^{m}\sum_{j=0}^{m} a_k b_j A^{k+j}=\sum_{k=0}^{m} a_k A^k\sum_{j=0}^{m} b_j A^j=P_1(A)\,P_2(A).$$

В частности, операторы $P(A)$ и $Q(A)$ всегда коммутативны, каковы бы ни были многочлены $P(\lambda)$ и $Q(\lambda)$.

Мы получили морфизм алгебры многочленов Π (*6.22б*) в алгебру линейных операторов $B(K)$ (*6.22г*). Вообще говоря, этот морфизм не есть эпиморфизм (хотя бы потому, что операторы вида $P(A)$ коммутируют друг с другом, а вся алгебра $B(K)$, за исключением тривиального случая $K=K_1$, некоммутативна).

е. Существует изоморфизм между алгеброй $\mathrm{B}(K_n)$ всех линейных операторов, действующих в пространстве K_n, и алгеброй $\mathrm{L}(K_n)$ всех матриц n-го порядка с элементами из поля K; мы осуществляем его, фиксируя в пространстве K_n базис $e_1, \ldots, e_n$ и ставя в соответствие каждому оператору $\mathrm{A} \in \mathrm{B}(K_n)$ его матрицу в этом базисе. Обе эти алгебры имеют одинаковую размерность n^2.

6.26. В коммутативной алгебре Π всех многочленов $P(\lambda)$ с коэффициентами из поля K (*6.22б*) совокупность всех многочленов вида $P(\lambda)\,Q_0(\lambda)$, где $Q_0(\lambda)$ — фиксированный многочлен, а $P(\lambda)$ — любой многочлен, очевидно, является идеалом.

Мы покажем, что *каждый идеал $I \neq 0$ алгебры* Π *имеет такую структуру*, т. е. получается из некоторого многочлена $Q_0(\lambda)$ умножением на любой многочлен $P(\lambda)$. Для доказательства найдем в идеале I отличный от 0 многочлен наименьшей возможной степени, например q, и обозначим его через $Q_0(\lambda)$. Мы утверждаем теперь, что любой многочлен $Q(\lambda) \in I$ имеет вид $P(\lambda)\,Q_0(\lambda)$, где $P(\lambda) \in \Pi$. Действительно, по правилу действий с многочленами *) можно написать

$$Q(\lambda) \equiv P(\lambda)\,Q_0(\lambda) + R(\lambda), \tag{3}$$

где $P(\lambda)$ — частное от деления $Q(\lambda)$ на $Q_0(\lambda)$, а $R(\lambda)$ — остаток, имеющий степень, меньшую степени делителя, т. е. меньшую, чем число q. Многочлены $Q(\lambda)$ и $Q_0(\lambda)$ принадлежат идеалу I. Но тогда, как видно из равенства (3), и многочлен $R(\lambda)$ принадлежит идеалу I. Так как степень $R(\lambda)$ меньше q, а многочлен $Q_0(\lambda)$ в идеале имеет наименьшую возможную степень q среди многочленов, отличных от 0, то $R(\lambda) \equiv 0$, что и утверждалось.

Многочлен $Q_0(\lambda)$ называется *порождающим идеал* I.

6.27. *Многочлен $Q_0(\lambda)$ определяется идеалом I единственным образом с точностью до числового множителя.* Действительно, если бы наряду с $Q_0(\lambda)$ тем же свойством

*) См., например, А. Г. Курош, Курс высшей алгебры, «Наука», 1965, гл. 5.

обладал многочлен $Q_1(\lambda)$, то, по доказанному, мы имели бы

$$Q_1(\lambda) = P_1(\lambda) Q_0(\lambda),$$
$$Q_0(\lambda) = P_0(\lambda) Q_1(\lambda).$$

Из этих равенств вытекает, что степени многочленов $Q_1(\lambda)$ и $Q_0(\lambda)$ одинаковы и что $P_1(\lambda)$ и $P_0(\lambda)$ не содержат λ, т. е. являются числами; это и утверждалось.

6.28. Пусть имеются многочлены $Q_1(\lambda), \ldots, Q_m(\lambda)$, не все равные 0 и не имеющие общего делителя степени $\geqslant 1$. Покажем, что существуют такие многочлены $P_1^0(\lambda), \ldots, P_m^0(\lambda)$, что

$$P_1^0(\lambda) Q_1(\lambda) + \ldots + P_m^0(\lambda) Q_m(\lambda) \equiv 1. \tag{4}$$

Действительно, пусть I — совокупность всех многочленов вида

$$P_1(\lambda) Q_1(\lambda) + \ldots + P_m(\lambda) Q_m(\lambda)$$

при любых $P_1(\lambda), \ldots, P_m(\lambda)$ из Π. Очевидно, совокупность I есть идеал в Π. По *6.26* идеал I порожден одним многочленом

$$Q_0(\lambda) = \sum_{k=1}^{m} P_k^0(\lambda) Q_k(\lambda). \tag{5}$$

При этом, в частности,

$$Q_1(\lambda) = S_1(\lambda) Q_0(\lambda), \ \ldots, \ Q_m(\lambda) = S_m(\lambda) Q_0(\lambda),$$

где $S_1(\lambda), \ldots, S_m(\lambda)$ — некоторые многочлены. Эти равенства показывают, что $Q_0(\lambda)$ есть общий делитель многочленов $Q_1(\lambda), \ldots, Q_m(\lambda)$. Но тогда из предположения следует, что степень многочлена $Q_0(\lambda)$ равна 0, т. е. что $Q_0(\lambda)$ есть постоянная a_0. При этом $a_0 \neq 0$, так как иначе $I = 0$. Умножая (5) на $1/a_0$, получаем (4), что и требуется.

§ 6.3. Каноническая форма матрицы произвольного оператора

6.31. Рассмотрим линейный оператор A в n-мерном пространстве K_n.

Указанное в *6.25д* отображение $\omega(P(\lambda)) = P(A)$ есть эпиморфизм (*6.23*) алгебры Π всех многочленов с коэффициентами из поля K в алгебру Π_A линейных операторов вида

$P(\mathrm{A})$ в пространстве K_n. В силу *6.25 г* алгебра Π_A изоморфна фактор-алгебре Π/I, где I — идеал, состоящий из всех многочленов $P(\lambda)$, для которых $\omega(P(\lambda)) = P(\mathrm{A}) = 0$. Мы выясним сейчас структуру этого идеала.

6.32. В *6.22* мы видели, что совокупность всех линейных операторов, действующих в пространстве K_n, представляет собой снова алгебру над тем же полем K, размерности n^2. Фиксируя оператор A, рассмотрим последовательность операторов

$$\mathrm{A}^0 = \mathrm{E},\ \mathrm{A},\ \mathrm{A}^2,\ \ldots,\ \mathrm{A}^m,\ \ldots$$

В этой последовательности первые n^2+1 членов линейно зависимы. Пусть, например,

$$\sum_{k=0}^{m} c_k \mathrm{A}^k = 0 \qquad (m \leqslant n^2).$$

Это означает, что в установленном выше соответствии (*6.31*) многочлену $Q(\lambda) = \sum_{k=0}^{m} c_k \lambda^k$ соответствует нулевой оператор.

Всякий многочлен $Q(\lambda)$, для которого оператор $Q(\mathrm{A})$ есть нулевой оператор, будем называть *аннулирующим многочленом оператора* A. Мы доказали, что *у всякого оператора* A *есть аннулирующий многочлен степени* $\leqslant n^2$.

6.33. Совокупность всех аннулирующих многочленов оператора A есть идеал в алгебре Π. В силу *6.26* имеется многочлен $Q_0(\lambda)$, определенный с точностью до множителя, такой, что все аннулирующие многочлены имеют вид $P(\lambda)\,Q_0(\lambda)$, где $P(\lambda)$ — любой многочлен из Π. В частности, $Q_0(\lambda)$ сам является аннулирующим многочленом. Среди всех аннулирующих многочленов он имеет наименьшую степень и поэтому называется *минимальным аннулирующим многочленом* для оператора A.

6.34. Теорема. *Пусть аннулирующий многочлен $Q(\lambda)$ оператора* A *разложен в произведение двух взаимно простых множителей:*

$$Q(\lambda) = Q_1(\lambda)\,Q_2(\lambda).$$

Тогда пространство K_n *можно разложить в прямую сумму двух подпространств,*

$$K_n = T_1 + T_2,$$

инвариантных относительно оператора A (*так что* $AT_1 \subset T_1$, $AT_2 \subset T_2$), *причем для любых* $x_1 \in T_1$, $x_2 \in T_2$

$$Q_1(A)\, x_2 = 0, \quad Q_2(A)\, x_1 = 0,$$

так что $Q_1(\lambda)\,[Q_2(\lambda)]$ *есть аннулирующий многочлен для оператора* A, *действующего в подпространстве* $T_1\,[T_2]$.

Доказательство. В силу *6.28* существуют такие многочлены $P_1(\lambda)$ и $P_2(\lambda)$, что

$$P_1(\lambda)\, Q_1(\lambda) + P_2(\lambda)\, Q_2(\lambda) \equiv 1.$$

Используя морфизм *6.25д*, находим

$$P_1(A)\, Q_1(A) + P_2(A)\, Q_2(A) \equiv E.$$

Обозначим через $T_k\ (k = 1,2)$ область значений оператора $Q_k(A)$, т. е. совокупность векторов вида $Q_k(A)\, x$, $x \in K_n$ (*4.61*). Очевидно, из $y = Q_k(A)\, x \in T_k$ следует $Ay = Q_k(A)\, Ax \in T_k$, так что подпространство T_k инвариантно относительно оператора A.

Мы имеем для любого $x_1 \in T_1$ и некоторого $y \in K_n$

$$Q_2(A)\, x_1 = Q_2(A)\, Q_1(A)\, y = Q(A)\, y = 0,$$

и аналогично для любого $x_2 \in T_2$ и некоторого $z \in K_n$

$$Q_1(A)\, x_2 = Q_1(A)\, Q_2(A)\, z = Q(A)\, z = 0.$$

Далее, для любого $x \in K_n$ имеет место равенство

$$x = Q_1(A)\, P_1(A)\, x + Q_2(A)\, P_2(A)\, x = x_1 + x_2,$$

где $x_k = Q_k(A)\, P_k(A)\, x \in T_k\ (k = 1, 2)$, оно показывает, что подпространства T_1 и T_2 в сумме дают все K_n.

Пусть $x_0 \in T_1 \cap T_2$. Тогда $Q_1(A)\, x_0 = Q_2(A)\, x_0 = 0$ и, следовательно,

$$x_0 = P_1(A)\, Q_1(A)\, x_0 + P_2(A)\, Q_2(A)\, x_0 = 0,$$

Таким образом, $T_1 \cap T_2 = 0$ и сумма $K_n = T_1 + T_2$ — прямая. Разумеется, не исключена возможность, что одно из подпространств T_1, T_2 состоит только из одного нулевого вектора.

6.35. Замечание. По построению оператор $Q_1(A)$ аннулирует подпространство T_2, а оператор $Q_2(A)$ аннулирует подпространство T_1. Покажем, что *всякий вектор x, аннулируемый оператором $Q_1(A)$, лежит в T_2 и всякий вектор x, аннулируемый оператором $Q_2(A)$, лежит в T_1*. Пусть $Q_1(A)x = 0$. Мы имеем $x = x_1 + x_2$, где $x_1 \in T_1$, $x_2 \in T_2$. Так как $Q_1(A)x_2 = 0$, то и $Q_1(A)x_1 = Q_1(A)x - Q_1(A)x_2 = 0$. Но и $Q_2(A)x_1 = 0$, поскольку $x_1 \in T_1$. Следовательно, $x_1 = P_1(A)Q_1(A)x_1 + P_2(A)Q_2(A)x_1 = 0$, $x = x_2 \in T_2$. Аналогично из $Q_2(A)x = 0$ следует $x \in T_1$, что и утверждалось.

6.36. Разлагая многочлены $Q_1(\lambda)$ и $Q_2(\lambda)$ далее на взаимно простые множители, получаем возможность разбивать пространство K_n на более мелкие инвариантные относительно оператора A подпространства, аннулируемые соответствующими множителями многочленов $Q_1(\lambda)$ и $Q_2(\lambda)$.

Пусть многочлен $P(\lambda)$ допускает в поле K разложение вида

$$P(\lambda) = \prod_{j=1}^{m} (\lambda - \lambda_j)^{r_j}, \tag{6}$$

где λ_j — все (различные) корни многочлена, а r_j — их кратности. Такое разложение всегда возможно, в частности, в поле C комплексных чисел. Разложение (6) есть разложение на m попарно взаимно простых множителей $(\lambda - \lambda_j)^{r_j}$. Применяя результат *6.34*, получаем следующее утверждение:

Теорема. *Если аннулирующий многочлен оператора A имеет вид (6), то пространство K_n разлагается в прямую сумму m подпространств $T_1, \ldots, T_m$, инвариантных относительно оператора A, причем подпространство T_k аннулируется оператором $B_k^{r_k}$, где*

$$B_k = A - \lambda_k E.$$

6.37. В каждом ненулевом пространстве T_k, согласно *6.14*, можно выбрать базис, в котором матрица оператора B_k (по построению нильпотентного в пространстве T_k) примет канонический вид (2). В этом же базисе матрица оператора

$A = B_k + \lambda_k E$ примет вид

$$\begin{Vmatrix}
\begin{array}{cccccc|}
\lambda_k & 1 & 0 & \ldots & 0 & 0 \\
0 & \lambda_k & 1 & \ldots & 0 & 0 \\
. & . & . & . & . & . \\
0 & 0 & 0 & \ldots & \lambda_k & 1 \\
0 & 0 & 0 & \ldots & 0 & \lambda_k \\ \hline
\end{array} & & & \\
& \boxed{\begin{array}{cccccc}
\lambda_k & 1 & 0 & \ldots & 0 & 0 \\
0 & \lambda_k & 1 & \ldots & 0 & 0 \\
. & . & . & . & . & . \\
0 & 0 & 0 & \ldots & \lambda_k & 1 \\
0 & 0 & 0 & \ldots & 0 & \lambda_k
\end{array}} & & \\
& & \ddots & \\
& & & \boxed{\lambda_k}
\end{Vmatrix}. \qquad (7)$$

Матрица оператора A во всем пространстве $K_n = T_1 + \ldots + T_m$ в базисе, который получается объединением всех канонических базисов, построенных в пространствах $T_1, \ldots T_m$, приобретает окончательную форму

$$J(A) = \begin{Vmatrix}
\begin{array}{cccc|}
\lambda_1 & 1 & \ldots & 0 \\
0 & \lambda_1 & \ldots & 0 \\
. & . & . & . \\
0 & 0 & \ldots & 1 \\
0 & 0 & \ldots & \lambda_1 \\ \hline
\end{array} & & & & & & \\
& \boxed{\begin{array}{cccc}
\lambda_1 & 1 & \ldots & 0 \\
0 & \lambda_1 & \ldots & 0 \\
. & . & . & . \\
0 & 0 & \ldots & 1 \\
0 & 0 & \ldots & \lambda_1
\end{array}} & & & & & \\
& & \ddots & & & & \\
& & & \boxed{\lambda_1} & & & \\
& & & & \ddots & & \\
& & & & & \boxed{\begin{array}{cccc}
\lambda_m & 1 & \ldots & 0 \\
0 & \lambda_m & \ldots & 0 \\
. & . & . & . \\
0 & 0 & \ldots & 1 \\
0 & 0 & \ldots & \lambda_m
\end{array}} & \\
& & & & & & \ddots \\
& & & & & & & \boxed{\lambda_m}
\end{Vmatrix}. \qquad (8)$$

Формулируем окончательный результат:

Теорема. *Для любого оператора* A *в n-мерном пространстве* K_n, *имеющего аннулирующий многочлен вида* (6) (*в частности, для любого оператора* A *в n-мерном комплексном пространстве* C_n), *существует базис, в котором матрица оператора* A *записывается в форме* (8). *Матрица* (8) *называется нормальной формой Жордана оператора* A, *а соответствующий базис — базисом Жордана.*

В случае $K_n = C_n$ комплексные числа $\lambda_1, \ldots, \lambda_m$ можно упорядочить по какому-либо правилу: например, в порядке возрастания модулей, а при равных модулях — в порядке возрастания аргумента θ, меняющегося в промежутке $0 \leqslant \theta < 2\pi$.

Для оператора A, действующего в пространстве $K_n \neq C_n$, представление (8) возможно не всегда. Мы рассмотрим в § 6.6 каноническую форму матрицы оператора A, действующего в пространстве $K_n = R_n$.

§ 6.4. Элементарные делители

6.41. Матрицу (8) можно задать таблицей

$$\left.\begin{array}{l} \lambda_1; \quad n_1^{(1)}, \ \ldots, \ n_{r_1}^{(1)} \\ \lambda_2; \quad n_1^{(2)}, \ \ldots, \ n_{r_2}^{(2)} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \lambda_m; \quad n_1^{(m)}, \ \ldots, \ n_{r_m}^{(m)} \end{array}\right\} \left(n_1^{(k)} \geqslant n_2^{(k)} \geqslant \ldots \geqslant n_{r_k}^{(k)}\right), \tag{9}$$

в которой для каждого диагонального числа λ_k указаны размеры $n_1^{(k)}, \ldots, n_{r_k}^{(k)}$ соответствующих «элементарных жордановых клеток» вида

$$\left.\left\| \begin{array}{ccccc} \lambda_k & 1 & 0 & \ldots & 0 \\ 0 & \lambda_k & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_k \end{array} \right\|\right\} n_j^{(k)} \text{ строк}, \tag{10}$$

встречающихся в матрице (8). Мы хотим выяснить, как построить таблицу (9) и тем самым восстановить вид матрицы $J(\mathrm{A})$ оператора A по известной его матрице A в каком-либо базисе пространства K_n.

6.42. Как мы знаем (*5.53*), характеристический многочлен оператора А не зависит от выбора базиса. Составим его для жорданова базиса; так как под главной диагональю стоят нули, мы получаем

$$\det(A-\lambda E)=\det\|J(\mathrm{A})-\lambda E\|=$$
$$=\prod_{k=1}^{m}(\lambda_k-\lambda)^{n_1^{(k)}+\ldots+n_{r_k}^{(k)}}. \qquad (11)$$

Мы видим, что числа λ_k суть корни характеристического многочлена, а суммы $r_k=n_1^{(k)}+\ldots+n_{r_k}^{(k)}$ — их кратности. Таким образом, вычисляя характеристический многочлен (что можно сделать по матрице A) и находя его корни, мы получаем величины λ_k и $r_k=n_1^{(k)}+\ldots+n_{r_k}^{(k)}$ таблицы (9).

6.43. Далее (здесь и в *6.44*) мы укажем, как по матрице A оператора А в исходном базисе вычислить сами числа $n_j^{(k)}$. Поскольку $J(\mathrm{A})$ и A — матрицы одного и того же оператора А, взятого в разных базисах, согласно *5.51* справедливо равенство

$$J(\mathrm{A})=T^{-1}AT,$$

где T — невырожденная матрица. Поэтому и

$$J(\mathrm{A})-\lambda E=T^{-1}(A-\lambda E)\,T.$$

Миноры фиксированного, например p-го, порядка матрицы $A-\lambda E$ представляют собой некоторые многочлены от λ степени $\leqslant p$. Обозначим через $I_p(\mathrm{A})$ идеал в алгебре Π, порожденный всеми этими минорами. Аналогичный смысл имеет идеал $I_p(J(\mathrm{A}))$. Покажем, что эти идеалы совпадают. Действительно, каждый минор p-го порядка матрицы $J(\mathrm{A})-\lambda E$ согласно *4.54* является суммой произведений миноров p-го порядка матриц $A-\lambda E$, T^{-1} и T. Но элементы матриц T и T^{-1} суть числа; таким образом, всякий минор p-го порядка матрицы $J(\mathrm{A})-\lambda E$ есть просто линейная комбинация миноров p-го порядка матрицы $A-\lambda E$ и тем самым входит в идеал $I_p(\mathrm{A})$. По симметрии каждый минор p-го порядка матрицы $A-\lambda E$ входит в идеал $I_p(J(\mathrm{A}))$. Тем самым $I_p(\mathrm{A})=$ $=I_p(J(\mathrm{A}))$, что и утверждалось.

Пусть $D_p(\lambda)$ — порождающий многочлен этого идеала; он может быть определен как общий наибольший делитель многочленов, порождающих идеал $I_p(\mathrm{A})$ (6.26). Таким образом, наибольший общий делитель миноров p-го порядка у матрицы $J(\mathrm{A})-\lambda E$ тот же, что и у миноров p-го порядка матрицы $A-\lambda E$, и поэтому может считаться известным. Вычислим непосредственно наибольший общий делитель миноров p-го порядка матрицы $J(\mathrm{A})-\lambda E$. Опять-таки вместо матрицы $J(\mathrm{A})-\lambda E$ при этом можно рассматривать матрицу вида $P(J(\mathrm{A})-\lambda E)Q$, где P и Q — обратимые числовые матрицы (не содержащие λ). Операции перестановок строк, столбцов, прибавления к одному столбцу другого с произвольным множителем в матрице $J(\mathrm{A})-\lambda E$ приводят как раз к такого рода матрицам (4.44). Мы утверждаем, что элементарную клетку

$$\left\|\begin{matrix} \lambda_k-\lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda_k-\lambda & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_k-\lambda \end{matrix}\right\|$$

указанными операциями можно преобразовать к виду

$$\left.\left\|\begin{matrix} 1 & & & & \\ & 1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \quad (\lambda_k-\lambda)^{n_j^{(k)}} \end{matrix}\right\|\right\} n_j^{(k)} \text{ строк.} \tag{12}$$

А именно, для получения требуемого результата следует вначале из второй строки вычесть первую, умноженную на $\lambda_k-\lambda$, из третьей — вторую, умноженную на $\lambda_k-\lambda$, и т. д.; мы получим матрицу $(p=n_j^{(k)})$

$$\left\|\begin{matrix} \lambda_k-\lambda & 1 & 0 & \dots & 0 \\ -(\lambda_k-\lambda)^2 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (-1)^{p-2}(\lambda_k-\lambda)^{p-1} & 0 & 0 & \dots & 1 \\ (-1)^{p-1}(\lambda_k-\lambda)^{p} & 0 & 0 & \dots & 0 \end{matrix}\right\|.$$

Если теперь из первого столбца вычесть второй, умноженный на $\lambda_k-\lambda$, затем третий, умноженный на $-(\lambda_k-\lambda)^2$, ..., $(p-1)$-й, умноженный на $(-1)^{p-2}(\lambda_k-\lambda)^{p-1}$, мы получим матрицу

$$\begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \\ (-1)^{p-1}(\lambda_k-\lambda)^p & 0 & 0 & \ldots & 0 \end{Vmatrix}, \tag{13}$$

из которой матрица (12) получается перестановкой столбцов.

Теперь подсчитаем общий наибольший делитель миноров p-го порядка у матрицы $\hat{J}(\lambda)$ с клетками вида (12) на главной диагонали. Так как у этой матрицы вне главной диагонали стоят нули, то отличными от нуля могут быть лишь миноры с одинаковым набором номеров строк и столбцов; такой минор равен произведению своих диагональных элементов.

В матрице $\hat{J}(\lambda)$ среди элементов на главной диагонали имеется некоторое число, положим N, биномов вида $(\lambda_k-\lambda)^{n_j^{(k)}}$, а остальные $n-N$ элементов главной диагонали равны 1. Число N есть полное число жордановых клеток в матрице $J(A)$, т. е. $N=r_1+\ldots+r_m$. С другой стороны, среди миноров до порядка $n-N$ заведомо имеются равные 1, откуда следует, что $D_p(\lambda)\equiv 1$ при $p\leqslant n-N$. Можно заменить матрицу $\hat{J}(\lambda)$ более простой диагональной матрицей

$$I(\lambda)=\begin{Vmatrix} (\lambda_1-\lambda)^{n_1^{(1)}} & & & & & & & \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & (\lambda_1-\lambda)^{n_{r_1}^{(1)}} & & & & \\ & & & & (\lambda_2-\lambda)^{n_1^{(2)}} & & & \\ & & & & & \cdot & & \\ & & & & & & \cdot & \\ & & & & & & & (\lambda_m-\lambda)^{n_{r_m}^{(m)}} \end{Vmatrix};$$

тогда многочлен $D_p(\lambda)$, сосчитанный для матрицы $\hat{J}(\lambda)$, будет совпадать с многочленом $D_{p-(n-N)}(\lambda)$, сосчитанным для матрицы $I(\lambda)$. Очевидно, наибольший общий делитель миноров p-го порядка матрицы $I(\lambda)$ имеет вид

$$D_p(\lambda) = \prod_{j=1}^{m} (\lambda_j - \lambda)^{m_j(p)}, \tag{14}$$

где $m_j \geqslant 0$. Спрашивается, каково значение, например, показателя $m_1(p)$. Эта величина есть наименьший показатель, с которым $\lambda_1 - \lambda$ входит во все миноры p-го порядка. Если $p \leqslant n_1^{(2)} + \ldots + n_{r_m}^{(m)}$, то имеется минор p-го порядка, не содержащий вообще $\lambda_1 - \lambda$, так что при этих p мы имеем $m_1(p) = 0$. Для $p = n_1^{(2)} + \ldots + n_{r_m}^{(m)} + 1$, учитывая, что показатели $n_1^{(1)}, \ldots, n_{r_1}^{(1)}$ идут в убывающем порядке, мы имеем

$$m_1(p) = n_{r_1}^{(1)}.$$

В дальнейшем на каждую единицу увеличения p показатель $m_1(p)$ будет увеличиваться соответственно на $n_{r_1-1}^{(1)}, n_{r_1-2}^{(1)}, \ldots$, наконец, при $p = n$ мы получим

$$m_1(p) = n_1^{(1)} + \ldots + n_{r_1}^{(1)}.$$

Аналогично,

$$m_j(p) = n_1^{(j)} + \ldots + n_{r_j}^{(j)} \qquad (1 \leqslant j \leqslant m).$$

6.44. Отношение $E_p(\lambda) = \dfrac{D_{p+1}(\lambda)}{D_p(\lambda)}$ называется *элементарным делителем оператора* A; вместе с многочленами $D_p(\lambda)$ элементарные делители не зависят от выбора базиса, и их можно вычислять по матрице оператора A в любом базисе.

Из сказанного в *6.43* видно, что элементарные делители имеют вид

$$E_{n-q}(\lambda) = \prod_{j=1}^{m} (\lambda_j - \lambda)^{n_q^{(j)}}, \qquad q = 1, 2, \ldots, n-1,$$

с корнями, кратности которых равны размерам (последовательных) жордановых клеток. Таким образом, вычислив элементарные делители, мы получаем числа $n_q^{(j)}$ и тем самым решаем задачу, поставленную в *6.43*.

6.45. Примеры.

а. У жордановой матрицы 10-го порядка

$$\left\|\begin{matrix} 1&1&0&&&&&&&\\ 0&1&1&&&&&&&\\ 0&0&1&&&&&&&\\ &&&1&1&&&&&\\ &&&0&1&&&&&\\ &&&&&1&&&&\\ &&&&&&2&1&&\\ &&&&&&0&2&&\\ &&&&&&&&2&1\\ &&&&&&&&0&2 \end{matrix}\right\|$$

имеются три клетки, отвечающие корню $\lambda_1=1$, размеров 3, 2 и 1, и две клетки, отвечающие корню $\lambda_2=2$, размеров 2 и 2; поэтому элементарные делители имеют значения

$$\begin{aligned} E_9(\lambda)&=(1-\lambda)^3(2-\lambda)^2,\\ E_8(\lambda)&=(1-\lambda)^2(2-\lambda)^2,\\ E_7(\lambda)&=1-\lambda,\\ E_6(\lambda)&=\ldots=E_1(\lambda)=1. \end{aligned}$$

б. У некоторой матрицы $A=\|a_{jk}\|$ 10-го порядка элементарные делители (вычисленные, как указано в *6.43* и *6.44*, по минорам матрицы $A-\lambda E$) оказались равными

$$\begin{aligned} E_9(\lambda)&=(3-\lambda)^2(4-\lambda)^3,\\ E_8(\lambda)&=(3-\lambda)^2(4-\lambda),\\ E_7(\lambda)&=4-\lambda,\\ E_6(\lambda)&=4-\lambda,\\ E_5(\lambda)&=\ldots=E_1(\lambda)=1, \end{aligned}$$

Напишем выражение матрицы $J(\mathrm{A})$. В соответствии с результатом *6.44* матрица $J(\mathrm{A})$ имеет две клетки, отвечающие корню $\lambda_1=3$, размеров 2 и 2, и четыре клетки, отвечающие корню $\lambda_2=4$, размеров 3, 1, 1, 1. Таким образом, мы имеем

$$J(\mathrm{A})=\left\|\begin{array}{cc|cc|ccc|c|c|c}
3 & 1 & & & & & & & & \\
0 & 3 & & & & & & & & \\
\hline
 & & 3 & 1 & & & & & & \\
 & & 0 & 3 & & & & & & \\
\hline
 & & & & 4 & 1 & & & & \\
 & & & & & 4 & 1 & & & \\
 & & & & & & 4 & & & \\
\hline
 & & & & & & & 4 & & \\
\hline
 & & & & & & & & 4 & \\
\hline
 & & & & & & & & & 4
\end{array}\right\|.$$

6.46. Итак, зная элементарные делители оператора A, мы можем определить все величины $n_j^{(k)}$, а вместе с ними — структуру жордановой матрицы оператора A. В частности, мы видим, что *жорданова матрица оператора* A *определяется однозначно самим оператором* A.

С другой стороны, поскольку элементарные делители оператора A определяются через миноры матрицы $A-\lambda E$ в любом базисе, *два эквивалентных оператора* A *и* B, *т. е. имеющих в двух (разных) базисах одну и ту же матрицу, имеют одну и ту же каноническую жорданову форму.*

Очевидно и обратное: *если два оператора имеют одну и ту же каноническую форму, то они эквивалентны.* Тем самым проблема эквивалентности линейных операторов (в комплексном пространстве), поставленная в начале этой главы, полностью решена.

§ 6.5. Некоторые следствия

6.51. Если известно, что оператор A приводится к диагональному виду, т. е. его матрица в некотором базисе имеет вид

$$A=\begin{Vmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & & & & & & & \lambda_m \end{Vmatrix},$$

то матрица A и есть жорданова матрица оператора A (все жордановы клетки имеют размер 1). В частности, все элементарные делители имеют простые корни. Обратно, если все элементарные делители некоторого оператора A имеют только простые корни, жорданова матрица J(A) имеет клетки только размера 1 и, следовательно, диагональна.

6.52. Имея жорданову форму оператора A, можно написать его минимальный аннулирующий многочлен. Пусть оператор B в некотором базисе $e_1, \ldots, e_p$ имеет матрицу

$$\begin{Vmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{Vmatrix}.$$

Это означает, что

$$\mathrm{B}e_1=0, \quad \mathrm{B}e_2=e_1, \ \ldots, \ \mathrm{B}e_p=e_{p-1},$$

откуда следует, что $\mathrm{B}^p x=0$ при любом $x=\sum_1^p c_k e_k$. Таким образом, аннулирующий многочлен оператора B имеет вид λ^p. Минимальный аннулирующий многочлен, как делитель λ^p, должен иметь вид λ^m, $m \leqslant p$; но поскольку $\mathrm{B}^{p-1}e_p = e_1 \neq 0$, мы видим, что λ^p и есть минимальный аннулирующий многочлен.

Пусть оператор A в том же базисе $e_1, \ldots, e_p$ имеет матрицу

$$\left\| \begin{matrix} \lambda_0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_0 \end{matrix} \right\|,$$

так что $\mathrm{A} = \mathrm{B} + \lambda_0 \mathrm{E}$. По доказанному $(\mathrm{A} - \lambda_0 \mathrm{E})^p = \mathrm{B}^p = 0$, так что для оператора A аннулирующим многочленом является $(\lambda_0 - \lambda)^p$; он является и минимальным в силу той же аргументации.

Пусть оператор A имеет матрицу

$$\left\| \begin{matrix} \boxed{\begin{matrix} \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 \\ 0 & 0 & \lambda_0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_0 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \lambda_0 & 1 & \ldots & 0 \\ 0 & \lambda_0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_0 \end{matrix}} \end{matrix} \right\|,$$

причем числа $p_1 \geqslant p_2 \geqslant \ldots \geqslant p_r$ суть размеры ее диагональных клеток. Многочлен $Q(\lambda)$, аннулирующий оператор A, должен по отдельности аннулировать каждую его клетку (ср. *4.52*). Этим свойством обладает многочлен $(\lambda_0 - \lambda)^{p_1}$. По указанным выше причинам он является и минимальным аннулирующим многочленом.

Наконец, в общем случае, когда оператор A обладает жордановой матрицей с таблицей

$$\begin{matrix} \lambda_1; & n_1^{(1)} \geqslant n_2^{(1)} \geqslant \ldots \geqslant n_{p_1}^{(1)}, \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ \lambda_m; & n_1^{(m)} \geqslant n_2^{(m)} \geqslant \ldots \geqslant n_{p_m}^{(m)}, \end{matrix}$$

аннулирующим многочленом является многочлен

$$Q(\lambda)=\prod_{k=1}^{m}(\lambda_k-\lambda)^{n_1^{(k)}}. \tag{15}$$

Он является и минимальным аннулирующим многочленом оператора A, так как ни один из показателей $n_1^{(k)}$ не может быть здесь понижен по указанным выше соображениям.

Итак, *минимальный аннулирующий многочлен матрицы A есть многочлен* (15). Его степень равна $n_1^{(1)}+\ldots+n_1^{(m)}$ — сумме размеров максимальных жордановых клеток, отвечающих каждому корню характеристического многочлена. Заметим, что это число не превосходит размера всей матрицы A (т. е. числа n — размерности пространства, в котором действует оператор). Характеристический многочлен $\det(A-\lambda E)$ оператора A (*6.42*) содержит многочлен $Q(\lambda)$ в качестве делителя и поэтому *также является аннулирующим* (это утверждение называется теоремой Гамильтона — Кэли). Вообще говоря, характеристический многочлен $\det(A-\lambda E)$ не является минимальным аннулирующим многочленом оператора A. Если оказывается, что минимальный многочлен совпадает с характеристическим, то это означает, что каждый характеристический корень используется только в одной жордановой клетке, размера, равного кратности корня.

§ 6.6. Вещественная жорданова форма

6.61. Рассмотрим оператор A в вещественном n-мерном пространстве R_n. Канонический базис, в котором матрица оператора A записывалась бы в жордановой форме (8), вообще говоря, не существует, хотя бы потому, что характеристический многочлен оператора A может иметь невещественные корни. Тем не менее в вещественном пространстве можно найти некоторую замену жордановой матрице (8).

Пусть $A=\|a_k^{(j)}\|$ есть матрица оператора A в некотором базисе $e_1, \ldots, e_n$ пространства R_n. Рассмотрим комплексное n-мерное пространство C_n, состоящее из векторов

$$x=\alpha_1 e_1+\ldots+\alpha_n e_n,$$

где $\alpha_1, \ldots, \alpha_n$ — произвольные *комплексные* числа. Матрица A

задает в пространстве C_n линейный оператор $\hat{A}$ по формулам

$$\hat{A}x = \sum_{j=1}^{n} \alpha_j A e_j = \sum_{j=1}^{n} \alpha_j \left(\sum_{k=1}^{n} a_k^{(j)} e_k \right);$$

для векторов x с вещественными составляющими α_j эти формулы определяют сам оператор A.

6.62. Начнем с оператора A, у которого аннулирующий многочлен имеет специальный вид

$$P(\lambda) = (\lambda^2 + \tau^2)^p,$$

где τ — положительное число.

Для оператора $\hat{A}$ определены операторные полиномы $Q(\hat{A})$ с комплексными коэффициентами. В частности, имеют смысл многочлены $(\hat{A} + i\tau E)^p$ и $(\hat{A} - i\tau E)^p$. Полином $P(\lambda) = (\lambda^2 + \tau^2)^p$ остается аннулирующим и для оператора $\hat{A}$. Разложению $(\lambda^2 + \tau^2)^p = (\lambda - i\tau)^p (\lambda + i\tau)^p$, согласно *6.34*, отвечает разложение пространства C_n в прямую сумму инвариантных относительно оператора $\hat{A}$ подпространств C_n^1 и C_n^2, в которых он имеет аннулирующие многочлены соответственно $(\lambda - i\tau)^p$ и $(\lambda + i\tau)^p$. Более того, если подпространство C_n^1 состоит из векторов

$$x = \alpha_1 e_1 + \ldots + \alpha_n e_n$$

с некоторыми $\alpha_1, \ldots, \alpha_n$, то подпространство C_n^2 состоит из векторов

$$\bar{x} = \bar{\alpha}_1 e_1 + \ldots + \bar{\alpha}_n e_n,$$

где числа $\bar{\alpha}_j$ комплексно сопряжены к числам α_j. Действительно, если $(\hat{A} - i\tau E)^p x = 0$, то, переходя к комплексно сопряженным числам во всех членах левой части, получаем, что $(\hat{A} + i\tau E)^p \bar{x} = 0$, и обратно; последним равенством, по *6.35*, однозначно определяется подпространство C_n^2. Отсюда следует, между прочим, что n четно, $n = 2m$, где m — размерность каждого из подпространств C_n^1 и C_n^2.

Пусть f_j^k — жорданов базис оператора $\hat{A}$ в пространстве C_n^1 *(6.37)*. Поскольку в этом базисе матрица оператора $\hat{A}$ имеет вид [6.37 (7)]

$$
\begin{array}{c}
n_1\left\{\begin{array}{c}\\ \\ \\ \\ \end{array}\right. \\
\\
n_q\left\{\begin{array}{c}\\ \\ \\ \\ \end{array}\right.
\end{array}
\left\|\begin{array}{ccc}
\begin{array}{|cccc|}
i\tau & 1 & & \\
 & i\tau & 1 & \\
 & & \ddots & \\
 & & & i\tau \\ \hline
\end{array} & & \\
 & \ddots & \\
 & & \begin{array}{|cccc}\hline
i\tau & 1 & & \\
 & i\tau & & \\
 & & \ddots & \\
 & & & i\tau
\end{array}
\end{array}\right\|,
$$

оператор $\hat{A}$ действует на базисные векторы по формулам

$$
\begin{aligned}
&\hat{A}f_1^1 = i\tau f_1^1, && \ldots, && \hat{A}f_1^q = i\tau f_1^q, \\
&\hat{A}f_2^1 = f_1^1 + i\tau f_2^1, && \ldots, && \hat{A}f_2^q = f_1^q + i\tau f_2^q, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\hat{A}f_{n_1}^1 = f_{n_1-1}^1 + i\tau f_{n_1}^1, && \ldots, && \hat{A}f_{n_q}^q = f_{n_q-1}^q + i\tau f_{n_q}^q.
\end{aligned}
$$

Для совокупности сопряженных векторов $\overline{f_j^k} \in C_n^2$ справедливы сопряженные формулы

$$
\begin{aligned}
&\hat{A}\overline{f_1^1} = -i\tau\overline{f_1^1}, && \ldots, && \hat{A}\overline{f_1^q} = -i\tau\overline{f_1^q}, \\
&\hat{A}\overline{f_2^1} = \overline{f_1^1} - i\tau\overline{f_2^1}, && \ldots, && \hat{A}\overline{f_2^q} = \overline{f_1^q} - i\tau\overline{f_2^q}, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\hat{A}\overline{f_{n_1}^1} = \overline{f_{n_1-1}^1} - i\tau\overline{f_{n_1}^1}, && \ldots, && \hat{A}\overline{f_{n_q}^q} = \overline{f_{n_q-1}^q} - i\tau\overline{f_{n_q}^q}.
\end{aligned}
$$

Мы видим, что векторы $\overline{f_j^k}$ образуют жорданов базис для оператора $\hat{A}$ в пространстве C_n^2. Таким образом, векторы f_j^k и $\overline{f_j^k}$ образуют жорданов базис для оператора $\hat{A}$ во всем пространстве C_n. Мы построим теперь базис в вещественном пространстве R_n, заменяя каждую пару комплексных векторов f_j^k и $\overline{f_j^k}$ на пару вещественных векторов $g_j^k = \frac{1}{2}(f_j^k + \overline{f_j^k})$ и $h_j^k = \frac{1}{2i}(f_j^k - \overline{f_j^k})$. Из формул

$$
\begin{aligned}
\hat{A}f_j^k &= f_{j-1}^k + i\tau f_j^k \qquad (f_0^k = \overline{f_0^k} = 0), \\
\hat{A}\overline{f_j^k} &= \overline{f_{j-1}^k} - i\tau\overline{f_j^k}
\end{aligned}
$$

получаем

$$\hat{\mathrm{A}}\left\{\frac{1}{2}(f_j^k+\overline{f_j^k})\right\}\equiv \mathrm{A}g_j^k=g_{j-1}^k-\tau h_j^k \qquad (g_0^k=h_0^k=0),$$
$$\hat{\mathrm{A}}\left\{\frac{1}{2i}(f_j^k-\overline{f_j^k})\right\}\equiv \mathrm{A}h_j^k=h_{j-1}^k+\tau g_j^k.$$

Поэтому для векторов g_j^k, h_j^k получаются формулы

$$\left.\begin{array}{l}
\mathrm{A}g_1^k=-\tau h_1^k,\\
\mathrm{A}h_1^k=\qquad \tau g_1^k,\\
\mathrm{A}g_2^k=\qquad\qquad g_1^k-\tau h_2^k,\\
\mathrm{A}h_2^k=\qquad h_1^k\qquad\qquad +\tau g_2^k,\\
\dots\dots\dots\dots\dots\dots\dots\dots\\
\mathrm{A}g_{n_k}^k=\qquad\qquad\qquad\qquad g_{n_k-1}^k-\tau h_n^k,\\
\mathrm{A}h_{n_k}^k=\qquad\qquad\qquad h_{n_k-1}^k\qquad\qquad +\tau g_{n_k}^k.
\end{array}\right\} \tag{16}$$

При этом, обратно,

$$f_j^k=g_j^k+ih_j^k, \quad \overline{f_j^k}=g_j^k-ih_j^k.$$

Таким образом, линейная (комплексная) оболочка всех векторов g_j^k и h_j^k та же, что и линейная оболочка всех векторов f_j^k и $\overline{f_j^k}$. А так как число тех и других одинаково, то векторы g_j^k и h_j^k линейно независимы над полем C вместе с векторами f_j^k и $\overline{f_j^k}$. Тем более векторы g_j^k и h_j^k линейно независимы над полем R, т. е. в вещественном линейном пространстве R_n.

Из формул (16) следует, что матрица оператора A в базисе g_j^k, h_j^k состоит из диагональных клеток вида

$$\left\|\begin{array}{cccccccccc}
0 & \tau & 1 & 0 & & & & & & \\
-\tau & 0 & 0 & 1 & & & & & & \\
 & & 0 & \tau & 1 & 0 & & & & \\
 & & -\tau & 0 & 0 & 1 & & & & \\
 & & & & 0 & \tau & & & & \\
 & & & & -\tau & 0 & & & & \\
 & & & & & & \ddots & & & \\
 & & & & & & 0 & \tau & 1 & 0 \\
 & & & & & & -\tau & 0 & 0 & 1 \\
 & & & & & & & & 0 & \tau \\
 & & & & & & & & -\tau & 0
\end{array}\right\| \tag{17}$$

размеров соответственно $2n_1, \ldots, 2n_q$.

6.63. Переходим к рассмотрению общего случая. Пусть A — линейный оператор в вещественном n-мерном пространстве R_n и $P(\lambda)$ — его аннулирующий многочлен. В вещественной области многочлен $P(\lambda)$ допускает разложение вида

$$P(\lambda)=\prod_{k=1}^{m}(\lambda_k-\lambda)^{r_k}\prod_{j=1}^{s}[(\lambda-\sigma_j)^2+\tau_j^2)]^{p_j},$$

где λ_k $(k=1, \ldots, m)$ — различные вещественные корни многочлена $P(\lambda)$, а $\sigma_j+i\tau_j=\lambda_j$; $\sigma_j-i\tau_j=\bar{\lambda}_j$ — различные невещественные корни; все числа τ_j положительны. В соответствии с общей теорией (*6.36*) пространство R_n допускает разложение в прямую сумму инвариантных относительно A подпространств

$$\mathrm{R}=\sum_{k=1}^{m}\mathrm{E}_k+\sum_{j=1}^{s}\mathrm{F}_j,$$

причем аннулирующим многочленом для оператора A в подпространстве E_k служит многочлен $(\lambda-\lambda_k)^{r_k}$, а в подпространстве F_j — многочлен $[(\lambda-\sigma_j)^2+\tau_j^2]^{p_j}$. В подпространстве E_k оператор A приводится к жордановой форме вида (7). В подпространстве F_j обозначим $\mathrm{B}_j=\mathrm{A}-\sigma_j\mathrm{E}$; тогда для оператора B_j в подпространстве F_j многочлен $(\lambda^2+\tau_j^2)^{p_j}$ будет аннулирующим и по предыдущему можно будет построить базис, в котором матрица оператора B_j будет состоять из клеток вида (17) (с заменой τ на τ_j). В этом же базисе матрица оператора $\mathrm{A}=\mathrm{B}_j+\sigma_j\mathrm{E}$ будет состоять из клеток

$$\left\|\begin{matrix} \sigma_j & \tau_j & 1 & 0 & & & & \\ -\tau_j & \sigma_j & 0 & 1 & & & & \\ & & \sigma_j & \tau_j & 1 & 0 & & \\ & & -\tau_j & \sigma_j & 0 & 1 & & \\ & & & & \ddots & & & \\ & & & & & & \sigma_j & \tau_j \\ & & & & & & -\tau_j & \sigma_j \end{matrix}\right\|. \qquad (18)$$

Итак, в пространстве R_n можно выбрать базис, в котором матрица оператора A состоит из диагональных клеток вида (10) и (18). Обозначим эту матрицу через $J_R(\mathrm{A})$ (*вещественная жорданова матрица*).

6.64. Структуру матрицы $J_R(\mathrm{A})$ можно восстановить по элементарным делителям оператора A, вычисленным по минорам матрицы $A-\lambda E$ в исходном базисе (§ 6.4). Поскольку многочлены $D_p(\lambda)$ и $E_p(\lambda)$ получаются из миноров матрицы $A-\lambda E$ рациональными операциями, многочлены $E_p(\lambda)$ имеют вещественные коэффициенты и, следовательно, имеют вид

$$E_{n-q}(\lambda)=\prod_{k=1}^{m}(\lambda-\lambda_k)^{n_q^{(k)}}\prod_{j=1}^{s}[(\lambda-\sigma_j)^2+\tau_j^2]^{p_q^{(j)}}$$

$$(q=1,\ 2,\ \ldots,\ n-1).$$

Каждому показателю $n_q^{(k)}$ отвечает жорданова клетка вида (10) размера $n_q^{(k)}$. Каждому показателю $p_q^{(j)}$ отвечает клетка вида (18) размера $2p_q^{(j)}$.

6.65. Формулируем окончательный результат в виде следующей теоремы:

Теорема. *Для любого оператора* A *в вещественном* n*-мерном пространстве* R_n *существует базис, в котором матрица* $J_R(\mathrm{A})$ *оператора* A *квазидиагональна и состоит из диагональных клеток вида* (10) *и* (18). *Здесь* λ_k — *вещественные характеристические корни оператора* A, $\sigma_j+i\tau_j$ *и* $\sigma_j-i\lambda_j$ — *комплексные его корни. Размеры клеток однозначно определяются по элементарным делителям оператора, как указано в 6.64.*

§ 6.7. Спектры, корпусы и многочлены

В различных вопросах анализа и алгебры встречается необходимость вычисления различных функций, в частности, многочленов, от заданных линейных операторов в конечномерном пространстве. Функции от операторов обладают рядом специфических свойств. В ближайших двух параграфах строится исчисление таких функций. Естественной арифметической моделью для функций от одного оператора является алгебра корпусов, с которой мы и начинаем теорию.

6.71. Будем называть *спектром* и обозначать символом S конечную совокупность точек $\lambda_1,\ \ldots,\ \lambda_m$. При этом будем считать, что каждой точке λ_k приписано под названием «кратности» некоторое натуральное число r_k $(k=1,\ \ldots,\ m)$.

Все это будем записывать так:

$$S = \{\lambda_1^{r_1}, \ldots, \lambda_m^{r_m}\}.$$

Предположим, что каждой точке λ_k поставлен в соответствие набор из r_k чисел из K, которые мы обозначим $f(\lambda_k)$, $f'(\lambda_k), \ldots, f^{(r_k-1)}(\lambda_k)$. Такой набор чисел мы будем называть *корпусом* f, заданным на S *). Пусть $Q(S)$ — совокупность всех корпусов на множестве S.

Введем в $Q(S)$ следующие операции:

а. Сложение. *Суммой* двух корпусов $f = \{f^{(j)}(\lambda_k)\}$ и $g = \{g^{(j)}(\lambda_k)\}$ будем называть корпус $f+g$, определенный набором чисел

$$(f+g)^{(j)}(\lambda_k) = f^{(j)}(\lambda_k) + g^{(j)}(\lambda_k), \qquad k = 1, \ldots, m, \quad j = 1, \ldots, r_k - 1.$$

б. Умножение на число. Пусть $\alpha \in K$; *произведением* αf корпуса $f = \{f^{(j)}(\lambda_k)\}$ на число α будем называть корпус, определяемый набором чисел

$$(\alpha f)^{(j)}(\lambda_k) = \alpha f^{(j)}(\lambda_k).$$

Очевидно, операции *а* и *б* превращают совокупность $Q(S)$ в линейное пространство. Нулем этого пространства являются корпус 0 со всеми нулевыми составляющими.

в. Введем теперь операцию умножения. *Произведением* fg корпусов $f = \{f^{(j)}(\lambda_k)\}$ и $g = \{g^{(j)}(\lambda_k)\}$ называется корпус, определяемый формулами

$$\left.\begin{array}{l} fg(\lambda_k) = f(\lambda_k)\, g(\lambda_k), \\ (fg)'(\lambda_k) = f(\lambda_k)\, g'(\lambda_k) + f'(\lambda_k)\, g(\lambda_k), \\ \dots\dots\dots\dots\dots\dots \\ (fg)^{(p)}(\lambda_k) = \sum\limits_{j=0}^{p} C_p^j f^{(j)}(\lambda_k)\, g^{(p-j)}(\lambda_k), \\ \dots\dots\dots\dots\dots\dots \end{array}\right\} \begin{array}{c} k = 1, \ldots, m, \\ p = 0, 1, \ldots, r_k - 1, \\ C_p^j = \dfrac{p!}{j!\,(p-j)!}. \end{array}$$

Легко проверить, что операция умножения коммутативна и удовлетворяет условиям *6.21* 1) — 3). Таким образом, $Q(S)$ — коммутативная *алгебра* над полем K. В этой алгебре есть единица, корпус e, который обладает свойством $ef = f$ для

*) Принят также термин *джет* (англ. jet означает «струя»).

любого $f \in Q(S)$. Именно, можно положить

$$e^{(j)}(\lambda_k) = \begin{cases} 1 \text{ при } j=0, \quad k=1, \ldots, m, \\ 0 \text{ при } 0 < j \leqslant r_k, \quad k=1, \ldots, m. \end{cases}$$

Для случая, когда точки $\lambda_1, \ldots, \lambda_m$ принадлежат полю K, *мы свяжем в этом параграфе алгебру $Q(S)$ с алгеброй всех многочленов $P(\lambda) = \sum_{k=0}^{p} a_k \lambda^k$ с коэффициентами из поля K.*

6.72. Будем предполагать, что поле K содержит бесконечное множество различных элементов. При этом условии установим вначале некоторые теоремы о восстановлении коэффициентов многочлена по его значениям.

а. Пусть $P(\lambda) = \sum_{k=0}^{p} a_k \lambda^k$ — многочлен с коэффициентами из поля K, аргумент которого λ также может принимать значения в поле K. Покажем, что значения $P(\lambda)$ позволяют однозначно восстановить его коэффициенты $a_0, \ldots, a_p$. Пусть $\lambda_0, \lambda_1, \ldots, \lambda_p$ — различные элементы поля K; рассмотрим равенства

$$\begin{gathered} a_0 + a_1\lambda_0 + \ldots + a_p\lambda_0^p = P(\lambda_0), \\ a_0 + a_1\lambda_1 + \ldots + a_p\lambda_1^p = P(\lambda_1), \\ \cdots\cdots\cdots\cdots \\ a_0 + a_1\lambda_p + \ldots + a_p\lambda_p^p = P(\lambda_p). \end{gathered}$$

Эти равенства можно рассматривать как уравнения относительно величин $a_0, a_1, \ldots, a_p$ с определителем, отличным от 0; по теореме *1.73* она обладает единственным решением, что и требуется.

б. Отсюда следует, что если два многочлена $P(\lambda) = \sum_{k=0}^{p} a_k \lambda^k$ и $Q(\lambda) = \sum_{k=0}^{p} b_k \lambda^k$ совпадают при каждом значении $\lambda \in K$, то $a_k = b_k$ $(k = 0, 1, \ldots, p)$.

6.73. Далее нам понадобится понятие производной от многочлена $P(\lambda)$, понятие высших производных и формула Тейлора. В анализе эти понятия вводятся для случая многочлена $P(\lambda)$ от вещественного (или комплексного) аргумента.

Здесь мы рассматриваем многочлены с аргументом λ, меняющимся в произвольном поле K; поэтому мы должны ввести соответствующие определения независимо (причем не опираясь на понятие предела, которого в поле K может и не быть).

а. Фиксируем точку $\mu \in K$ и напишем равенство

$$\sum_{k=0}^{p} a_k \lambda^k = \sum_{k=0}^{p} a_k [\mu + (\lambda - \mu)]^k = \sum_{k=0}^{p} \frac{b_k(\mu)}{k!} (\lambda - \mu)^k, \quad (19)$$

где $\frac{b_k(\mu)}{k!}$ $(k = 0, 1, \ldots, p)$ — многочлены от μ, которые получаются после развертывания $[\mu + (\lambda - \mu)]^k$ по степеням μ и $(\lambda - \mu)$ и приведения подобных членов. Многочлены $b_k(\mu)$ получают следующие наименования:

$b_0(\mu) \equiv \sum_{k=0}^{p} a_k \mu^k \equiv P(\mu)$ — сам многочлен $P(\mu)$;

$b_1(\mu) \equiv \sum_{k=1}^{p} k a_k \mu^{k-1} = P'(\mu)$ — первая производная многочлена $P(\mu)$,

$b_2(\mu) \equiv \sum_{k=2}^{p} k(k-1) a_k \mu^{k-2} = P''(\mu)$ — вторая производная многочлена $P(\mu)$,

. . . .

$b_p(\mu) \equiv p(p-1) \ldots 1 \cdot a_p = P^{(p)}(\mu)$ — p-я производная многочлена $P(\mu)$.

Для многочлена степени p положим $P^{(q)}(\mu) \equiv 0$ при $q > p$.

Равенство (19) в приведенных обозначениях принимает форму

$$P(\lambda) = \sum_{k=0}^{p} \frac{1}{k!} P^{(k)}(\mu) (\lambda - \mu)^k \quad (20)$$

и называется *формулой Тейлора для многочленов* $P(\lambda)$.

б. В частности, для многочлена $P(\lambda) = (\lambda - a)^p$ $(a \in K)$ мы имеем

$$P(a) = P'(a) = \ldots = P^{(p-1)}(a) = 0, \quad P^{(p)}(\lambda) \equiv p!, \quad P^{(q)}(\lambda) = 0$$

при $q > p$.

в. Более общим образом, если $P(\lambda)=(\lambda-a)^p Q(\lambda)$, мы имеем также

$$Q(\lambda)=\sum_{k=0}^{r} b_k(\lambda-a)^k, \quad P(\lambda)=\sum_{k=0}^{r} b_k(\lambda-a)^{k+p},$$

так что

$$P(a)=P'(a)=\ldots=P^{(p-1)}(a)=0. \tag{21}$$

г. Обратно, если известно, что выполнены условия (21), то это означает, что

$$P(\lambda)=\sum_{k=0}^{s}\frac{1}{k!}P^{(k)}(a)(\lambda-a)^k=$$

$$=(\lambda-a)^p\sum_{k=p}^{s}\frac{1}{k!}P^{(k)}(a)(\lambda-a)^{k-p}=(\lambda-a)^p Q(\lambda),$$

где $Q(\lambda)$ — некоторый новый многочлен.

6.74. Заметим, что представление многочлена $P(\lambda)$ в форме

$$\sum_{k=0}^{p} a_k\lambda^k\equiv P(\lambda)=\sum_{k=0}^{p} b_k(\mu)(\lambda-\mu)^k,$$

где $b_k(\mu)$ — некоторые многочлены от μ, может быть лишь единственным. Действительно, зафиксируем $\mu=\mu_0$ и придадим величине λ последовательно различные значения $\lambda_0, \ldots, \lambda_p$. Тогда $\tau=\lambda-\mu$ примет последовательно различные значения $\lambda_0-\mu_0, \lambda_1-\mu_0, \ldots, \lambda_p-\mu_0$, и при этих значениях τ оказываются известными значениями многочлена $\sum_{k=0}^{p} b_k(\mu_0)\tau^k$ (которые равны $P(\lambda_0), \ldots, P(\lambda_p)$); в силу *6.72а* величины $b_k(\mu_0)$ определены однозначно. Так как это верно при любом $\mu=\mu_0\in K$, то однозначно определены и сами многочлены $b_k(\mu)$ $(k=0, 1, \ldots, p)$.

6.75а. Пусть даны два многочлена $P(\lambda)$ и $Q(\lambda)$. Проверим справедливость формул

$$(P+Q)^{(k)}(\mu)=P^{(k)}(\mu)+Q^{(k)}(\mu), \tag{22}$$

$$(PQ)^{(k)}(\mu)=\sum_{j=0}^{k} C_k^j P^{(j)}(\mu)Q^{(k-j)}(\mu), \tag{23}$$

где $C_k^j=\dfrac{k!}{j!\,(k-j)!}$ $(k=0, 1, \ldots)$.

Действительно, мы имеем по определению

$$(P+Q)(\lambda)=\sum_{k=0}^{p}\frac{1}{k!}(P+Q)^{(k)}(\mu)(\lambda-\mu)^k,$$

$$P(\lambda)=\sum_{k=0}^{p}\frac{1}{k!}P^{(k)}(\mu)(\lambda-\mu)^k,$$

$$Q(\lambda)=\sum_{k=0}^{p}\frac{1}{k!}Q^{(k)}(\mu)(\lambda-\mu)^k,$$

$$P(\lambda)+Q(\lambda)=\sum_{k=0}^{p}\frac{1}{k!}\left[P^{(k)}(\mu)+Q^{(k)}(\mu)\right](\lambda-\mu)^k.$$

В силу теоремы единственности *6.74* получаем формулы (22). Далее, аналогично

$$(PQ)(\lambda)=\sum_{k=0}^{p}\frac{1}{k!}(PQ)^{(k)}(\mu)(\lambda-\mu)^k;$$

с другой стороны,

$$P(\lambda)=\sum_{j=0}^{p}\frac{1}{j!}P^{(j)}(\mu)(\lambda-\mu)^j,\quad Q(\lambda)=\sum_{s=0}^{p}\frac{1}{s!}Q^{(s)}(\mu)(\lambda-\mu)^s,$$

$$P(\lambda)Q(\lambda)=\sum_{j=0}^{p}\sum_{k=0}^{p}\frac{1}{s!j!}P^{(j)}(\mu)Q^{(s)}(\mu)(\lambda-\mu)^{s+j}=$$
$$=\sum_{k=0}^{2p}\left\{\sum_{j=0}^{k}\frac{1}{j!(k-1)!}P^{(j)}(\mu)Q^{(k-j)}(\mu)\right\}(\lambda-\mu)^k;$$

в силу теоремы единственности *6.74* имеем

$$\frac{1}{k!}(PQ)^{(k)}(\mu)=\sum_{j=0}^{k}\frac{1}{j!(k-j)!}P^{(j)}(\mu)Q^{(k-j)}(\mu),$$

откуда и вытекает (23).

б. В частности, из формулы (23) вытекает следующее важное предложение:

Если $P^{(k)}(\mu)=0$ при $k=0, 1, \ldots, m$, то для любого многочлена $Q(\lambda)$ также $(PQ)^{(k)}(\mu)=0$ при $k=0, 1, \ldots, m$.

6.76. Пусть теперь заданы некоторый спектр $S=\{\lambda_1^{r_1}, \ldots, \lambda_m^{r_m}\}$, $\lambda_j \in K$, и алгебра соответствующих корпусов $Q(S)$ (*6.71*).

Каждому многочлену $P(\lambda)$ отвечает следующий корпус P из $Q(S)$: точке λ_k ставятся в соответствие числа $P(\lambda_k)$, $P'(\lambda_k)$, $\ldots$, $P^{(r_k-1)}(\lambda_k)$, где $P^{(j)}(\lambda)$ — производные от многочлена $P(\lambda)$, определенные в *6.73а*. Формулы (22) и (23) показывают, что обычным действиям сложения и умножения многочленов отвечают определенные в *6.71* действия над корпусами. Таким образом, отображение $P(\lambda) \longrightarrow P$ есть морфизм (*6.24*) алгебры многочленов Π в алгебру корпусов $Q(S)$. Покажем, что этот морфизм является *эпиморфизмом*; иначе говоря, для любого корпуса f можно найти такой многочлен $P(\lambda)$, что

$$f(\lambda_k)=P(\lambda_k),\ f'(\lambda_k)=P'(\lambda_k),\ \ldots,\ f^{(r_k-1)}(\lambda_k)=P^{(r_k-1)}(\lambda_k) \quad (k=1, \ldots, m).$$

Достаточно рассмотреть случай, когда числа $f^{(j)}(\lambda_k)$ отличны от нуля только для какого-нибудь одного значения $k=k_1$, а для остальных значений равны 0; если мы решим поставленную задачу для такого случая, т. е. для каждого $k=1, \ldots, m$ построим многочлен $P_k(\lambda)$, удовлетворяющий условиям

$$P_k(\lambda_k)=f(\lambda_k),\ \ldots,\ P_k^{(r_k-1)}(\lambda_k)=f^{(r_k-1)}(\lambda_k), \tag{24}$$

$$P_k^{(j)}(\lambda_s)=0, \quad s \neq k, \quad j=1, \ldots, r_s-1, \tag{25}$$

то искомое решение можно будет получить по формуле

$$P(\lambda)=P_1(\lambda)+\ldots+P_m(\lambda).$$

Итак, нам требуется найти многочлен $P_k(\lambda)$, удовлетворяющий условиям (24)—(25). Будем искать его в форме

$$P_k(\lambda)=Q_k(\lambda)\,R_k(\lambda), \tag{26}$$

где $Q_k(\lambda)$ — новый искомый многочлен, а

$$R_k(\lambda)=\prod_{s \neq k}(\lambda-\lambda_s)^{r_s}. \tag{27}$$

В силу *6.72б* и *6.75б* мы имеем

$$R_k^{(j)}(\lambda_s)=0 \quad (s \neq k,\ j=0, 1, \ldots, r_s-1);$$

отсюда, каков бы ни был многочлен $Q_k(\lambda)$, мы получаем, снова применяя *6.75б*, что

$$P_k^{(j)}(\lambda_s)=0 \quad (s\neq k,\ j=0,\ 1,\ \ldots,\ r_s-1),$$

так что условие (25) заведомо выполнено.

Мы должны подчинить многочлен $P_k(\lambda)$ условиям (24). Заметим, что

$$R_k(\lambda_k)=\prod_{s\neq k}(\lambda_k-\lambda_s)^{r_s}\neq 0.$$

Поэтому из условия

$$f(\lambda_k)=P_k(\lambda_k)=Q_k(\lambda_k)\,R_k(\lambda_k)$$

величина $Q_k(\lambda_k)$ однозначно определяется. Далее из условия

$$f'(\lambda_k)=P_k'(\lambda_k)=Q_k'(\lambda_k)\,R_k(\lambda_k)+Q_k(\lambda_k)\,R_k'(\lambda_k)$$

при известном уже $Q_k(\lambda_k)$ однозначно определяется $Q_k'(\lambda_k)$; продолжая таким образом дальше, мы получаем возможность однозначно определить все числа $Q_k(\lambda_k), Q_k'(\lambda_k), \ldots, Q_k^{(r_k-1)}(\lambda_k)$. А имея их, мы можем определить и искомый многочлен $Q_k(\lambda)$ по формуле Тейлора

$$Q_k(\lambda)=\sum\frac{1}{j!}\,Q_k^{(j)}(\lambda_k)\,(\lambda-\lambda_k)^j. \tag{28}$$

Возвращаясь по цепочке наших рассуждений, мы видим, что многочлен $P_k(\lambda)$, определенный формулами (26)—(28), удовлетворяет требуемым условиям (24)—(25).

6.77. Применяя *6.25г*, получаем, что алгебра $Q(S)$ всех корпусов на данном спектре изоморфна фактор-алгебре Π/I, где I — идеал в Π, образованный теми многочленами, для которых

$$P^{(j)}(\lambda_k)=0, \quad k=1,\ \ldots,\ m, \quad j=1,\ \ldots,\ r_k.$$

Из *6.73г* следует, что каждый многочлен $P(\lambda)\in I$ делится на многочлен

$$T(\lambda)=\prod_{k=1}^{m}(\lambda-\lambda_k)^{r_k}; \tag{29}$$

из *6.73в* следует, что каждый многочлен $P(\lambda)$, делящийся на $T(\lambda)$, входит в I. Идеал I, как и всякий идеал в алгебре Π, порождается входящим в него многочленом наименьшей степени (*6.26*). Таким многочленом является сам многочлен $T(\lambda)$. Следовательно, *алгебра* $Q(S)$ *изоморфна фактор-алгебре* Π/I, где I — идеал, порожденный многочленом $T(\lambda)$.

6.78. Применим результат *6.77* к решению следующей задачи: *описать все обратимые* (*6.21*) *элементы алгебры* $Q(S)$.

Очевидно, корпус f, для которого хотя бы при одном значении j имеет место равенство $f(\lambda_k)=0$, не может быть обратимым, поскольку в этом случае для любого корпуса g мы имеем

$$(fg)(\lambda_k)=f(\lambda_k)\,g(\lambda_k)=0\neq 1=e(\lambda_k).$$

Рассмотрим корпус f, для которого $f(\lambda_k)\neq 0$ $(k=1, \ldots, m)$. Пусть $P(\lambda)$ — многочлен, для которого (*6.76*)

$$f(\lambda_k)=P(\lambda_k), \ \ldots, \ f^{(r_k-1)}(\lambda_k)=P^{(r_k-1)}(\lambda_k) \quad (k=1, \ldots, m).$$

Этот многочлен, таким образом, не имеет общих множителей с многочленом $T(\lambda)$ (29); поэтому, согласно *6.28*, существуют такие многочлены $Q(\lambda)$ и $S(\lambda)$, что

$$P(\lambda)\,Q(\lambda)+T(\lambda)\,S(\lambda)\equiv 1. \tag{30}$$

Пусть q — корпус, отвечающий многочлену $Q(\lambda)$. Применяя к равенству (30) эпиморфизм $\Pi\to Q(S)$, построенный в *6.76*, и используя тот факт, что при этом эпиморфизме многочлен $T(\lambda)$ переходит в 0, мы получаем

$$f\cdot q=1,$$

т. е. корпус f обратим в алгебре $Q(S)$. Итак, *корпус* $f\in Q(S)$ *обратим*.

Как мы знаем из *6.21*, для обратимого корпуса u разрешимо, притом единственным образом, любое уравнение вида

$$ux=v,$$

где v — известный корпус, а x — неизвестный корпус. Для частного $x=\dfrac{v}{u}$ можно найти явное выражение, последовательно решая уравнения

$$\begin{gathered} u(\lambda_k)\, x(\lambda_k) = v(\lambda_k), \\ u(\lambda_k)\, x'(\lambda_k) + u'(\lambda_k)\, x(\lambda_k) = v'(\lambda_k), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \sum_{j=0}^{r} C_r^j u^{(j)}(\lambda_k)\, x^{(r-j)}(\lambda_k) = v^{(r)}(\lambda_k), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ k = 1, \ldots, m, \quad r = 0, 1, \ldots, r_k. \end{gathered}$$

6.79а. Спектр $S = \{\lambda_1^{r_1}, \ldots, \lambda_m^{r_m}\}$ с комплексными $\lambda_1, \ldots, \lambda_m$ называется *симметричным*, если для любого невещественного $\lambda_k = \sigma_k + i\tau_k$ в S содержится и комплексно сопряженное число $\bar\lambda_k = \sigma_k - i\tau_k$ с той же кратностью r_k. Корпус $f = \{f^{(j)}(\lambda_k)\}$ на симметричном спектре S называется *симметричным*, если числа $f^{(j)}(\lambda_k)$ комплексно сопряжены числам $f^{(j)}(\bar\lambda_k)$ $(j = 0, 1, \ldots, r_k)$. Если $P(\lambda)$ — многочлен с вещественными коэффициентами, то на симметричном спектре S корпус, образованный числами $P^{(j)}(\lambda_k)$ $(k = 1, \ldots, m, j = 1, \ldots, r_k)$, симметричен, поскольку производные $P^{(j)}(\lambda)$ также имеют вещественные коэффициенты и поэтому

$$P^{(j)}(\bar\lambda_k) = \overline{P^{(j)}(\lambda_k)}. \tag{31}$$

Обратно, для симметричного корпуса $f = \{f^{(j)}(\lambda_k)\}$ на симметричном спектре $S = \{\lambda_1^{r_1}, \ldots, \lambda_m^{r_m}\}$ всегда можно найти многочлен $P_0(\lambda)$ с вещественными коэффициентами, для которого $P_0^{(j)}(\lambda_k) = f^{(j)}(\lambda_k)$ $(k = 1, \ldots, m, j = 1, \ldots, r_k)$. Действительно, по *6.76* можно построить многочлен $P(\lambda)$ с комплексными коэффициентами, удовлетворяющий требуемым условиям. Обозначим через $\bar P(\lambda)$ многочлен с комплексно сопряженными коэффициентами. Тогда

$$\frac{1}{2}\left[P^{(j)}(\lambda_k) + \bar P^{(j)}(\lambda_k)\right] = \frac{1}{2}\left[P^{(j)}(\lambda_k) + \overline{P^{(j)}(\bar\lambda_k)}\right] = \\ = \frac{1}{2}\left[f^{(j)}(\lambda_k) + \overline{f^{(j)}(\bar\lambda_k)}\right] = f^{(j)}(\lambda_k),$$

т. е. многочлен $P_0(\lambda) = \frac{1}{2}\left[P(\lambda) + \bar P(\lambda)\right]$, имеющий вещественные коэффициенты, удовлетворяет требуемому условию.

б. Симметричные корпусы f на симметричном спектре S, очевидно, образуют алгебру над полем вещественных чисел.

В силу *6.25г* эта алгебра *изоморфна фактор-алгебре* Π/I, *где* Π — *алгебра всех многочленов с вещественными коэффициентами*, а $I \subset \Pi$ — *идеал, образованный теми многочленами* $P(\lambda) \in \Pi$, *для которых*

$$P^{(j)}(\lambda_k) = 0, \qquad k = 1, \ldots, m, \; j = 1, \ldots, r_k,$$

т. е. идеал, порожденный (*вещественным*) *многочленом*

$$T(\lambda) = \prod_{k=1}^{m} (\lambda - \lambda_k)^{r_k}.$$

§ 6.8. Функции от оператора и их матричная запись

В этом параграфе строится исчисление функций от операторов: для каждого линейного оператора A, действующего в n-мерном пространстве C_n (или R_n), указываются матрицы любого многочлена $P(A)$ или рациональной функции $\frac{P(A)}{Q(A)}$ и правила действий над ними. В последнем пункте дается расширение операторного исчисления на аналитические функции.

6.81. Пусть дан оператор A в пространстве K_n. Алгебра Π_A всех операторов $P(A)$, где $P(\lambda)$ есть некоторый многочлен, как мы знаем из *6.31—6.33*, изоморфна фактор-алгебре Π/I_A, где Π — алгебра всех многочленов, а I_A — идеал, порожденный минимальным аннулирующим многочленом $T(\lambda)$ оператора A. Пусть известно, что многочлен $T(\lambda)$ допускает в поле K разложение на множители

$$T(\lambda) = \prod_{k=1}^{m} (\lambda - \lambda_k)^{r_k}. \tag{32}$$

В силу *6.77* фактор-алгебра Π/I_A изоморфна алгебре $Q(S)$ всех корпусов, определенных на спектре $S = S_A = \{\lambda_1^{r_1}, \ldots, \lambda_m^{r_m}\}$ (спектре оператора A).

Следовательно, и *алгебра* Π_A *изоморфна алгебре* $Q(S_A)$. Явный вид этого изоморфизма можно получить следующим образом: каждому корпусу $f \in Q(S_A)$ отвечает класс многочленов $P(\lambda) \in \Pi$ таких, что

$$P^{(j)}(\lambda_k) = f^{(j)}(\lambda_k) \tag{33}$$
$$(k = 1, \ldots, m, \ j = 0, 1, \ldots, r_{k-1}),$$

и каждому из этих многочленов отвечает один и тот же вполне определенный оператор $P(\mathrm{A}) \in \Pi_{\mathrm{A}}$, который мы будем обозначать $f(\mathrm{A})$.

Далее мы укажем явный вид матрицы оператора $P(\mathrm{A})$ при заданных значениях (33), если матрица A дана в жордановой форме.

6.82. Пусть известно, что в некотором базисе пространства K_n оператор A записывается матрицей n-го порядка специального вида

$$\begin{Vmatrix} \lambda_0 & 1 & \ldots & 0 \\ 0 & \lambda_0 & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & \ldots & \lambda_0 \end{Vmatrix}, \tag{34}$$

т. е. имеет вид $\lambda_0\mathrm{E} + \mathrm{B}$, где B имеет матрицу

$$B = \begin{Vmatrix} 0 & 1 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & \ldots & 0 \end{Vmatrix}.$$

Из *4.74б* мы знаем, что матрица B^k имеет вид

$$B^k = \begin{array}{c} \begin{array}{ccccccccc} & & & & k+1 & & & & \end{array} \\ \begin{Vmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ . & . & . & . & . & . & . \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \\ . & . & . & . & . & . & . \\ 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \end{Vmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ k+1 \\ \\ \\ \end{array}$$

(диагональ из единиц удалена от главной диагонали на k шагов).

Пусть $P(\lambda)$ — произвольный многочлен степени p; применяя формулу Тейлора (20), мы можем написать

$$P(\lambda)=\sum_{k=0}^{p}\frac{1}{k!}P^{(k)}(\lambda_0)(\lambda-\lambda_0)^k.$$

Заменяя λ на символ оператора A, получаем равенство

$$P(\mathrm{A})=\sum_{k=0}^{p}\frac{1}{k!}P^{(k)}(\lambda)(\mathrm{A}-\lambda_0\mathrm{E})^k=\sum_{k=0}^{p}\frac{1}{k!}P^{(k)}(\lambda_0)\,\mathrm{B}^k.$$

Учитывая выражения матриц B^k, находим матрицу

$$P(\mathrm{A})=\left\|\begin{matrix} P(\lambda_0) & P'(\lambda_0) & \frac{1}{2!}P''(\lambda_0) & \ldots & \frac{1}{(n-1)!}P^{(n-1)}(\lambda_0) \\ 0 & P(\lambda_0) & P'(\lambda_0) & \ldots & \frac{1}{(n-2)!}P^{(n-2)}(\lambda_0) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & P(\lambda_0) \end{matrix}\right\|. \quad (35)$$

Заметим, что для построения матрицы $P(\mathrm{A})$ от многочлена $P(\lambda)$ нам понадобились только значения $P(\lambda_0)$, $P'(\lambda_0)$, ..., $P^{(n-1)}(\lambda_0)$, где n — порядок матрицы A.

6.83. Пусть теперь A записывается квазидиагональной матрицей A n-го порядка, состоящей из m диагональных клеток вида (34), где λ_0 пробегает значения $\lambda_1, \ldots, \lambda_m$, с размерами соответственно $n_1, \ldots, n_m$. В силу правил действий с квазидиагональными матрицами (*4.52*) вычисление $P(\mathrm{A})$ можно вести независимо для каждой диагональной клетки. Применяя *6.82*, получаем: *матрица* $P(\mathrm{A})$ *получается путем замены каждой диагональной клетки* (34) *матрицы* A *на клетку* (35). Таким образом, для построения матрицы $P(\mathrm{A})$ нам требуются в данном случае значения $P^{(j)}(\lambda_k)$, $k=1, \ldots, m$, $j=1, \ldots, n_m$.

6.84. Пусть $K=C$ есть поле комплексных чисел. Тогда для любого оператора A, действующего в пространстве $\mathrm{K}_n=\mathrm{C}_n$, минимальный аннулирующий полином имеет разложение (32), и существует базис, в котором матрица оператора A записывается квазидиагональной матрицей с клетками вида

$$\left\| \begin{matrix} \lambda_k & 1 & 0 & \ldots & 0 \\ 0 & \lambda_k & 1 & \ldots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_k \end{matrix} \right\| \tag{36}$$

размера $n_k^{(j)} \leqslant r_k$ (*6.37*).

Таким образом, оператор A определяет спектр

$$S_{\mathrm{A}} = \{\lambda_1^{r_1}, \ \ldots, \ \lambda_m^{r_m}\}.$$

Если f — любой корпус на спектре S_{A},

$$f = \{f^{(j)}(\lambda_k), \qquad k = 1, \ \ldots, \ m, \ j = 1, \ \ldots, \ r_k\},$$

то соответствующий оператор $f(\mathrm{A})$ в силу рассуждений *6.81—6.83* имеет вид квазидиагональной матрицы, в которой каждая клетка вида (36) заменена на клетку

$$\left\| \begin{matrix} f(\lambda_k) & f'(\lambda_k) & \frac{1}{2!} f''(\lambda_k) & \ldots & \frac{1}{(n_k^{(j)}-1)!} f^{(n_k^j - 1)}(\lambda_k) \\ 0 & f(\lambda_k) & f'(\lambda_k) & \ldots & \frac{1}{(n_k^j - 2)!} f^{(n_k^j - 2)}(\lambda_k) \\ . & . & . & . & . \\ 0 & 0 & 0 & \ldots & f(\lambda_k) \end{matrix} \right\|. \tag{37}$$

Этим задача изоморфизма алгебры Π_{A} с алгеброй $Q(S_{\mathrm{A}})$ решена в явном виде.

6.85a. Остановимся еще на функциях от оператора, матрица которого в некотором базисе имеет вид (*6.62*)

$$\left\| \begin{matrix} \sigma & \tau & 1 & 0 & & & & & \\ -\tau & \sigma & 0 & 1 & & & & & \\ & & \sigma & \tau & & & & & \\ & & -\tau & \sigma & & & & & \\ & & & & \cdot & & & & \\ & & & & & \cdot & & & \\ & & & & & & \cdot & & \\ & & & & & & & \sigma & \tau \\ & & & & & & & -\tau & \sigma \end{matrix} \right\|, \tag{38}$$

где σ и τ — элементы из поля K.

Положим

$$E=\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}, \quad \Lambda=\begin{Vmatrix} \sigma & \tau \\ -\tau & \sigma \end{Vmatrix}.$$

Матрицу A можно записать в форме блок-матрицы с двустрочными квадратными блоками:

$$A=\begin{Vmatrix} \Lambda & E & 0 & \ldots & 0 & 0 \\ 0 & \Lambda & E & \ldots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & \Lambda & E \\ 0 & 0 & 0 & \ldots & 0 & \Lambda \end{Vmatrix}=\begin{Vmatrix} \Lambda & 0 & \ldots & 0 & 0 \\ 0 & \Lambda & \ldots & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & \ldots & \Lambda & 0 \\ 0 & 0 & \ldots & 0 & \Lambda \end{Vmatrix}+\begin{Vmatrix} 0 & E & 0 & \ldots & 0 & 0 \\ 0 & 0 & E & \ldots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots & 0 & E \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{Vmatrix}.$$

Поэтому, так же как и выше, и на основании правила умножения блок-матриц (*4.51*) матрицу $P(\mathrm{A})$ можно записать в форме блок-матрицы с блоками того же размера

$$P(\mathrm{A})=\begin{Vmatrix} P(\Lambda) & P'(\Lambda) & \frac{1}{2}P''(\Lambda) & \ldots & \frac{1}{(m-1)!}P^{(m-1)}(\Lambda) \\ 0 & P(\Lambda) & P'(\Lambda) & \ldots & \frac{1}{(m-2)!}P^{(m-2)}(\Lambda) \\ . & . & . & . & . \\ 0 & 0 & 0 & \ldots & P(\Lambda) \end{Vmatrix}. \quad (39)$$

б. Если матрица A квазидиагональна и состоит из диагональных клеток вида (34) и (38), то, так же как в *6.83*, выводим, что матрица $P(\mathrm{A})$ получается заменой каждой диагональной клетки матрицы A на соответствующую клетку вида (35) или (39).

в. Отметим случай, когда в *а* $K=R$ и числа σ и τ вещественны. В этом случае для вещественного многочлена $P(\lambda)$ можно явно указать вид матриц $P^{(k)}(\Lambda)$, фигурирующих в формуле (39). Действительно, если ввести матрицу

$$I=\begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix},$$

то, как легко проверить, мы имеем $I^2=-E$, так что алгебра вещественных матриц

$$\Lambda=\sigma E+\tau I=\begin{Vmatrix} \sigma & \tau \\ -\tau & \sigma \end{Vmatrix}=\begin{Vmatrix} \operatorname{Re}\lambda & \operatorname{Im}\lambda \\ -\operatorname{Im}\lambda & \operatorname{Re}\lambda \end{Vmatrix}, \quad \lambda=\sigma+i\tau,$$

изоморфна обычной алгебре комплексных чисел λ (ср. *4.74а*). Поэтому для любого многочлена $P(\lambda)$ с вещественными коэффициентами

$$P(\Lambda)=P(\sigma E+\tau I)=\begin{Vmatrix} \operatorname{Re}P(\lambda) & \operatorname{Im}P(\lambda) \\ -\operatorname{Im}P(\lambda) & \operatorname{Re}P(\lambda)\end{Vmatrix},\quad \lambda=\sigma+i\tau,$$

и соответственно

$$P^{(k)}(\Lambda)=P^{(k)}(\sigma E+\tau I)=\begin{Vmatrix} \operatorname{Re}P^{(k)}(\lambda) & \operatorname{Im}P^{(k)}(\lambda) \\ -\operatorname{Im}P^{(k)}(\lambda) & \operatorname{Re}P^{(k)}(\lambda)\end{Vmatrix}.$$

6.86. Пусть $K=R$; тогда для любого оператора А, действующего в пространстве $\mathrm{K}_n=\mathrm{R}_n$, минимальный аннулирующий многочлен $T(\lambda)$ имеет вещественные коэффициенты и, следовательно, определяет симметричный спектр S_{A}. Алгебра операторов $P(\mathrm{A})$ изоморфна фактор-алгебре Π/I_T, где Π — алгебра многочленов с вещественными коэффициентами, а I_T — идеал, порожденный многочленом $T(\lambda)$. Эта фактор-алгебра изоморфна алгебре симметричных корпусов на спектре S_T (*6.79*). С другой стороны, для оператора А существует базис, в котором матрица A оператора А квазидиагональна с диагональными клетками вида (34) или (38). Если теперь f — любой симметричный корпус на спектре S_{A}, то соответствующая ему матрица $f(A)$, в соответствии со сказанным выше, получается заменой каждой клетки (34) на (37) и каждой клетки (38) размера, положим, $2m$, на блок-матрицу

$$\begin{Vmatrix} f(\Lambda) & f'(\Lambda) & \dots & \frac{1}{(m-1)!}f^{(m-1)}(\Lambda) \\ 0 & f(\Lambda) & \dots & \frac{1}{(m-2)!}f^{(m-2)}(\Lambda) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(\Lambda)\end{Vmatrix},$$

где двустрочные квадратные блоки $f^{(k)}(\Lambda)$ имеют вид

$$f^{(k)}(\Lambda)=\begin{Vmatrix} \operatorname{Re}f^{(k)}(\lambda) & \operatorname{Im}f^{(k)}(\lambda) \\ -\operatorname{Im}f^{(k)}(\lambda) & \operatorname{Re}f^{(k)}(\lambda)\end{Vmatrix}.$$

6.87. Пусть А — линейный оператор в пространстве C_n; какие операторы $P(\mathrm{A})$ (где $P(\lambda)$ многочлен) обратимы?

Из выражения матрицы оператора $P(\mathrm{A})$ в жордановом базисе оператора А (36) видно, что

$$\det P(\mathrm{A}) = \prod_{k=1}^{m} [P(\lambda_k)]^{n_k}, \quad \sum_{n=1}^{k} n_k = n.$$

Поэтому оператор $P(\mathrm{A})$ обратим в алгебре $\mathrm{B}(\mathrm{C}_n)$ всех линейных операторов, действующих в пространстве C_n, тогда и только тогда, когда $P(\lambda_k) \neq 0$, $k = 1, \ldots, m$. Покажем, что при выполнении этого условия обратный оператор $[P(\mathrm{A})]^{-1}$ имеется уже в алгебре Π_{A}. Действительно, в данном случае корпус p, отвечающий многочлену $P(\lambda)$ в алгебре корпусов $Q(S_{\mathrm{A}})$, т. е. составленный из чисел

$$P^{(j)}(\lambda_k) \qquad (k = 1, \ldots, m; \ j = 0, \ldots, r_k - 1),$$

обратим в алгебре $Q(S_{\mathrm{A}})$ (*6.78*); а так как алгебра $Q(S_{\mathrm{A}})$ изоморфна алгебре Π_{A} (*6.81*), то оператор $P(\mathrm{A})$ обратим в алгебре Π_{A}.

По тем же соображениям изоморфизма между алгебрами $Q(S_{\mathrm{A}})$ и Π_{A} в этом случае для любого многочлена $Q(\lambda)$ разрешимо в алгебре Π_{A} уравнение

$$P(\mathrm{A}) \cdot z(\mathrm{A}) = Q(\mathrm{A})$$

с неизвестным многочленом $z(\lambda)$. В соответствии с результатами *6.78* и *6.84* матрица искомого оператора $z(\mathrm{A})$ в жордановом базисе оператора А получается заменой каждой клетки (36) на клетку вида

$$\left\| \begin{matrix} \dfrac{Q(\lambda_k)}{P(\lambda_k)} & \left(\dfrac{Q(\lambda)}{P(\lambda)}\right)'_{\lambda=\lambda_k} & \dfrac{1}{2}\left(\dfrac{Q(\lambda)}{P(\lambda)}\right)''_{\lambda=\lambda_k} & \ldots \\ 0 & \dfrac{Q(\lambda_k)}{P(\lambda_k)} & \left(\dfrac{Q(\lambda)}{P(\lambda)}\right)'_{\lambda=\lambda_k} & \ldots \\ 0 & 0 & \dfrac{Q(\lambda_k)}{P(\lambda_k)} & \ldots \\ \cdot & \cdot & \cdot & \cdot \end{matrix} \right\|. \tag{40}$$

6.88. Полученный результат можно истолковать еще и следующим образом. Пусть $S = \{\lambda_1^{r_1}, \ldots, \lambda_m^{r_m}\}$ — некоторый спектр в комплексной плоскости. Обозначим через $T(S)$ совокупность всех комплексных рациональных функций $f(\lambda) = \dfrac{Q(\lambda)}{P(\lambda)}$, где $Q(\lambda)$ и $P(\lambda)$ — многочлены, причем многочлен $P(\lambda)$ не имеет корней в точках множества S. В совокупности $T(S)$ введем по обычным правилам операции сложения, умножения на комплексные числа и умножения функций друг на друга, после чего совокупность $T(S)$ становится алгеброй над полем C. Далее заметим, что каждая функция $f(\lambda)\in T(S)$ обладает производными $f'(\lambda)$, $f''(\lambda)$, ... в обычном смысле анализа. Если поставить в соответствие каждой функции $f(\lambda)\in T(S)$ корпус

$$f=\{f^{(j)}(\lambda_k),\ k=1,\ \ldots,\ m;\ j=0,\ 1,\ \ldots,\ r_k-1\},$$

где $f^{(j)}(\lambda)$ означает обычную производную от функции $f(\lambda)$, то это соответствие будет морфизмом алгебры $T(S)$ рациональных функций в алгебру $Q(S)$ корпусов на спектре S, и именно эпиморфизмом, поскольку корпусы, отвечающие даже только многочленам $Q(\lambda)$, заполняют всю алгебру $Q(S)$ (*6.76*). Пусть теперь спектр $S=S_{\mathrm{A}}$ есть спектр некоторого оператора A, действующего в пространстве C_n. Тогда алгебра Π_{A} операторов $P(\mathrm{A})$ изоморфна алгебре корпусов $Q(S_{\mathrm{A}})$, и мы можем имеющийся эпиморфизм $T(S_{\mathrm{A}})\to Q(S_{\mathrm{A}})$ продолжить до эпиморфизма $T(S_{\mathrm{A}})\to\Pi_{\mathrm{A}}$. Иными словами, мы можем поставить в соответствие каждой рациональной функции $f(\lambda)\in T(S)$ линейный оператор $f(\mathrm{A})\in\Pi_{\mathrm{A}}$ так, что соответствие $f(\lambda)\to f(\mathrm{A})$ будет снова эпиморфизмом. Матрица оператора $f(\mathrm{A})$ строится по приведенному выше правилу (40).

6.89*. Вместо алгебры рациональных функций мы можем рассмотреть алгебру аналитических функций. Именно, пусть $W(S)$ — совокупность функций $f(\lambda)$, аналитических в точках $\lambda_1,\ \ldots,\ \lambda_m$ (т. е. аналитических каждая в некоторой окрестности указанных точек). Совокупность $W(S)$ с обычными операциями сложения и умножения есть снова алгебра над полем C (содержащая алгебру $T(S)$). У аналитических функций также имеются производные любого порядка (в обычном смысле анализа); пользуясь ими, мы можем распространить эпиморфизм $T(S_{\mathrm{A}})\to\Pi_{\mathrm{A}}$, построенный в *6.88*, до эпиморфизма $W(S_{\mathrm{A}})\to\Pi_{\mathrm{A}}$. Существенно, что в сферу действия этого эпиморфизма втягиваются теперь многие трансцендентные функции, встречающиеся в анализе, например, $e^{t\lambda}$, $\sin\omega\lambda$, $\cos\omega\lambda$ и др. Если обозначить через $f(\mathrm{A})$ оператор, отвечающий функции $f(\lambda)\in W(S_{\mathrm{A}})$, то матрица его в жордановом базисе оператора A вычисляется по тому же правилу (37). Отметим, например, формулу

$$e^{(t_1+t_2)\,\mathrm{A}} = e^{t_1 \mathrm{A}} e^{t_2 \mathrm{A}},$$

вытекающую непосредственно из тождества $e^{(t_1+t_2)\lambda} = e^{t_1\lambda}e^{t_2\lambda}$ и из того факта, что отображение $W(S_{\mathrm{A}}) \to \Pi_{\mathrm{A}}$ является эпиморфизмом.

Результаты *6.87—6.89*, относящиеся к линейным операторам в комплексном пространстве, можно перенести на линейные операторы в вещественном пространстве, используя вещественную жорданову форму и метод *6.85—6.86*. В таком переносе уже не требуется использование никаких новых идей и мы можем предоставить его читателю.

ЗАДАЧИ

1. Матрица оператора A в некотором базисе $e_1, \ldots, e_n$ имеет вид

$$\left\| \begin{matrix} \lambda & & & & & & \\ 1 & \lambda & & & & & \\ & 1 & \lambda & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & \lambda & \\ & & & & & 1 & \lambda \end{matrix} \right\|.$$

В каком базисе она будет иметь нормальную жорданову форму?

2. Доказать, что матрица A и матрица A' (полученная транспонированием матрицы A) всегда эквивалентны.

3. Найти жорданову форму матрицы

$$\left\| \begin{matrix} -2 & -1 & -1 & 3 & 2 \\ -4 & 1 & -1 & 3 & 2 \\ 1 & 1 & 0 & -3 & -2 \\ -4 & -2 & -1 & 5 & 1 \\ 1 & 1 & 1 & -3 & 0 \end{matrix} \right\|.$$

4. Эквивалентны ли операторы, заданные матрицами

$$A = \left\| \begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{matrix} \right\|, \quad B = \left\| \begin{matrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{matrix} \right\|?$$

5. Найти элементарные делители матриц n-го порядка

$$A_1=\begin{Vmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ . & . & . & . \\ 0 & 0 & \dots & 1 \end{Vmatrix}, \quad A_2=\begin{Vmatrix} 1 & 2 & 3 & \dots & n \\ 0 & 1 & 2 & \dots & n-1 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & 1 \end{Vmatrix},$$

$$A_3=\begin{Vmatrix} n & n-1 & n-2 & \dots & 1 \\ 0 & n & n-1 & \dots & 2 \\ 0 & 0 & n & \dots & 3 \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & n \end{Vmatrix}, \quad A_4=\begin{Vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 2 & 2 & \dots & 2 \\ 0 & 0 & 3 & \dots & 3 \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & n \end{Vmatrix}.$$

6. Показать, что все матрицы

$$A=\begin{Vmatrix} \alpha & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & \alpha & a_{23} & \dots & a_{2n} \\ 0 & 0 & \alpha & \dots & a_{3n} \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & \alpha \end{Vmatrix}$$

с любыми элементами a_{12}, a_{13}, ... эквивалентны, если элементы a_{12}, a_{23}, ..., $a_{n-1,\,n}$ отличны от нуля.

7. Найти жорданову форму матрицы A, удовлетворяющей уравнению $P(A)=0$, где многочлен $P(\lambda)$ не имеет кратных корней.

8. Найти жорданову форму матрицы A, удовлетворяющей уравнению $P(A)=0$, где $P(\lambda)$ — произвольный многочлен.

9. Если аннулирующий многочлен оператора A есть многочлен 2-й степени, то любой вектор x пространства R лежит в инвариантной (относительно A) плоскости или прямой.

10. Найти все матрицы, коммутирующие с матрицей

$$A_m(a)=\left.\begin{Vmatrix} a & 1 & 0 & \dots & 0 & 0 \\ 0 & a & 1 & \dots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & a & 1 \\ 0 & 0 & 0 & \dots & 0 & a \end{Vmatrix}\right\} m \text{ строк.}$$

11. Найти все $m\times n$-матрицы B, удовлетворяющие условию

$$BA_n(a)=A_m(a)B.$$

12. Найти все матрицы, коммутирующие с квазидиагональными матрицами вида

$$\begin{Vmatrix} A_{m_1}(a) & 0 & \dots & 0 \\ 0 & A_{m_2}(a) & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & A_{m_k}(a) \end{Vmatrix}.$$

13. Найти все матрицы, коммутирующие с квазидиагональными матрицами вида

$$\begin{Vmatrix} A_{m_1}(a_1) & 0 & \ldots & 0 \\ 0 & A_{m_2}(a_2) & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & A_{m_k}(a_k) \end{Vmatrix},$$

где все числа $a_1, a_2, \ldots, a_k$ различны.

14. Найти все матрицы, коммутирующие с общей жордановой матрицей (8).

15. При каком условии всякая матрица, коммутирующая с некоторой матрицей A, есть многочлен от A?

16. Показать, что в вещественном пространстве R_n размерности $n \geqslant 2$ каждый линейный оператор имеет инвариантную плоскость.

БИЛИНЕЙНЫЕ И КВАДРАТИЧНЫЕ ФОРМЫ

В этой главе мы будем изучать числовые линейные функции двух векторных аргументов. В отличие от случая линейных числовых функций одного аргумента теория линейных числовых функций двух аргументов — билинейных форм — имеет богатое геометрическое содержание. Положив в выражении билинейной формы второй аргумент равным первому, мы получаем новый важный класс функций одного переменного, уже нелинейных — квадратичных форм.

В §§ 7.1—7.8 рассмотрения ведутся в линейном пространстве K над произвольным числовым полем K, в § 7.9 — в вещественном линейном пространстве.

§ 7.1. Билинейные формы

7.11. Числовая функция $A(x, y)$ от двух векторных аргументов x, y в линейном пространстве K называется *билинейной функцией* или *билинейной формой*, если она является линейной функцией от x при каждом фиксированном значении y и линейной функцией от y при каждом фиксированном значении x.

Иными словами, $A(x, y)$ есть билинейная форма от x и y, если для любых x, y и z из K и любого $\alpha \in K$ удовлетворяются равенства

$$\left.\begin{aligned} A(x+z, y) &= A(x, y) + A(z, y), \\ A(\alpha x, y) &= \alpha A(x, y), \\ A(x, y+z) &= A(x, y) + A(x, z), \\ A(x, \alpha y) &= \alpha A(x, y). \end{aligned}\right\} \qquad (1)$$

Первые два из этих равенств означают линейность функции $A(x, y)$ по первому аргументу, последние два — линейность по второму аргументу.

Из определения билинейной формы, используя равенства (1), легко получим общую формулу

$$\mathrm{A}\left(\sum_{i=1}^{k}\alpha_i x_i,\ \sum_{j=1}^{m}\beta_j y_j\right)=\sum_{i=1}^{k}\sum_{j=1}^{m}\alpha_i\beta_j\,\mathrm{A}(x_i,\ y_j), \qquad (2)$$

в которой $x_1, x_2, \ldots, x_k, y_1, \ldots, y_m$ — произвольные векторы пространства K, а $\alpha_1, \alpha_2, \ldots, \alpha_k, \beta_1, \beta_2, \ldots, \beta_m$ — любые числа из K.

Билинейные формы, заданные в бесконечномерных пространствах, называют обычно *билинейными функционалами*.

7.12. Примеры.

а. Если $L_1(x)$ и $L_2(x)$ — линейные формы, то $\mathrm{A}(x, y)=L_1(x)\cdot L_2(y)$ является, очевидно, билинейной формой от x и y.

б. В n-мерном линейном пространстве с фиксированным базисом $e_1, e_2, \ldots, e_n$ примером билинейной формы является функция

$$\mathrm{A}(x,\ y)=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}\xi_i\eta_k,$$

где $x=\sum_{i=1}^{n}\xi_i e_i$, $y=\sum_{k=1}^{n}\eta_k e_k$ — произвольные векторы и a_{ik} $(i,\ k=1,\ 2,\ \ldots,\ n)$ — фиксированные числа.

7.13. Общий вид билинейной формы в n-мерном линейном пространстве. Пусть в n-мерном линейном пространстве K_n задана билинейная форма $\mathrm{A}(x, y)$. Выберем в K_n произвольный базис $e_1, e_2, \ldots, e_n$. Положим $\mathrm{A}(e_i,\ e_k)=a_{ik}$ $(i,\ k=1,\ 2,\ \ldots,\ n)$. Тогда для любых

$$x=\sum_{i=1}^{n}\xi_i e_i, \quad y=\sum_{k=1}^{n}\eta_k e_k$$

согласно формуле (2)

$$\mathrm{A}(x,\ y)=\mathrm{A}\left(\sum_{i=1}^{n}\xi_i e_i,\quad \sum_{k=1}^{n}\eta_k e_k\right)=$$

$$=\sum_{i=1}^{n}\sum_{k=1}^{n}\xi_i\eta_k\,\mathrm{A}(e_i,\ e_k)=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}\xi_i\eta_k; \qquad (3)$$

Таким образом, в примере 6 мы имели самый общий вид билинейной функции в n-мерном линейном пространстве.

Коэффициенты a_{ik} образуют квадратную матрицу.

$$A = A_{(e)} = \begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Vmatrix} = \| a_{ik} \|,$$

которую мы будем называть *матрицей билинейной формы* $\mathrm{A}(x, y)$ *в базисе* $\{e\} = \{e_1, e_2, \dots, e_n\}$.

7.14. Симметричные билинейные формы. Билинейная форма называется *симметричной*, если для любых векторов x и y

$$\mathrm{A}(x, y) = \mathrm{A}(y, x).$$

Если билинейная форма $\mathrm{A}(x, y)$ в n-мерном пространстве K_n симметрична, то

$$a_{ik} = \mathrm{A}(e_i, e_k) = \mathrm{A}(e_k, e_i) = a_{ki};$$

следовательно, матрица $A_{(e)}$ симметричной билинейной формы в любом базисе $e_1, e_2, \dots, e_n$ пространства K_n совпадает с транспонированной матрицей $A'_{(e)}$. Легко проверить, что верно и обратное: если в некотором базисе $\{e\} = \{e_1, e_2, \dots, e_n\}$ $A'_{(e)} = A_{(e)}$, то форма $\mathrm{A}(x, y)$ симметрична; в самом деле, в этом случае

$$\mathrm{A}(y, x) = \sum_{i,\,k=1}^{n} a_{ik}\eta_i\xi_k = \\ = \sum_{i,\,k=1}^{n} a_{ki}\eta_i\xi_k = \sum_{k,\,i=1}^{n} a_{ik}\xi_i\eta_k = \mathrm{A}(x, y),$$

что и утверждалось.

В частности, мы получаем: если матрица билинейной формы $\mathrm{A}(x, y)$, вычисленная в некотором базисе, совпадает с транспонированной матрицей, то в любом другом базисе пространства K_n матрица этой формы также совпадает с транспонированной.

Матрицу, совпадающую с транспонированной матрицей, мы будем называть в дальнейшем *симметричной*.

7.15. Преобразование матрицы билинейной формы при переходе к новому базису.

а. При переходе к новому базису матрица билинейной формы, разумеется, изменяется; найдем закон ее изменения. Пусть $A_{(e)} = \| a_{ji} \|$ — матрица билинейной формы $\mathrm{A}(x, y)$ в базисе $\{e\} = \{e_1, e_2, \ldots, e_n\}$ и $A_{(f)} = \| b_{ik} \|$ — матрица той же формы в базисе $\{f\} = \{f_1, f_2, \ldots, f_n\}$ ($i, j, k, l = 1, 2, \ldots, n$), и пусть формулы перехода от одного базиса к другому имеют вид

$$f_i = \sum_{j=1}^{n} p_j^{(i)} e_j \qquad (i = 1, 2, \ldots, n)$$

с матрицей перехода $P = \| p_j^{(i)} \|$. В таком случае

$$b_{ik} = \mathrm{A}(f_i, f_k) = \mathrm{A}\left(\sum_{j=1}^{n} p_j^{(i)} e_j, \ \sum_{l=1}^{n} p_l^{(k)} e_l\right) =$$

$$= \sum_{j,\ l=1}^{n} p_j^{(i)} p_l^{(k)} \mathrm{A}(e_j, e_l) = \sum_{j,\ l=1}^{n} p_j^{(i)} p_l^{(k)} a_{jl}.$$

Полученную формулу мы запишем в виде

$$b_{ik} = \sum_{j=1}^{n} \sum_{l=1}^{n} p_i^{(j)\prime} a_{jl} p_l^{(k)}, \tag{4}$$

где $p_i^{(j)\prime} = p_j^{(i)}$ — элемент матрицы P', транспонированной по отношению к матрице P. Формула (4) отвечает следующему соотношению между матрицами *(4.43)*:

$$A_{(f)} = P' A_{(e)} P. \tag{5}$$

б. Так как матрицы P и P' невырождены, то в силу *4.67* ранг матрицы $A_{(f)}$ равен рангу матрицы $A_{(e)}$; следовательно, *ранг матрицы билинейной формы не зависит от выбора базиса.*

Поэтому имеет смысл понятие *ранга билинейной формы*, определяемого как ранг матрицы этой формы в любом базисе пространства K_n.

Если форма $\mathrm{A}(x, y)$ имеет ранг n, равный размерности пространства K_n, она называется *невырожденной формой.*

в. Пусть $\mathrm{A}(x, y)$ — невырожденная форма; покажем, что для каждого вектора $x_0 \neq 0$ существует вектор $y_0 \in \mathrm{K}_n$, для которого $\mathrm{A}(x_0, y_0) \neq 0$. Допустим противное, т. е. что

$A(x_0, y) \equiv 0$ при каждом $y \in K_n$. Построим базис $e_1, \ldots, e_n$ в пространстве K_n так, что $e_1 = x_0$. Тогда в матрице формы $A(x, y)$ в этом базисе мы будем иметь при любом $m = 1, \ldots, n$

$$a_{1m} = A(e_1, e_m) = A(x_0, e_m) = 0,$$

так что вся первая строка матрицы состоит из нулей. Но тогда ранг этой матрицы $< n$, что противоречит предположению о невырожденности формы. Это доказывает утверждение.

г. Заметим, что форма $A(x, y)$, невырожденная во всем пространстве K, может быть вырожденной на подпространстве $K' \subset K$. Так, в пространстве R_2, где $x = (\xi_1, \xi_2)$, $y = (\eta_1, \eta_2)$, форма

$$A(x, y) = \xi_1\eta_1 - \xi_2\eta_2$$

невырождена; однако на подпространстве $R'_2 \subset R_2$, где $\xi_1 = \xi_2$ (и $\eta_1 = \eta_2$), она тождественно равна 0.

д. Для определителей наших матриц в силу теоремы об определителе произведения матриц *(4.75)* получается соотношение

$$\det A_{(f)} = \det A_{(e)} (\det C)^2. \tag{6}$$

§ 7.2. Квадратичные формы

В аналитической геометрии на плоскости одна из основных задач состоит в приведении общего уравнения кривой 2-го порядка к каноническому виду путем перехода к некоторой новой системе координат. Уравнение центральной кривой 2-го порядка с центром в начале координат $x = 0$, $y = 0$, как известно, имеет вид

$$Ax^2 + 2Bxy + Cy^2 = D. \tag{7}$$

Преобразование координат производится по формулам

$$x = a_{11}x' + a_{12}y',$$
$$y = a_{21}x' + a_{22}y',$$

где a_{11}, a_{12}, a_{21}, a_{22} — некоторые числа (обычно синусы и косинусы угла поворота осей). В результате уравнение (7) приобретает более простой вид:

$$A'x'^2 + B'y'^2 = D.$$

Аналогичная задача может быть поставлена в пространстве любого числа измерений. Теория квадратичных форм, излагаемая далее, основной своей целью имеет решение этой задачи и задач, связанных с нею.

7.21. Введем следующее определение.

Квадратичной формой в линейном пространстве K называется функция A(x, x) от одного векторного аргумента $x \in$ K, которая получается из произвольной билинейной формы A(x, y) заменой y на x.

В n-мерном линейном пространстве K_n с фиксированным базисом $\{e\} = \{e_1, e_2, \ldots, e_n\}$ каждая квадратичная форма в силу формулы (2) имеет вид

$$A(x, x) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} \xi_i \xi_k, \tag{8}$$

где $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x относительно базиса $\{e\}$. И наоборот, *если задана функция* A(x, x) *от вектора* x, *определяемая в базисе* $\{e\}$ *формулой* (8), *то эта функция представляет собой квадратичную форму от вектора* x.

Действительно, мы можем ввести билинейную форму

$$B(x, y) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} \xi_i \eta_k,$$

где $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора y относительно базиса $\{e\}$; тогда очевидно, что квадратичная форма B(x, x) совпадает с функцией A(x, x).

7.22. Заметим, что в двойной сумме (8) можно совершить приведение некоторых подобных членов: при $i \neq k$ мы имеем

$$a_{ik} \xi_i \xi_k + a_{ki} \xi_k \xi_i = (a_{ik} + a_{ki}) \xi_i \xi_k = b_{ik} \xi_i \xi_k,$$

где

$$b_{ik} = a_{ik} + a_{ki}.$$

Для $i = k$ полагаем

$$b_{ii} = a_{ii}.$$

В результате двойную сумму можно записать с меньшим числом слагаемых:

$$A(x, x) = \sum_{k=1}^{n} \sum_{i \leqslant k} b_{ik} \xi_i \xi_k.$$

Отсюда следует, что две различные билинейные формы

$$\mathrm{A}(x, y)=\sum_{i, k=1}^{n} a_{ik}\xi_i\eta_k \quad \text{и} \quad \mathrm{C}(x, y)=\sum_{i, k=1}^{n} c_{ik}\xi_i\eta_k$$

можно иногда после замены y на x привести к одной и той же квадратичной форме; для этого нужно только, чтобы имело место равенство $a_{ik}+a_{ki}=c_{ik}+c_{ki}$ для любых i и k.

Таким образом, вообще говоря, нельзя однозначно восстановить по квадратичной форме породившую ее билинейную форму. В одном случае исходная билинейная форма восстанавливается однозначно: именно, если известно, что она была симметричной. В самом деле, если $a_{ik}=a_{ki}$, то из уравнения $a_{ik}+a_{ki}=b_{ik}$ (при $i\neq k$) коэффициенты a_{ik} однозначно определяются:

$$a_{ik}=a_{ki}=\frac{b_{ik}}{2} \tag{9}$$

и при $i=k$

$$a_{ii}=b_{ii},$$

а вместе со всеми a_{ik} однозначно определяется и вся билинейная форма. Это утверждение можно доказать и не прибегая к базису и координатам; в самом деле, по определению билинейной формы

$$\mathrm{A}(x+y,\ x+y)=\mathrm{A}(x,\ x)+\mathrm{A}(x,\ y)+\mathrm{A}(y,\ x)+\mathrm{A}(y,\ y)$$

и при условии симметрии

$$\mathrm{A}(x,\ y)=\frac{1}{2}[\mathrm{A}(x,\ y)+\mathrm{A}(y,\ x)]=$$

$$=\frac{1}{2}[\mathrm{A}(x+y,\ x+y)-\mathrm{A}(x,\ x)-\mathrm{A}(y,\ y)];$$

таким образом, значение билинейной формы $\mathrm{A}(x,\ y)$ для любой пары векторов x, y однозначно определяется по значениям квадратичной формы на векторах x, y и $x+y$.

С другой стороны, чтобы получить из билинейных форм все возможные квадратичные формы, достаточно иметь одни лишь симметричные билинейные формы. В самом деле, если $\mathrm{A}(x,\ y)$ — произвольная билинейная форма, то

$$\mathrm{A}_1(x,\ y)=\frac{1}{2}[\mathrm{A}(x,\ y)+\mathrm{A}(y,\ x)]$$

есть симметричная билинейная форма, и

$$\mathrm{A}_1(x, x) = \frac{1}{2}[\mathrm{A}(x, x) + \mathrm{A}(x, x)] = \mathrm{A}(x, x),$$

т. е. квадратичные формы $\mathrm{A}_1(x, x)$ и $\mathrm{A}(x, x)$ совпадают.

7.23. В силу всех этих соображений при использовании билинейных форм для изучения свойств квадратичных форм достаточно ограничиться рассмотрением одних только симметричных билинейных форм и соответственно симметричных матриц $\|a_{jk}\|$, $a_{jk} = a_{kj}$.

Симметричная матрица $A = \|a_{jk}\|$ симметричной билинейной формы $\mathrm{A}(x, y)$, отвечающей квадратичной форме $\mathrm{A}(x, x)$, называется *матрицей формы* $\mathrm{A}(x, x)$.

При изменении базиса матрица A квадратичной формы $\mathrm{A}(x, x)$, совпадающая с матрицей соответствующей симметричной билинейной формы $\mathrm{A}(x, y)$, меняется так же, как и эта последняя:

$$A_{(f)} = P' A_{(e)} P,$$

где P — матрица перехода от базиса $\{e\}$ к базису $\{f\}$. В частности, *ранг матрицы квадратичной формы не зависит от выбора базиса.* Поэтому можно говорить о ранге квадратичной формы $\mathrm{A}(x, x)$, подразумевая под ним ранг матрицы этой формы в любом базисе пространства K_n. Квадратичная форма ранга n, равного размерности пространства, называется *невырожденной*.

§ 7.3. Приведение квадратичной формы к каноническому виду

7.31. Пусть дана произвольная квадратичная форма $\mathrm{A}(x, x)$ в n-мерном линейном пространстве K_n. Покажем, что *в пространстве K_n существует базис $\{f\} = \{f_1, f_2, \ldots, f_n\}$, в котором для каждого вектора $x = \sum_{k=1}^{n} \eta_k f_k$ значение квадратичной формы $\mathrm{A}(x, x)$ вычисляется по формуле*

$$\mathrm{A}(x, x) = \lambda_1 \eta_1^2 + \lambda_2 \eta_2^2 + \ldots + \lambda_n \eta_n^2, \tag{10}$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ — некоторые фиксированные числа.

Всякий базис, обладающий этим свойством, будем называть *каноническим базисом формы* $A(x, x)$, а выражение (10) — *каноническим видом формы* $A(x, x)$; в частности, числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ будем называть *каноническими коэффициентами формы* $A(x, x)$.

Пусть $\{e_1, e_2, \ldots, e_n\}$ — произвольный базис пространства K_n; если $x = \sum_{k=1}^{n} \xi_k e_k$, то, как мы уже знаем, форма $A(x, x)$ имеет следующий вид:

$$A(x, x) = \sum_{k=1}^{n} \sum_{i \leqslant k} b_{ik} \xi_i \xi_k. \tag{11}$$

Согласно *5.32* наше утверждение будет доказано, если мы сможем написать формулы

$$\left.\begin{aligned} \eta_1 &= p_{11}\xi_1 + p_{12}\xi_2 + \ldots + p_{1n}\xi_n, \\ \eta_2 &= p_{21}\xi_1 + p_{22}\xi_2 + \ldots + p_{2n}\xi_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \eta_n &= p_{n1}\xi_1 + p_{n2}\xi_2 + \ldots + p_{nn}\xi_n \end{aligned}\right\} \tag{11'}$$

с невырожденной матрицей $P = \|p_{ik}\|$ и такие, что, выражая координаты $\{\xi\}$ в формуле (11) через величины $\{\eta\}$, мы преобразуем формулу (11) к виду (10).

Будем вести доказательство индукцией по числу координат, фактически входящих в формулу (11) (т. е. с отличными от нуля коэффициентами).

Предположим, что каждая форма, содержащая $m-1$ координат (например, $\xi_1, \xi_2, \ldots, \xi_{m-1}$), может быть приведена к каноническому виду (10), с $n = m-1$ преобразованием (11′), также с $n = m-1$. Если в формулу (11) фактически входит лишь одна координата, например ξ_1, т. е. формула (11) имеет вид

$$A(x, x) = b_{11}\xi_1^2,$$

то это предположение, очевидно, выполняется (можно взять $p_{11} \neq 0$ произвольно).

Пусть теперь имеется форма (11), фактически содержащая m координат $\xi_1, \xi_2, \ldots, \xi_m$. Допустим вначале, что среди чисел $b_{11}, b_{22}, \ldots, b_{mm}$ имеется число, отличное от нуля; предположим для определенности, что $b_{mm} \neq 0$. Выделим в форме (11) группу членов, содержащих координату ξ_m;

эта группа имеет вид

$$b_{1m}\xi_1\xi_m + b_{2m}\xi_2\xi_m + \ldots + b_{m-1,\,m}\,\xi_{m-1}\xi_m + b_{mm}\xi_m^2 =$$
$$= b_{mm}\left(\frac{b_{1m}}{2b_{mm}}\xi_1 + \frac{b_{2m}}{2b_{mm}}\xi_2 + \ldots + \frac{b_{m-1,\,m}}{2b_{mm}}\xi_{m-1} + \xi_m\right)^2 +$$
$$+ \mathrm{A}_1(x,\ x) \qquad (12)$$

где через $\mathrm{A}_1(x,\ x)$ обозначена квадратичная форма, зависящая только от величин $\xi_1, \xi_2, \ldots, \xi_{m-1}$. Рассмотрим следующее преобразование координат:

$$\begin{aligned}
\tau_1 &= \xi_1,\\
\tau_2 &= \xi_2,\\
&\ldots\ldots\ldots\ldots\ldots\\
\tau_{m-1} &= \xi_{m-1},\\
\tau_m &= \frac{b_{1m}}{2b_{mm}}\xi_1 + \frac{b_{2m}}{2b_{mm}}\xi_2 + \ldots + \frac{b_{m-1,\,m}}{2b_{mm}}\xi_{m-1} + \xi_m.
\end{aligned}$$

Матрица этого преобразования невырождена (ее определитель равен 1). В новых координатах форма $\mathrm{A}(x,\ x)$ очевидно, принимает вид

$$\mathrm{A}(x,\ x) = \mathrm{B}(x,\ x) + b_{mm}\tau_m^2,$$

где квадратичная форма $\mathrm{B}(x,\ x)$ зависит только от величин $\tau_1, \tau_2, \ldots, \tau_{m-1}$. В силу предположения индукции существует новое преобразование

$$\left.\begin{aligned}
\eta_1 &= p_{11}\tau_1 + p_{12}\tau_2 + \ldots + p_{1,\,m-1}\tau_{m-1},\\
\eta_2 &= p_{21}\tau_1 + p_{22}\tau_2 + \ldots + p_{2,\,m-1}\tau_{m-1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\eta_{m-1} &= p_{m-1,\,1}\tau_1 + p_{m-1,\,2}\tau_2 + \ldots + p_{m-1,\,m-1}\tau_{m-1}
\end{aligned}\right\} \qquad (13)$$

с невырожденной матрицей $P = \|p_{ij}\|$, которое приводит форму $\mathrm{B}(x,\ x)$ к каноническому виду

$$\mathrm{B}(x,\ x) = \lambda_1\eta_1^2 + \lambda_2\eta_2^2 + \ldots + \lambda_{m-1}\eta_{m-1}^2.$$

Если мы добавим к равенствам (13) еще одно равенство

$$\eta_m = \tau_m,$$

то получим невырожденное преобразование координат $\tau_1, \tau_2, \ldots, \tau_m$ в координаты $\eta_1, \eta_2, \ldots, \eta_m$, после которого форма $\mathrm{A}(x, x)$ примет канонический вид

$$\mathrm{A}(x, x) = \mathrm{B}(x, x) + b_{mm}\tau_m^2 = \\ = \lambda_1\eta_1^2 + \lambda_2\eta_2^2 + \ldots + \lambda_{m-1}\eta_{m-1}^2 + b_{mm}\eta_m^2.$$

Прямой переход от координат $\{\xi\}$ к координатам $\{\eta\}$ в силу *5.33* осуществляется с помощью матрицы, равной произведению матрицы перехода от координат $\{\tau\}$ к координатам $\{\eta\}$ на матрицу перехода от координат $\{\xi\}$ к координатам $\{\tau\}$. Так как обе эти $m\times m$- матрицы невырождены, то и $m\times m$-матрица-произведение тоже невырождена.

Нам остается рассмотреть случай, когда в форме $\mathrm{A}(x, x)$ с m координатами $\xi_1, \xi_2, \ldots, \xi_m$ все числа a_{11}, $a_{22}, \ldots, a_{mm}$ равны нулю. Рассмотрим один из членов $a_{ij}\xi_i\xi_j$ с отличным от нуля коэффициентом a_{ij}; например, пусть $a_{12}\neq 0$. Произведем следующее преобразование координат (для удобства рассуждения выпишем переход от новых координат к старым):

$$\left.\begin{aligned} \xi_1 &= \xi_1' + \xi_2', \\ \xi_2 &= \xi_1' - \xi_2', \\ \xi_j &= \xi_j' \ (j = 3, \ldots, m). \end{aligned}\right\} \tag{14}$$

Определитель матрицы преобразования (14) равен -2, и, таким образом, это преобразование снова невырожденное. Член $a_{12}\xi_1\xi_2$ преобразуется следующим образом:

$$a_{12}\xi_1\xi_2 = a_{12}\xi_1'^2 - a_{12}\xi_2'^2;$$

поэтому в преобразованной форме появятся сразу два квадрата координат с ненулевыми коэффициентами (очевидно, что сократиться с остающимися членами эти квадраты не смогут, ибо все остающиеся члены содержат координату ξ_i' с $i > 2$). Таким образом, в координатах ξ_i' к форме (11) уже можно применить наш индуктивный метод.

Итак, форма (11), с любым числом $m\leqslant n$ фактически входящих в нее координат ξ_j, приводится к виду (10) преобразованием (11′), где n заменено на m. Дописывая при необходимости равенства $\eta_{m+1} = \xi_{m+1}, \ldots, \eta_n = \xi_n$, мы можем дополнить систему (11′) до требуемой системы из n уравнений с невырожденной матрицей $P = \| p_{ij} \|$, $i, j = 1, \ldots, n$, чем доказательство и завершается.

Идея нашего доказательства — последовательное выделение полных квадратов — может быть применена и для фактического приведения данной квадратичной формы к каноническому виду. В § 7.5 будет описан другой метод, позволяющий непосредственно получить векторы искомого канонического базиса и канонический вид формы.

7.32. Пример. Приведем к каноническому виду форму

$$A(x, x) = \xi_1^2 + 6\xi_1\xi_2 + 5\xi_2^2 - 4\xi_1\xi_3 - 12\xi_2\xi_3 + 4\xi_3^2 - 4\xi_2\xi_4 - 8\xi_3\xi_4 - \xi_4^2.$$

Дополняем группу членов, содержащую ξ_1, до полного квадрата и полагаем

$$\eta_1 = \xi_1 + 3\xi_2 - 2\xi_3.$$

Тогда форма преобразуется к виду

$$A(x, x) = \eta_1^2 - 4\xi_2^2 - 4\xi_2\xi_4 - 8\xi_3\xi_4 - \xi_4^2.$$

Далее дополняем группу членов, содержащую ξ_2, до полного квадрата и полагаем

$$\eta_2 = 2\xi_2 + \xi_4,$$

после чего мы будем иметь

$$A(x, x) = \eta_1^2 - \eta_2^2 - 8\xi_3\xi_4.$$

Квадраты координат ξ_3 и ξ_4 отсутствуют. Поэтому мы полагаем

$$\xi_3 = \eta_3 - \eta_4,$$
$$\xi_4 = \eta_3 + \eta_4,$$

так что $\xi_3\xi_4 = \eta_3^2 - \eta_4^2$.

Таким образом, форма $A(x, x)$ приведена к каноническому виду

$$A(x, x) = \eta_1^2 - \eta_2^2 - 8\eta_3^2 + 8\eta_4^2$$

преобразованием

$$\eta_1 = \xi_1 + 3\xi_2 - 2\xi_3, \quad \eta_2 = 2\xi_2 + \xi_4,$$
$$\eta_3 = \frac{1}{2}\xi_3 + \frac{1}{2}\xi_4, \quad \eta_4 = -\frac{1}{2}\xi_3 + \frac{1}{2}\xi_4,$$

которое, как видно из построения, невырождено.

7.33а. Ни канонический базис, ни канонический вид квадратичной формы не определены однозначно. Например, любая перестановка векторов канонического базиса приводит вновь к каноническому базису. В § 7.5, между прочим, будет показано, что для данной квадратичной формы можно построить канонический базис, взяв первый вектор этого базиса в пространстве произвольно (за некоторыми редкими исключениями). Далее, если форма $A(x, x)$ записана в каноническом виде

$$A(x, x) = \lambda_1\eta_1^2 + \lambda_2\eta_2^2 + \ldots + \lambda_n\eta_n^2$$

($\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора x), то преобразование координат

$$\begin{aligned} \eta_1 &= \alpha_1\tau_1, \\ \eta_2 &= \alpha_2\tau_2, \\ &\cdot\ \cdot\ \cdot\ \cdot \\ \eta_n &= \alpha_n\tau_n \end{aligned}$$

($\alpha_1, \alpha_2, \ldots, \alpha_n$ — фиксированные числа, все отличные от нуля, $\tau_1, \tau_2, \ldots, \tau_n$ — новые координаты) приводит форму $A(x, x)$ к новому виду, также каноническому, но с другими коэффициентами:

$$A(x, x) = (\lambda_1\alpha_1^2)\,\tau_1^2 + (\lambda_2\alpha_2^2)\,\tau_2^2 + \ldots + (\lambda_n\alpha_n^2)\,\tau_n^2.$$

Поэтому встает вопрос об описании всех канонических видов, к которым можно привести данную квадратичную форму. Этот вопрос будет уточнен, если мы сузим определение канонического вида (например, как это будет сделано в вещественном пространстве, *7.93*) или сузим класс допустимых преобразований координат (например, как это будет сделано в евклидовом пространстве (*10.12*)).

б. Заметим, что в приведенном примере число коэффициентов, отличных от 0, осталось неизменным.

Вообще, число ненулевых канонических коэффициентов, очевидно, есть ранг матрицы квадратичной формы в соответствующем каноническом базисе. Поскольку ранг матрицы квадратичной формы не зависит от выбора базиса (*7.23*), *число ненулевых канонических коэффициентов квадратичной формы не зависит от выбора канонического базиса.* Это число, очевидно, совпадает с *рангом квадратичной формы* (*7.23*). Зная матрицу формы $A(x, x)$ в каком-либо базисе $\{e\}$, мы можем предсказать число ее ненулевых канонических коэффициентов — это число есть ранг формы $A(x, x)$, который можно вычислить как ранг матрицы формы $A(x, x)$ в базисе $\{e\}$. В частности, у невырожденной формы (*7.23*) в любом каноническом базисе все ее канонические коэффициенты отличны от 0.

§ 7.4. Канонический базис билинейной формы

7.41а. Вектор x_1 называется *сопряженным* с вектором y_1 относительно билинейной формы $A(x, y)$, если

$$A(x_1, y_1) = 0.$$

б. Пусть $\|a_{jk}\|$ — матрица формы $A(x, y)$ в каком-либо базисе $e_1, \ldots, e_n$. Если $x_1 = \sum \xi_k e_k$, $y_1 = \sum \eta_j e_j$, то условие сопряженности векторов x_1 и y_1 записывается в виде

$$A(x_1, y_1) = \sum_{j,\, k=1}^{n} a_{jk} \xi_j \eta_k = 0.$$

в. *Если векторы $x_1, x_2, \ldots, x_k$ сопряжены с вектором y_1, то также сопряжен с вектором y_1 любой вектор подпространства $L(x_1, x_2, \ldots, x_k)$ — линейной оболочки векторов $x_1, x_2, \ldots, x_k$.*

Действительно, в силу свойств билинейной формы

$$A(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k, y_1) =$$
$$= \alpha_1 A(x_1, y_1) + \alpha_2 A(x_2, y_1) + \ldots + \alpha_k A(x_k, y_1) = 0.$$

Если вектор y_1 сопряжен с каждым вектором некоторого подпространства $K' \subset K$, мы будем называть этот вектор *сопряженным к подпространству* K'.

г. Совокупность K'' всех векторов $y_1 \in K$, сопряженных к подпространству K', очевидно, является, подпространством пространства K. Это подпространство K'' мы будем называть *сопряженным* к K'.

7.42. Базис $e_1, e_2, \ldots, e_n$ пространства K называется *каноническим базисом для билинейной формы* $A(x, y)$, если базисные векторы взаимно сопряжены: $A(e_i, e_k) = 0$ при $i \neq k$.

Пример. В пространстве V_3 в качестве билинейной формы $A(x, y)$ рассмотрим скалярное произведение векторов x и y. Сопряженность векторов относительно этой билинейной формы равнозначна, очевидно, их ортогональности. Каноническим базисом в этом случае служит любой ортогональный базис пространства V_3.

7.43. Матрица билинейной формы в каноническом базисе имеет диагональный вид, так как $a_{ik} = A(e_i, e_k) = 0$ при $i \neq k$. Диагональная матрица совпадает со своей транспонированной, поэтому билинейная форма, обладающая каноническим базисом, должна быть симметричной*). Покажем, *что каждая симметричная билинейная форма* $\mathrm{A}(x, y)$ *обладает каноническим базисом.*

Рассмотрим квадратичную форму $\mathrm{A}(x, x)$, соответствующую данной билинейной форме $\mathrm{A}(x, y)$. Мы знаем, что в пространстве K существует базис $e_1, e_2, \ldots, e_n$, относительно которого квадратичная форма $\mathrm{A}(x, x)$ записывается в каноническом виде

$$\mathrm{A}(x, x) = \sum_{i=1}^{n} \lambda_i \xi_i^2.$$

Соответствующая симметричная билинейная форма $\mathrm{A}(x, y)$ согласно формуле (9) имеет канонический вид

$$\mathrm{A}(x, y) = \sum_{i=1}^{n} \lambda_i \xi_i \eta_i. \tag{15}$$

$\left(\text{где } y = \sum_{i=1}^{n} \eta_i e_i\right)$; ее матрица, следовательно, диагональная. Но это и означает, что базис $e_1, e_2, \ldots, e_n$ является каноническим для формы $\mathrm{A}(x, y)$.

7.44. В аналитической геометрии доказывается, что геометрическое место середин хорд кривой 2-го порядка, параллельных заданному вектору, есть прямая линия. Приведем доказательство этой теоремы. На плоскости x_1, x_2 уравнение кривой 2-го порядка имеет вид

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0,$$

или

$$\mathrm{A}(x, x) + \mathrm{L}(x) + c = 0,$$

где

$$\mathrm{A}(x, x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

есть квадратичная, а

$$\mathrm{L}(x) = b_1x_1 + b_2x_2$$

— линейная форма от вектора $x = (x_1, x_2)$.

*) Вспомним, что симметричность или несимметричность билинейной формы — факт, не зависящий от выбора базиса (*7.14*).

Пусть вектор x определяет середину хорды, параллельной фиксированному вектору e. Это означает, что для некоторого $t \neq 0$ удовлетворяются равенства

$$\left.\begin{aligned} &\mathrm{A}(x+te,\ x+te)+\mathrm{L}(x+te)+c=0,\\ &\mathrm{A}(x-te,\ x-te)+\mathrm{L}(x-te)+c=0. \end{aligned}\right\} \tag{16}$$

Пусть $\mathrm{A}(x, y)$ — симметричная билинейная форма, отвечающая квадратичной форме $\mathrm{A}(x, x)$; тогда равенства (16) можно записать так:

$$\mathrm{A}(x,\ x)+2t\,\mathrm{A}(x,\ e)+t^2\mathrm{A}(e,e)+\mathrm{L}(x)+t\,\mathrm{L}(e)+c=0,$$
$$\mathrm{A}(x,\ x)-2t\,\mathrm{A}(x,\ e)+t^2\mathrm{A}(e,e)+\mathrm{L}(x)-t\,\mathrm{L}(e)+c=0.$$

Вычитая второе из первого и сокращая на $2t$, находим

$$2\mathrm{A}(x,\ e)+\mathrm{L}(e)=0. \tag{17}$$

Полученное уравнение линейно относительно x и, следовательно, определяет прямую в плоскости x_1, x_2, что и доказывает теорему. Если x' — другая точка этой же прямой, так что

$$2\mathrm{A}(x',\ e)+\mathrm{L}(e)=0, \tag{18}$$

то, вычитая (17) из (18), мы получаем

$$\mathrm{A}(x-x',\ e)=0,$$

т. е. вектор e и вектор $x-x'$, определяющий направление полученной прямой, сопряжены друг другу в смысле *7.41* относительно билинейной формы $\mathrm{A}(x, y)$.

7.45. Пусть $e_1, \ldots, e_k$ — канонический базис формы $\mathrm{A}(x, y)$ в k-мерном подпространстве $\mathrm{K}' \subset \mathrm{K}$. Пусть $\varepsilon_1, \ldots, \varepsilon_k$ — соответствующие канонические коэффициенты. Выразим числа $\mathrm{A}(x, e_j)$ через координаты вектора $x \in \mathrm{K}'$. Мы имеем

$$\mathrm{A}(x,\ e_j)=\mathrm{A}\left(\sum_{j=1}^{k}\xi_j e_j,\ e_i\right)=$$
$$=\sum_{j=1}^{k}\xi_j\,\mathrm{A}(e_j,\ e_i)=\xi_i\,\mathrm{A}(e_i,\ e_i)=\varepsilon_i\xi_i,$$

так что числа $\mathrm{A}(x, e_i)$ однозначно определяются координатами вектора x. Если форма $\mathrm{A}(x, y)$ невырождена в подпространстве K', то числа ε_i отличны от нуля; в этом случае справедливо и обратное — значения формы $\mathrm{A}(x, e_i)$ однозначно определяют координаты вектора x.

§ 7.5. Построение канонического базиса по методу Якоби

7.51. Построение канонического базиса, приведенное в *7.31*, имеет тот недостаток, что оно не дает возможности непосредственно, по элементам матрицы $A_{(f)}$ симметричной билинейной формы $\mathrm{A}(x, y)$ в заданном базисе $\{f\} = \{f_1, f_2, \ldots, f_n\}$ указать коэффициенты λ_i и координаты векторов канонического базиса. Метод Якоби, излагаемый далее, позволяет находить эти коэффициенты и координаты векторов искомого канонического базиса. Но при этом на матрицу $A_{(f)}$ мы наложим следующее дополнительное условие: угловые миноры матрицы $A_{(f)}$ до $(n-1)$-го порядка включительно

$$\delta_1 = a_{11},\ \delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ldots$$

$$\ldots, \delta_{n-1} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1,n-1} \\ a_{21} & a_{22} & \ldots & a_{2,n-1} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n-1} \end{vmatrix}$$

все должны быть о т л и ч н ы м и о т н у л я.

7.52. Векторы $e_1, e_2, \ldots, e_n$ мы построим по формулам

$$\left.\begin{aligned} e_1 &= f_1, \\ e_2 &= \alpha_1^{(1)} f_1 + f_2, \\ e_3 &= \alpha_1^{(2)} f_1 + \alpha_2^{(2)} f_2 + f_3, \\ &\cdots\cdots\cdots\cdots \\ e_{k+1} &= \alpha_1^{(k)} f_1 + \alpha_2^{(k)} f_2 + \alpha_3^{(k)} f_3 + \ldots + \alpha_k^{(k)} f_k + f_{k+1}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ e_n &= \alpha_1^{(n-1)} f_1 + \alpha_2^{(n-1)} f_2 + \alpha_3^{(n-1)} f_3 + \ldots + \alpha_{n-1}^{(n-1)} f_{n-1} + f_n, \end{aligned}\right\} \tag{19}$$

где коэффициенты $\alpha_i^{(k)}$ $(i = 1, 2, \ldots, k;\ k = 1, 2, \ldots, n-1)$ еще должны быть определены.

Заметим прежде всего, что переход от векторов $f_1, f_2, \ldots, f_k$ к векторам $e_1, e_2, \ldots, e_k$ совершается при помощи матрицы

$$\begin{Vmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ \alpha_1^{(1)} & 1 & 0 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_1^{(k-1)} & \alpha_2^{(k-1)} & \alpha_3^{(k-1)} & \ldots & \alpha_{k-1}^{(k-1)} & 1 \end{Vmatrix}$$

с определителем, равным единице; поэтому для $k = 1, 2, \ldots, n$ векторы $f_1, f_2, \ldots, f_k$ могут быть линейно выражены через $e_1, e_2, \ldots, e_k$, и, следовательно, линейная оболочка $L[f_1, f_2, \ldots, f_k]$ совпадает с линейной оболочкой $L[e_1, e_2, \ldots, e_k]$.

Коэффициенты $\alpha_i^{(k)}$ $(i = 1, 2, \ldots, k)$ мы подчиним условию, чтобы вектор e_{k+1} был сопряжен с подпространством $L(e_1, e_2, \ldots, e_k)$. Для этого *необходимо и достаточно*, чтобы выполнялись равенства

$$A(e_{k+1}, f_1) = 0, \quad A(e_{k+1}, f_2) = 0, \ldots, A(e_{k+1}, f_k) = 0. \tag{20}$$

Действительно, из условий (20) вытекает, что вектор e_{k+1} сопряжен с линейной оболочкой векторов $f_1, f_2, \ldots, f_k$, которая по доказанному совпадает с линейной оболочкой векторов $e_1, e_2, \ldots, e_k$. Обратно, если вектор e_{k+1} сопряжен с подпространством $L[e_1, e_2, \ldots, e_k]$, то он сопряжен с каждым вектором этого подпространства и, в частности, с векторами $f_1, f_2, \ldots, f_k$; поэтому выполняются равенства (20).

Подставляя в формулы (20) выражение e_{k+1} (19) и пользуясь определением билинейной формы, получаем систему уравнений относительно величин $\alpha_i^{(k)}$ $(i = 1, 2, \ldots, k)$:

$$\left.\begin{aligned} A(e_{k+1}, f_1) &= \alpha_1^{(k)} A(f_1, f_1) + \alpha_2^{(k)} A(f_2, f_1) + \ldots \\ &\qquad \ldots + \alpha_k^{(k)} A(f_k, f_1) + A(f_{k+1}, f_1) = 0, \\ A(e_{k+1}, f_2) &= \alpha_1^{(k)} A(f_1, f_2) + \alpha_2^{(k)} A(f_2, f_2) + \ldots \\ &\qquad \ldots + \alpha_k^{(k)} A(f_k, f_2) + A(f_{k+1}, f_2) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ A(e_{k+1}, f_k) &= \alpha_1^{(k)} A(f_1, f_k) + \alpha_2^{(k)} A(f_2, f_k) + \ldots \\ &\qquad \ldots + \alpha_k^{(k)} A(f_k, f_k) + A(f_{k+1}, f_k) = 0. \end{aligned}\right\} \tag{21}$$

Эта неоднородная система уравнений с коэффициентами $A(f_i, f_j) = a_{ij}$ $(i, j = 1, 2, \ldots, k)$ имеет по условию отличный от нуля определитель и поэтому однозначно разрешима; следовательно, можно определить величины $\alpha_i^{(k)}$ и тем самым построить искомый вектор e_{k+1}.

Для определения всех коэффициентов $\alpha_i^{(k)}$ и всех векторов e_k нужно при каждом k разрешить соответствующую систему (21), следовательно, всего разрешить $n-1$ систем линейных уравнений.

7.53. Обозначим координаты вектора x в построенном базисе $e_1, e_2, \ldots, e_n$ через $\xi_1, \xi_2, \ldots, \xi_n$ и координаты вектора y в этом базисе соответственно через $\eta_1, \eta_2, \ldots, \eta_n$. Билинейная форма $\mathrm{A}(x, y)$ в этом базисе принимает вид

$$\mathrm{A}(x, y) = \sum_{i=1}^{n} \lambda_i \xi_i \eta_i. \tag{22}$$

Чтобы вычислить коэффициенты λ_i, будем рассуждать следующим образом. Рассмотрим нашу билинейную форму $\mathrm{A}(x, y)$ только в подпространстве $\mathrm{L}_m = \mathrm{L}(e_1, e_2, \ldots, e_m)$, где $m \leqslant n$. Форма $\mathrm{A}(x, y)$ в базисе $f_1, f_2, \ldots, f_n$ подпространства L_m имеет, очевидно, матрицу

$$\begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{Vmatrix},$$

а в базисе $e_1, e_2, \ldots, e_m$ — матрицу

$$\begin{Vmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & \lambda_m \end{Vmatrix}.$$

Матрица перехода от базиса $f_1, f_2, \ldots, f_m$ к базису $e_1, e_2, \ldots, e_m$, отвечающая формулам перехода (19), имеет, как мы видели, определитель, равный 1. В силу формулы

7.15 (6) мы должны иметь

$$\det\begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{Vmatrix} = \det\begin{Vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{Vmatrix},$$

или, используя обозначения угловых миноров *7.51*,

$$\delta_m = \lambda_1\lambda_2 \ldots \lambda_m \qquad (m = 1, 2, \ldots, n). \tag{23}$$

Из формулы (23) непосредственно вытекает, что

$$\lambda_1 = \delta_1 = a_{11}, \quad \lambda_2 = \frac{\delta_2}{\delta_1}, \quad \lambda_3 = \frac{\delta_3}{\delta_2}, \ldots, \lambda_n = \frac{\delta_n}{\delta_{n-1}}. \tag{24}$$

Формулы (24) позволяют найти коэффициенты билинейной формы в каноническом базисе, не вычисляя самого базиса.

7.54. Обратим внимание на k-ю формулу в системе (19); запишем ее в форме

$$f_{k+1} = -\alpha_1^{(k)} f_1 - \ldots - \alpha_k^{(k)} f_k + e_{k+1} = g_k + e_{k+1}.$$

В этой формуле вектор g_k лежит в подпространстве $\mathrm{L}(f_1, \ldots, f_k)$, а e_{k+1} сопряжен к этому подпространству. Коэффициенты $\alpha_1^{(k)}, \ldots, \alpha_k^{(k)}$ определяются, причем единственным образом, из системы (21), при условии, что $\det\|\mathrm{A}(f_i, f_j)\| \neq 0$ $(i, j = 1, \ldots, k)$, или, что то же, при условии, что форма $\mathrm{A}(x, y)$ на подпространстве $\mathrm{L}(f_1, \ldots, f_k)$ невырождена. Так как вектор f_{k+1} был в этом построении произвольным, то, обозначая $f = f_{k+1}$, $g = g_k$, $h = e_{k+1}$, $\mathrm{L}(f_1, \ldots, f_k) = \mathrm{K}' \subset \mathrm{K}$, мы приходим к следующей теореме:

Теорема. *Пусть вектор f не принадлежит к подпространству $\mathrm{K}' \subset \mathrm{K}$, на котором форма $\mathrm{A}(x, y)$ невырождена. Тогда существует (и единственно) разложение*

$$f = g + h, \tag{25}$$

где $g \in \mathrm{K}'$, а h сопряжен с пространством K'.

7.55. Обозначим через K'' подпространство, сопряженное к подпространству K' относительно формы $\mathrm{A}(x, y)$.

Наличие разложения (25) вместе с его единственностью показывает, что все пространство K есть прямая сумма подпространств K′ и K″ (*2.45*). Итак, имея подпространство $K' \subset K$, на котором (заданная во всем пространстве K) форма $A(x, y)$ невырождена, можно осуществить прямое разложение

$$K = K' + K'',$$

где подпространство K″ сопряжено с K′ относительно формы $A(x, y)$.

§ 7.6. Сопряженные линейные операторы

7.61. Фиксируем в пространстве K_n билинейную невырожденную симметричную форму, которую будем обозначать в этом параграфе через (x, y). Пусть A и B — линейные операторы, действующие в пространстве K_n. Образуем функции $A(x, y)$ и $B(x, y)$ от двух векторов x, y по формуле

$$A(x, y) = (Ax, y), \quad B(x, y) = (x, By).$$

Проверим, что полученные функции $A(x, y)$ и $B(x, y)$ — билинейные формы. Действительно, на основании определений линейного оператора (*4.21*) и билинейной формы (*7.11*) имеют место равенства

$$A(x_1 + x_2, y) = (A(x_1 + x_2), y) = (Ax_1 + Ax_2, y) =$$
$$= (Ax_1, y) + (Ax_2, y) = A(x_1, y) + A(x_2, y),$$
$$A(\alpha x, y) = (A(\alpha x), y) = (\alpha A(x), y) = \alpha (Ax, y) = \alpha A(x, y),$$

которые показывают, что $A(x, y)$ линейна по первому аргументу. Линейность $A(x, y)$ по второму аргументу следует из того, что (x, y) линейна по y. Таким образом, $A(x, y)$ — билинейная форма. Аналогично и $B(x, y)$ — билинейная форма. Пусть $e_1, \ldots, e_n$ — канонический базис для формы (x, y):

$$(e_j, e_k) = 0 \text{ при } j \neq k, \ (e_m, e_m) = \varepsilon_m \in K, \ \varepsilon_m \neq 0.$$

Сравним в этом базисе матрицу оператора A и формы $A(x, y)$. Матрица $\|a_k^{(j)}\|$ оператора A определяется из формул

$$Ae_j = \sum_{k=1}^{n} a_k^{(j)} e_k, \qquad j = 1, \ldots, n$$

(здесь верхний индекс указывает номер строки, нижний — номер столбца). Матрица $\|a_{jk}\|$ формы $A(x, y)$ (первый индекс — номер строки, второй — номер столбца) определяется из формул

$$a_{jm} = \mathrm{A}(e_j, e_m) = (\mathrm{A}\,e_j, e_m) = \left(\sum_{k=1}^{n} a_k^{(j)} e_k, e_m\right) = \\ = a_m^{(j)}(e_m, e_m) = \varepsilon_m a_m^{(j)}. \quad (26)$$

Следовательно, m-й столбец матрицы $\|a_{jm}\|$ получается умножением m-го столбца матрицы $\|a_m^{(j)}\|$ при каждом $m = 1, \ldots, n$ на канонический коэффициент ε_m формы (x, y). Для матрицы $\|b_k^{(j)}\|$ оператора B (в том же базисе $e_1, \ldots, e_n$) и матрицы $\|b_{jk}\|$ формы $\mathrm{B}(x, y)$ аналогично получаем

$$b_{jm} = \mathrm{B}(e_j, e_m) = (e_j, \mathrm{B}e_m) = \\ = \left(e_j, \sum_{k=1}^{n} b_k^{(m)} e_k\right) = b_j^{(m)}(e_j, e_j) = \varepsilon_j b_j^{(m)}, \quad (27)$$

т. е. j-я строка матрицы $\|b_{jm}\|$ получается умножением j-го столбца матрицы оператора B при каждом $j = 1, \ldots, n$ на соответствующий канонический коэффициент ε_j.

7.62. Пусть, обратно, в пространстве K_n заданы билинейные формы $\mathrm{A}(x, y)$ и $\mathrm{B}(x, y)$; мы утверждаем, что существуют линейные операторы A и B такие, что

$$\mathrm{A}(x, y) = (\mathrm{A}x, y), \quad \mathrm{B}(x, y) = (x, \mathrm{B}y),$$

причем они определены единственным образом.

Для доказательства зададим операторы A и B в каком-либо базисе $e_1, \ldots, e_n$ матрицами соответственно из чисел

$$a_j^{(m)} = \frac{1}{\varepsilon_m}\mathrm{A}(e_j, e_m), \quad b_j^{(m)} = \frac{1}{\varepsilon_j}\mathrm{B}(e_j, e_m).$$

По этим операторам построим формы $\mathrm{A}_1(x, y) = (\mathrm{A}x, y)$ и $\mathrm{B}_1(x, y) = (x, \mathrm{B}y)$. По доказанному, в базисе $e_1, \ldots, e_n$ матрица формы $\mathrm{A}_1(x, y)$ совпадает с матрицей формы $\mathrm{A}(x, y)$, а матрица формы $\mathrm{B}_1(x, y)$ совпадает с матрицей формы $\mathrm{B}(x, y)$. Но тогда для любых x и y из K_n

$$(\mathrm{A}x, y) \equiv \mathrm{A}_1(x, y) \equiv \mathrm{A}(x, y), \quad (x, \mathrm{B}y) \equiv \mathrm{B}_1(x, y) \equiv \mathrm{B}(x, y),$$

так что операторы A и B удовлетворяют требуемым условиям. Для доказательства единственности нам достаточно проверить, что если некоторый оператор A удовлетворяет условию

$$(Ax, y) \equiv 0 \text{ (при любых } x, y \text{ из } K_n), \qquad (28)$$

то

$$Ax \equiv 0 \qquad (29)$$

при каждом $x \in K_n$, т. е. A — нулевой оператор.

Пусть для некоторого $x_0 \in K_n$ мы имеем $Ax_0 \neq 0$. Тогда, поскольку форма (x, y) невырождена, в силу *7.15в* существует вектор $y_0 \in K_n$ такой, что $(Ax_0, y_0) \neq 0$; это противоречит (28), откуда и следует требуемая единственность.

7.63. Теперь мы установим следующую важную теорему.

Теорема. *Если в пространстве K_n выделена невырожденная билинейная симметричная форма (x, y), то для каждого линейного оператора A существует и единствен линейный оператор A′, удовлетворяющий уравнению*

$$(Ax, y) = (x, A'y)$$

при любых x и y из K_n. Матрица оператора A′ в любом каноническом базисе формы (x, y) получается из матрицы оператора A транспонированием, умножением m-й строки на ε_m и делением j-го столбца на ε_j $(j, m = 1, \ldots, n)$.

Доказательство. По заданному оператору A построим форму $A(x, y) = (Ax, y)$ и затем определим оператор A′ из уравнения

$$(Ax, y) \equiv A(x, y) = (x, A'y).$$

Существование и единственность оператора A′ следуют из *7.62*. Матрица $\| a_m^{(j)} \|$ оператора A, матрица $\| a_{jm} \|$ формы $A(x, y)$ и матрица $\| a_j'^{(m)} \|$ оператора A′ связаны в любом каноническом базисе формы (x, y) формулами (26)—(27)

$$a_m^{(j)} = \frac{a_{jm}}{\varepsilon_m}, \qquad a_j'^{(m)} = \frac{a_{jm}}{\varepsilon_j},$$

так что

$$a_j'^{(m)} = \frac{a_{jm}}{\varepsilon_j} = \frac{\varepsilon_m}{\varepsilon_j} a_m^{(j)}, \qquad (30)$$

и теорема полностью доказана.

Оператор A' называется *сопряженным к оператору* A *относительно формы* (x, y).

7.64. Операция перехода к сопряженному оператору $A \longrightarrow A'$ обладает следующими свойствами:

а) $(A')' = A$ для любого оператора A;

б) $(A + B)' = A' + B'$ для любых A и B;

в) $(\lambda A)' = \lambda A'$ для любого оператора A и любого $\lambda \in K$;

г) $(AB)' = B'A'$ для любых A и B.

Равенство а) вытекает из определения $(A')'$:

$$(x, (A')' y) = (A'x, y) = (x, Ay)$$

и единственности оператора, определяемого билинейной формой *(7.62)*. Аналогично получаются и остальные утверждения; так, б) вытекает из соотношений

$$(x, (A + B)' y) = ((A + B) x, y) = (Ax, y) + (Bx, y) = \\ = (x, A'y) + (x, B'y) = (x, (A' + B') y);$$

в) вытекает из соотношений

$$(x, (\lambda A)' y) = (\lambda Ax, y) = \lambda (Ax, y) = \lambda (x, A'y) = (x, \lambda A'y),$$

и, наконец, г) — из соотношений

$$(x, (AB)' y) = (ABx, y) = (Bx, A'y) = (x, B'A'y).$$

7.65. Отметим еще одну связь между операторами A и A'. Пусть подпространство $K' \subset K_n$ инвариантно относительно оператора A; это означает *(4.81)*, что каждый вектор $x \in K'$ переводится оператором A снова в вектор из того же подпространства K'. Пусть K'' — подпространство, сопряженное к K' *(7.41г)*. Покажем, что оператор A' инвариантен относительно подпространства K''. Пусть $y \in K''$, так что $(y, x) = 0$ для любого $x \in K'$. Мы имеем $(A'y, x) = (y, Ax) = 0$, так как Ax вместе с x лежит в K'; но это означает, что вектор $A'y$ сопряжен всем векторам $x \in K'$ и входит тем самым в K'', что и требуется.

§ 7.7. Изоморфизм пространств с выделенной билинейной формой

7.71. Определение. Два линейных пространства K′ и K″ над одним и тем же числовым полем K с выделенными в них билинейными симметричными формами $\mathrm{A}(x', y')$ и $\mathrm{A}(x'', y'')$ называются *А-изоморфными*, если они изоморфны как линейные пространства над полем K *(2.71)*,— т. е. существует взаимно однозначное отображение (морфизм) $\omega x' = x''$, сохраняющее линейные операции,— и, кроме того, для соответствующих друг другу пар элементов $\omega x' = x''$, $\omega y' = y''$ значения форм $\mathrm{A}(x', y')$ и $\mathrm{A}(x'', y'')$ совпадают:

$$\mathrm{A}(\omega x', \omega y') = \mathrm{A}(x', y').$$

7.72. Имеет место следующая теорема:

Теорема. *Для того чтобы конечномерные линейные пространства* K′ *и* K″, *снабженные формами* $\mathrm{A}(x', y')$ *и* $\mathrm{A}(x'', y'')$, *были* А-*изоморфны, необходимо и достаточно, чтобы* K′ *и* K″ *были равной размерности* n *и чтобы в* K′ *и* K″ *существовали канонические базисы форм* $\mathrm{A}(x', y')$ *и* $\mathrm{A}(x'', y'')$ *с одинаковым набором канонических коэффициентов* $\varepsilon_1, \ldots, \varepsilon_n$.

Доказательство. Предположим, что пространства K′ и K″, снабженные формами $\mathrm{A}(x', y')$ и $\mathrm{A}(x'', y'')$, являются А-изоморфными. Тогда они изоморфны как линейные пространства и, следовательно, имеют равную размерность *(2.71)*. Возьмем произвольный канонический базис $e_1, \ldots, e_n$ формы $\mathrm{A}(x', y')$ в пространстве K′. Мы имеем

$$\mathrm{A}(e_i', e_j') = \begin{cases} 0 & \text{при } i \neq j, \\ \varepsilon_i & \text{при } i = j. \end{cases}$$

Пусть $e_1'', \ldots, e_n''$ — векторы в K″, отвечающие векторам $e_1', \ldots, e_n'$ пространства K′ в силу имеющегося А-изоморфизма. По условию

$$\mathrm{A}(e_i'', e_j'') = \begin{cases} 0 & \text{при } i \neq j, \\ \varepsilon_i & \text{при } i = j. \end{cases}$$

Мы видим, что у формы $\mathrm{A}(x'', y'')$ имеется канонический базис в пространстве K″ с тем же набором канонических коэффициентов $\varepsilon_1, \ldots, \varepsilon_n$, что и у формы $\mathrm{A}(x', y')$. Таким образом, необходимость условия теоремы установлена.

Для доказательства достаточности рассмотрим канонические базисы $e_1', \ldots, e_n'$ (в K') и $e_1'', \ldots, e_n''$ (в K'') с одинаковым набором канонических коэффициентов, так что

$$\mathrm{A}(e_i', e_j') = \mathrm{A}(e_i'', e_j'') = \begin{cases} 0 & \text{при } i \neq j, \\ \varepsilon_i & \text{при } i = j. \end{cases}$$

Положим для любого $x' = \sum \xi_j e_j' \in \mathrm{K}'$

$$x'' = \omega(x') = \sum \xi_j e_j'' \in \mathrm{K}''$$

с теми же координатами ξ_j. Такое соответствие определит изоморфизм пространств K' и K'' *(2.74)*. Мы имеем, далее, при $y' = \sum \eta_k e_k'$, $y'' = \omega(y') = \sum \eta_k e_k''$

$$\mathrm{A}(x', x'') = \sum \varepsilon_i \xi_k \eta_k = \mathrm{A}(y', y''),$$

так что изоморфизм ω является и А-изоморфизмом. Теорема доказана.

7.73. Пусть имеется n-мерное пространство K_n с фиксированной невырожденной билинейной формой $\mathrm{A}(x, y)$. Рассмотрим А-*изоморфизм пространства* K_n, т. е. линейное обратимое отображение $y \to \mathrm{Q}x$, не изменяющее форму $\mathrm{A}(x, y)$:

$$\mathrm{A}(\mathrm{Q}x, \mathrm{Q}y) = \mathrm{A}(x, y). \tag{31}$$

Будем далее форму $\mathrm{A}(x, y)$ короче обозначать через (x, y).

Обозначая через Q' оператор, сопряженный к оператору Q относительно формы (x, y) *(7.63)*, имеем

$$(\mathrm{Q}x, \mathrm{Q}y) = (\mathrm{Q}'\mathrm{Q}x, y). \tag{32}$$

Так как оператор Q' невырожден вместе с оператором Q, из (32) следует

$$\mathrm{Q}'\mathrm{Q} = \mathrm{E}, \tag{33}$$

т. е. Q' является обратным оператором для оператора Q.

Обратно, из (33) следует (32) и далее (31), так что равенство (33) полностью определяет класс операторов, не изменяющих форму (x, y). Эти операторы будем называть *инвариантными относительно формы* (x, y).

7.74. Вместе с оператором Q обратный оператор $Q^{-1} = Q'$ также является инвариантным, так как для любых x и y

$$(Q'x,\ Q'y) = (QQ'x,\ y) = (x,\ y).$$

Произведение двух инвариантных операторов Q и T также является инвариантным оператором, так как для любых x и y

$$(QTx,\ QTy) = (Tx,\ Ty) = (x,\ y).$$

7.75. Применяя инвариантный оператор Q к векторам $e_1, \ldots, e_n$ канонического базиса формы (x, y) с коэффициентами $\varepsilon_1, \ldots, \varepsilon_n$, мы получаем векторы

$$f_1 = Qe_1, \ \ldots, \ f_n = Qe_n. \tag{34}$$

При этом

$$(f_j,\ f_k) = (Qe_j,\ Qe_k) = (e_j,\ e_k) = \begin{cases} \varepsilon_j & \text{при } k = j, \\ 0 & \text{при } k \neq j. \end{cases}$$

Таким образом, $f_1, \ldots, f_n$ *есть также канонический базис формы* (x, y) *с теми же каноническими коэффициентами* $\varepsilon_1, \ldots, \varepsilon_n$.

Обратно, если $f_1, \ldots, f_n$ есть канонический базис формы (x, y) с теми же каноническими коэффициентами $\varepsilon_1, \ldots, \varepsilon_n$, что и в базисе $e_1, \ldots, e_n$, то оператор Q, определенный по формулам (34), является инвариантным. Действительно, мы имеем

$$(Qe_j,\ Qe_k) = (f_j,\ f_k) = (e_j,\ e_k) = \begin{cases} \varepsilon_j & \text{при } k = j, \\ 0 & \text{при } k \neq j. \end{cases}$$

Мы видим, что равенство (31) справедливо для любой пары базисных векторов. Отсюда, по линейности, оно справедливо для *любой* пары векторов x, y из K_n, что и требуется.

Итак, инвариантный оператор Q характеризуется тем, что *всякий канонический базис пространства* K_n (относительно формы (x, y)) *он переводит снова в канонический базис с теми же самыми каноническими коэффициентами.*

7.76. Найдем условия, определяющие матрицу любого инвариантного оператора в каноническом базисе формы (x, y). Пусть $e_1, \ldots, e_n$ — такой базис и $\varepsilon_1, \ldots, \varepsilon_n$ — соответствующие канонические коэффициенты. Пусть, далее,

$Q = \| q_i^{(j)} \|$ — матрица оператора Q в этом базисе. Матрица сопряженного оператора имеет вид (7.63)

$$Q' = \| q_i^{\prime(j)} \|, \quad q_i^{\prime(j)} = \frac{\varepsilon_j}{\varepsilon_i} q_j^{(i)}.$$

Равенство (33) теперь можно записать в координатной форме

$$\sum_{i=1}^{n} q_i^{\prime(j)} q_k^{(i)} = \sum_{j=1}^{n} \frac{\varepsilon_j}{\varepsilon_i} q_j^{(i)} q_k^{(i)} = \delta_j^{(k)} = \begin{cases} 1 \text{ при } j = k, \\ 0 \text{ при } j \neq k. \end{cases}$$

Иначе говоря,

$$\sum_{j=1}^{n} \frac{1}{\varepsilon_i} q_j^{(i)} q_k^{(i)} = \begin{cases} \dfrac{1}{\varepsilon_j} \text{ при } j = k, \\ 0 \text{ при } j \neq k. \end{cases} \tag{35}$$

Равенство (35) эквивалентно (33) и поэтому также может служить определением инвариантного оператора Q.

Итак, матрица инвариантного оператора в любом каноническом базисе формы (x, y) (*инвариантная матрица*) характеризуется тем, что сумма квадратов элементов ее j-го столбца с коэффициентами $\varepsilon_1^{-1}, \ldots, \varepsilon_n^{-1}$ равна числу ε_j^{-1} $(j = 1, \ldots, n)$, а сумма произведений соответствующих элементов двух различных столбцов с коэффициентами $\varepsilon_1, \ldots, \varepsilon_n$ равна 0. Так как из (33) следует также $QQ' = E$, то имеют место и равенства

$$\sum_{k=1}^{m} q_k^{(j)} q_m^{(k)} = \sum_{k=1}^{m} \frac{\varepsilon_k}{\varepsilon_m} q_k^{(j)} q_k^{(m)} = \delta_m^{(j)},$$

или

$$\sum_{k=1}^{m} \varepsilon_k q_k^{(j)} q_k^{(m)} = \begin{cases} \varepsilon_m \text{ при } j = m, \\ 0 \text{ при } j \neq m. \end{cases} \tag{35'}$$

Мы получаем вторую характеристику инвариантной матрицы: сумма квадратов элементов ее j-й строки с коэффициентами $\varepsilon_1, \ldots, \varepsilon_n$ равна числу ε_j, сумма произведений соответствующих элементов двух различных строк также с коэффициентами $\varepsilon_1, \ldots, \varepsilon_n$ равна 0,

§ 7.8*. Полилинейные формы

7.81. По аналогии с билинейными формами можно рассматривать линейные функции от большего числа векторов (трех, четырех и более). Все они называются полилинейными формами.

Определение. Функция $A(x_1, \ldots, x_k)$ от k векторных аргументов $x_1, \ldots, x_k$, меняющихся в линейном пространстве K, называется *полилинейной* (точнее, *k-линейной*) формой, если она линейна по любому аргументу x_j $(j=1, \ldots, k)$ при фиксированных значениях остальных аргументов $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k$.

Полилинейная форма $A(x_1, x_2, \ldots, x_k)$ называется *симметричной*, если она не изменяется при перемене местами любых двух своих аргументов, и *антисимметричной*, если при перемене местами двух своих аргументов она изменяет знак.

Примером полилинейной антисимметричной формы от трех векторов x, y, z (трилинейной формы) в пространстве V_3 является смешанное произведение этих векторов.

Примером полилинейной антисимметричной формы от n векторов в пространстве K_n

$$\begin{aligned} x_1 &= \{a_{11}, a_{12}, \ldots, a_{1n}\}, \\ x_2 &= \{a_{21}, a_{22}, \ldots, a_{2n}\}, \\ &\cdots\cdots\cdots\cdots \\ x_n &= \{a_{n1}, a_{n2}, \ldots, a_{nn}\} \end{aligned}$$

является определитель

$$A(x_1, x_2, \ldots, x_n) = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}. \tag{36}$$

Несколько более общим примером является произведение определителя (36) на фиксированное число $\alpha \in K$.

7.82. Покажем, что *всякая полилинейная антисимметричная форма* $A(x_1, x_2, \ldots, x_n)$ *от n векторов $x_1, x_2, \ldots, x_n$ в n-мерном линейном пространстве K_n с фиксированным базисом $e_1, e_2, \ldots, e_n$ равна определителю* (36) *с некоторым постоянным множителем* $\alpha \in K$.

Обозначим через α величину $\mathrm{A}(e_1, e_2, \ldots, e_n)$. Тогда легко подсчитать величину $\mathrm{A}(e_{i_1}, e_{i_2}, \ldots, e_{i_n})$, где $i_1, i_2, \ldots, i_n$ — любые целые числа от 1 до n. Если среди этих чисел имеются два равных, то величина $\mathrm{A}(e_{i_1}, e_{i_2}, \ldots, e_{i_n})$ равна нулю, ибо при перестановке соответствующих аргументов она не изменяется, в то время как по свойству антисимметрии она должна изменить знак. Если все числа $i_1, i_2, \ldots, i_n$ различны, то, произведя столько перестановок соседних аргументов, сколько беспорядков в последовательности индексов $i_1, i_2, \ldots, i_n$ *), — обозначим это число через N, — можно добиться нормального расположения аргументов; отсюда

$$\mathrm{A}(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = (-1)^N \alpha.$$

Пусть теперь

$$x_i = \sum_{j=1}^{n} a_{ij} e_j \quad (i = 1, 2, \ldots, n)$$

есть произвольная система n векторов пространства K_n. Составим полилинейную форму $\mathrm{A}(x_1, x_2, \ldots, x_n)$:

$$\begin{aligned}
\mathrm{A}(x_1, x_2, \ldots, x_n) &= \\
&= \mathrm{A}\left(\sum_{i_1=1}^{n} a_{1i_1} e_{i_1}, \sum_{i_2=1}^{n} a_{2i_2} e_{i_2}, \ldots, \sum_{i_n=1}^{n} a_{ni_n} e_{i_n}\right) = \\
&= \sum_{i_1, i_2, \ldots, i_n=1}^{n} a_{1i_1} a_{2i_2} \ldots a_{ni_n} \mathrm{A}(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) = \\
&= \alpha \sum_{i_1, i_2, \ldots, i_n=1}^{n} (-1)^N a_{1i_1} a_{2i_2} \ldots a_{ni_n}.
\end{aligned}$$

Так как в каждом слагаемом получившейся суммы N означает число беспорядков в расположении вторых индексов элементов при нормальном порядке первых индексов, то каждое слагаемое есть один из членов определителя (36) с положенным этому члену знаком. Сумма всех этих членов равна поэтому самому определителю (36). Таким образом, наше утверждение доказано.

---

*) Ср. доказательство теоремы *4.54*.

В частности, мы показали, что смешанное произведение трех векторов x, y, z пространства V_3 в любом базисе записывается как определитель 3-го порядка из координат этих векторов с коэффициентом, равным смешанному произведению базисных векторов.

§ 7.9. Квадратичные и билинейные формы в вещественном пространстве

7.91. У вещественных чисел определены знаки ($+$ или $-$), поэтому для билинейных и квадратичных форм в вещественном пространстве теория может быть продвинута несколько далее, чем в пространстве над произвольным полем K.

Согласно общей теореме *7.31* квадратичная форма $\mathrm{A}(x,x)$ в некотором базисе приводится к каноническому виду

$$\mathrm{A}(x,\ x)=\lambda_1\eta_1^2+\ldots+\lambda_n\eta_n^2.$$

Среди чисел $\lambda_1, \ldots, \lambda_n$ имеется столько отличных от нуля, каков ранг формы $\mathrm{A}(x,\ x)$ (*7.33 б*). Они положительны или отрицательны. Оказывается, что число положительных и число отрицательных канонических коэффициентов также не изменяются при изменении канонического базиса.

Теорема. (Теорема инерции квадратичных форм.) *Число положительных и число отрицательных коэффициентов в каноническом виде квадратичной формы* $\mathrm{A}(x,\ x)$ *являются инвариантами формы* (*т. е. не зависят от выбора канонического базиса*).

Доказательство. Пусть задана квадратичная форма $\mathrm{A}(x,\ x)$. В некотором базисе $\{e\}=\{e_1, e_2, \ldots, e_n\}$ она имеет вид

$$\mathrm{A}(x,\ x)=\sum_{i,\ k=1}^{n} a_{ik}\xi_i\xi_k,$$

где $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x относительно базиса $\{e\}$. Допустим, что она обладает двумя каноническими базисами $\{f\}=\{f_1, f_2, \ldots, f_n\}$ и $\{g\}=\{g_1, g_2, \ldots, g_n\}$. Обозначим через $\eta_1, \eta_2, \ldots, \eta_n$ координаты вектора x в базисе $\{f\}$ и через $\tau_1, \tau_2, \ldots, \tau_n$ — в базисе $\{g\}$. Соответствующие формулы преобразования координат пусть будут следующими:

$$\left.\begin{array}{ll}\eta_1 = b_{11}\xi_1 + b_{12}\xi_2 + \ldots + b_{1n}\xi_n; & \tau_1 = c_{11}\xi_1 + c_{12}\xi_2 + \ldots + c_{1n}\xi_n,\\ \eta_2 = b_{21}\xi_1 + b_{22}\xi_2 + \ldots + b_{2n}\xi_n; & \tau_2 = c_{21}\xi_1 + c_{22}\xi_2 + \ldots + c_{2n}\xi_n,\\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\\ \eta_n = b_{n1}\xi_1 + b_{n2}\xi_2 + \ldots + b_{nn}\xi_n; & \tau_n = c_{n1}\xi_1 + c_{n2}\xi_2 + \ldots + c_{nn}\xi_n.\end{array}\right\} \quad (37)$$

В базисе $\{f\}$ форма $\mathrm{A}(x, x)$ имеет вид

$$\mathrm{A}(x, x) = \alpha_1\eta_1^2 + \ldots + \alpha_k\eta_k^2 - \alpha_{k+1}\eta_{k+1}^2 - \ldots - \alpha_m\eta_m^2, \quad (38)$$

а в базисе $\{g\}$ соответственно

$$\mathrm{A}(x, x) = \\ = \beta_1\tau_1^2 + \beta_2\tau_2^2 + \ldots + \beta_p\tau_p^2 - \beta_{p+1}\tau_{p+1}^2 - \ldots - \beta_q\tau_q^2. \quad (39)$$

Числа $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_q$ предполагаются положительными. Мы покажем, что $k = p$, $m = q$. Приравнивая правые части равенств (38) и (39) и перенося отрицательные члены в противоположные части равенства, мы получаем

$$\alpha_1\eta_1^2 + \alpha_2\eta_2^2 + \ldots + \alpha_k\eta_k^2 + \beta_{p+1}\tau_{p+1}^2 + \ldots + \beta_q\tau_q^2 = \\ = \alpha_{k+1}\eta_{k+1}^2 + \ldots + \alpha_m\eta_m^2 + \beta_1\tau_1^2 + \ldots + \beta_p\tau_p^2. \quad (40)$$

Допустим, что $k < p$. Рассмотрим тогда векторы x, удовлетворяющие условиям:

$$\left.\begin{array}{c}\eta_1 = 0, \quad \eta_2 = 0, \ \ldots, \ \eta_k = 0,\\ \tau_{p+1} = 0, \ \ldots, \ \tau_q = 0, \quad \tau_{q+1} = 0, \ \ldots, \ \tau_n = 0.\end{array}\right\} \quad (41)$$

Этих условий, очевидно, меньше чем n, так как $k < p$. Подставляя выражения $\eta_1, \ldots, \eta_k, \tau_{p+1}, \ldots, \tau_n$ через координаты $\{\xi\}$ по формулам (37), мы получаем однородную систему линейных уравнений относительно координат $\{\xi\}$; число уравнений меньше числа неизвестных, и, следовательно, эта однородная система допускает ненулевое решение $x = \{\xi_1, \xi_2, \ldots, \xi_n\}$. Но, с другой стороны, всякий вектор x, удовлетворяющий условиям (41), в силу равенства (40) удовлетворяет и условиям

$$\tau_1 = \tau_2 = \ldots = \tau_p = 0.$$

Вектор, для которого $\tau_1 = \tau_2 = \ldots = \tau_p = \tau_{p+1} = \ldots = \tau_n = 0$, необходимо есть нуль-вектор, и для него все координаты $\{\xi\}$ также должны быть равны нулю. Полученное противоречие показывает, что предположение $k < p$ не может осуществиться.

В силу полной симметрии чисел k и p в рассматриваемой задаче также не может осуществиться и предположение $p < k$. Отсюда $k = p$. Далее, если рассмотреть условия

$$\tau_1 = 0, \quad \tau_2 = 0, \ \ldots, \ \tau_p = 0,$$
$$\eta_{k+1} = 0, \ \ldots, \ \eta_m = 0, \quad \eta_{m+1} = 0, \ \ldots, \ \eta_n = 0,$$

тем же приемом можно опровергнуть предположение $m < q$ и по симметрии $q < m$. Таким образом, окончательно получаем, что $k = p$, $m = q$, что и требовалось.

7.92. Полное число членов, входящих в канонический вид квадратичной формы $A(x, x)$, т. е. ее ранг *(7.33б)*, называется также ее *индексом инерции*; число положительных членов называется *положительным индексом инерции*, число отрицательных членов — *отрицательным индексом инерции*. Если при этом положительный индекс инерции равен размерности пространства, форма называется *положительно определенной*; иными словами, квадратичная форма положительно определенная, если все n ее канонических коэффициентов положительны. Тем самым *положительно определенная форма в каждой точке пространства, кроме начала координат, принимает положительное значение*. И обратно, если некоторая квадратичная форма в n-мерном пространстве принимает всюду, кроме начала координат, положительные значения, то ее ранг равен n и положительный индекс инерции также равен n, т. е. форма положительно определенная. Действительно, для формы ранга, меньшего n, или имеющей меньшее, чем n, число положительных канонических коэффициентов, легко указать точки в пространстве, отличные от начала координат, где эта форма принимает отрицательное значение или нулевое.

Например, форма ранга 2 в трехмерном пространстве

$$A(x, x) = \xi_1^2 + \xi_3^2$$

принимает нулевое значение на любом векторе с координатами $\xi_1 = 0$, $\xi_2 \neq 0$, $\xi_3 = 0$. Форма ранга 3 в трехмерном пространстве

$$A(x, x) = \xi_1^2 - \xi_2^2 + \xi_3^2$$

принимает на тех же векторах отрицательные значения.

7.93. Теорема инерции, доказанная нами для квадратичных форм, непосредственно переносится и на симметричные билинейные формы: именно, *число положительных и отрицательных коэффициентов канонического вида* (22) *билинейной формы* $\mathrm{A}(x, y)$ *не зависит от выбора канонического базиса*. Поэтому для симметричной билинейной формы имеют смысл понятия положительного и отрицательного индексов инерции. Значения положительного и отрицательного индексов инерции билинейной формы $\mathrm{A}(x, y)$ и, следовательно, квадратичной формы $\mathrm{A}(x, x)$ могут быть определены по знакам угловых миноров матрицы формы в каком-либо базисе (если только эти миноры отличны от нуля) по формулам (24).

Заметим, что *в вещественном пространстве* R_n *всегда можно для данной квадратичной формы* $\mathrm{A}(x, x)$ *найти канонический базис такой, что все канонические коэффициенты будут числами* $+1$ *или* -1. Для этого, приведя форму $\mathrm{A}(x, x)$ к виду $\mathrm{A}(x, x)=\lambda_1\tau_1^2+\ldots+\lambda_p\tau_p^2-\mu_1\tau_{p+1}^2-\ldots-\mu_q\tau_{p+q}^2$, где $\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q$ положительны, мы совершим еще одно преобразование координат

$$\xi_1=\sqrt{\lambda_1}\tau_1, \ \ldots, \ \xi_p=\sqrt{\lambda_p}\tau_p,$$
$$\xi_{p+1}=\sqrt{\mu_1}\tau_{p+1}, \ \ldots, \ \xi_{p+q}=\sqrt{\mu_q}\tau_{p+q}.$$

После этого преобразования форма принимает вид

$$\mathrm{A}(x, x)=\xi_1^2+\ldots+\xi_p^2-\xi_{p+1}^2-\ldots-\xi_{p+q}^2.$$

Отсюда легко следует, что *числа p и q в вещественном пространстве — единственные инварианты квадратичной формы* $\mathrm{A}(x, x)$ *и соответствующей билинейной формы* $\mathrm{A}(x, y)$:

Теорема. *Два пространства R' и R'' с выделенными в них симметричными билинейными формами $\mathrm{A}(x', y')$ и $\mathrm{A}(x'', y'')$ являются A-изоморфными тогда и только тогда, когда пространства R' и R'' имеют одинаковую размерность, а индексы инерции p', q' формы $\mathrm{A}(x', y')$ совпадают соответственно с индексами инерции p'', q'' формы $\mathrm{A}(x'', y'')$.*

Доказательство непосредственно следует из приведенных выше соображений и теоремы *7.72*.

7.94. Следующие определения соответствуют определениям, приведенным в *7.92* для квадратичных форм.

Билинейная форма $\mathrm{A}(x, y)$ называется *невырожденной*, если ее ранг равен числу измерений пространства, иными словами, если в канонической записи формы $\mathrm{A}(x, y)$

$$\mathrm{A}(x, y) = \lambda_1 \xi_1 \eta_1 + \ldots + \lambda_n \xi_n \eta_n$$

все коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_n$ отличны от нуля. Если все эти коэффициенты, кроме того, положительны, форма $\mathrm{A}(x, y)$ называется *положительно определенной*. Положительно определенная билинейная форма $\mathrm{A}(x, y)$ характеризуется тем, что соответствующая квадратичная форма $\mathrm{A}(x, x)$, согласно *7.92*, принимает при каждом $x \neq 0$ положительное значение.

По самому определению положительно определенная форма в пространстве R_n невырождена. Но так как в силу свойства $\mathrm{A}(x, x) > 0$ она остается положительно определенной и на любом подпространстве $\mathrm{R}' \subset \mathrm{R}_n$, то в отличие от общей билинейной формы (*7.15г*) положительно определенная форма остается невырожденной на любом подпространстве $\mathrm{R}' \subset \mathrm{R}_n$. Если $f_1, \ldots, f_k$ — любые k линейно независимых векторов, определитель

$$D = \begin{vmatrix} A(f_1, f_1) & \ldots & A(f_1, f_k) \\ \cdot & \cdot & \cdot \\ A(f_k, f_1) & \ldots & A(f_k, f_k) \end{vmatrix}$$

заведомо отличен от нуля. Ниже мы увидим, что число D всегда положительно.

7.95. Одним из важных примеров симметричной положительно определенной билинейной формы в пространстве V_3 является скалярное произведение векторов x и y. Действительно, из определения скалярного произведения непосредственно вытекают соотношения

$$(x, y) = (y, x),$$
$$(x, x) = |x|^2 > 0 \quad \text{при} \quad x \neq 0;$$

первое из них показывает, что билинейная форма (x, y) симметрична; второе — что соответствующая квадратичная форма принимает при каждом $x \neq 0$ положительное значение и, следовательно, билинейная форма (x, y) положительно определена.

В дальнейшем симметричные положительно определенные билинейные формы будут играть исключительную роль: именно, используя такие формы, мы в общем линейном пространстве получим возможность ввести понятия длин векторов и углов между векторами (гл. 8).

7.96. Возникает вопрос, как по матрице билинейной симметричной формы $\mathrm{A}(x, y)$ судить о том, является ли она положительно определенной. Ответ на этот вопрос дается следующей теоремой.

Теорема. *Для того чтобы симметричная матрица $A = \|a_{ik}\|$ определяла положительно определенную билинейную форму* $\mathrm{A}(x, y)$, *необходимо и достаточно, чтобы все угловые миноры матрицы* $\|a_{ik}\|$

$$a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \ldots, \quad \det\|a_{ik}\|$$

были положительными.

Доказательство. Если все угловые миноры матрицы A положительны, то в силу формул (24) будут положительны и все канонические коэффициенты λ_k формы $\mathrm{A}(x, y)$ в некотором каноническом базисе; таким образом, форма $\mathrm{A}(x, y)$ положительно определена.

Пусть, обратно, форма $\mathrm{A}(x, y)$ положительно определена; покажем, что тогда угловые миноры матрицы $\|a_{ik}\|$ положительны. Действительно, угловой минор

$$M = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{vmatrix} \tag{42}$$

определяет матрицу $\|a_{ik}\|$ $(i, k = 1, 2, \ldots, m)$ формы $\mathrm{A}(x, y)$ на подпространстве L_m, порожденном первыми m базисными векторами. Так как на подпространстве L_m форма $\mathrm{A}(x, y)$ положительно определена ($\mathrm{A}(x, x) > 0$ при $x \neq 0$), то в подпространстве L_m существует канонический базис, в котором форма $\mathrm{A}(x, y)$ записывается в каноническом виде с положительными каноническими коэффициентами. В частности, и определитель формы $\mathrm{A}(x, y)$ в этом базисе, равный произведению канонических коэффициентов, также положителен. Учитывая связь определителей матрицы билинейной формы в разных базисах (*7.15* (6)), мы видим, что определитель формы $\mathrm{A}(x, y)$ в исходном базисе подпространства

L_m тоже положителен. Но определитель формы $A(x, y)$ в исходном базисе подпространства L_m есть как раз минор M. Следовательно, $M > 0$, и теорема полностью доказана.

Замечание. Вместо углового минора во второй части доказательства можно было взять любой диагональный минор (*5.53*) без существенного изменения рассуждений. Таким образом, *у матрицы положительно определенной формы любой диагональный минор положителен.*

7.97. Для положительно определенной формы $A(x, y)$ всегда существует канонический базис $e_1, \ldots, e_n$, в котором все канонические коэффициенты равны $+1$ (*7.93*). Поэтому *два n-мерных вещественных пространства R'_n и R''_n с выделенными в них положительно определенными формами $A(x', y')$ и $A(x'', y'')$* в силу теоремы *7.72 всегда A-изоморфны.*

7.98. В приложениях линейной алгебры к анализу (именно, в теории условных экстремумов) часто требуется решить следующую задачу: зная матрицу $A = \| a_{ik} \|$ билинейной симметричной формы $A(x, y)$, узнать, будет ли эта форма положительно определенной на подпространстве, заданном системой k независимых линейных уравнений

$$\sum_{j=1}^{n} b_{ij}\xi_j = 0 \qquad (i = 1, 2, \ldots, k;\ k < n).$$

Оказывается, что необходимым и достаточным условием для этого является положительность угловых миноров матрицы

$$\Delta = (-1)^k \begin{Vmatrix} 0 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ 0 & 0 & \ldots & 0 & b_{21} & b_{22} & \ldots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & b_{k1} & b_{k2} & \ldots & b_{kn} \\ b_{11} & b_{21} & \ldots & b_{k1} & a_{11} & a_{12} & \ldots & a_{1n} \\ b_{12} & b_{22} & \ldots & b_{k2} & a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{1n} & b_{2n} & \ldots & b_{kn} & a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix}$$

порядка $2k+1, 2k+2, \ldots, k+n$ *).

---

*) См. заметку Р. Я. Шостака в журнале «Успехи математических наук», 1954, т. 9, вып. 2 (60), стр. 199—206.

ЗАДАЧИ

1. Образуют ли элементы матрицы билинейной формы тензор (*5.61*), и если да, то какого типа?

2. Преобразовать квадратичную форму

$$\xi_1\xi_2+\xi_2\xi_3+\xi_3\xi_1$$

к каноническому виду.

3. Пусть p — положительный индекс инерции квадратичной формы $\mathrm{A}(x, x)$ (заданной в пространстве R_n) и q — ее отрицательный индекс инерции. Пусть заданы p положительных чисел $\lambda_1, \lambda_2, \ldots, \lambda_p$ и q отрицательных чисел $\mu_1, \mu_2, \ldots, \mu_q$. Показать, что существует базис, в котором форма $\mathrm{A}(x, x)$ принимает вид

$$\mathrm{A}(x, x)=\lambda_1\tau_1^2+\lambda_2\tau_2^2+\ldots+\lambda_p\tau_p^2+\mu_1\tau_{p+1}^2+\ldots+\mu_q\tau_{p+q}^2.$$

4. Показать, что у квадратичной формы ранга r всегда имеется по крайней мере один диагональный минор ранга r, отличный от нуля.

5. Преобразовать билинейную форму

$$\mathrm{A}(x, y)=\xi_1\eta_1+\xi_1\eta_2+\xi_2\eta_1+2\xi_2\eta_2+2\xi_2\eta_3+2\xi_3\eta_2+5\xi_3\eta_3$$

к каноническому виду.

6. Применить метод Якоби для преобразования билинейной формы

$$\mathrm{A}(x, y)=\xi_1\eta_1-\xi_1\eta_2-\xi_2\eta_1+\xi_1\eta_3+\xi_3\eta_1+2\xi_2\eta_3+2\xi_3\eta_2+\xi_3\eta_3+\xi_2\eta_2$$

к каноническому виду.

7. Сформулировать условия, при которых симметричная матрица $\|a_{ik}\|$ определяет отрицательно определенную билинейную форму.

8. Дана симметричная матрица $A=\|a_{ik}\|$, обладающая свойствами

$$a_{11}>0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0, \ \ldots, \ \det\|a_{ik}\|>0.$$

Доказать, что $a_{nn}>0$.

9. Доказать теорему: полилинейная антисимметричная форма от $n+1$ векторов в n-мерном пространстве K_n тождественно равна нулю.

10. Доказать теорему: полилинейная антисимметричная форма $\mathrm{A}(x_1, \ldots, x_{n-1})$ от $n-1$ векторов в n-мерном пространстве записывается в любом базисе как определитель, первые $n-1$ строк которого заполнены координатами векторов-аргументов, а последняя, n-я, строка заполнена n фиксированными числами.

11. Доказать, что всякая антисимметричная билинейная форма $\mathrm{A}(x, y)\not\equiv 0$ может быть всегда приведена к «каноническому виду»

$$\mathrm{A}(x, y)=\sigma_1\tau_2-\sigma_2\tau_1+\sigma_3\tau_4-\sigma_4\tau_3+\ldots+\sigma_{2k-1}\tau_{2k}+\sigma_{2k}\tau_{2k-1}.$$

12. Доказать теорему: вещественная квадратичная форма $\mathrm{A}(x, x)=\sum a_{jk}x_jx_k$ неотрицательна при всех $x\in R_n$ тогда и только тогда, когда *все* диагональные миноры матрицы $A=\|a_{jk}\|$ неотрицательны.

Примечание. Для матрицы $\begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix}$ угловые миноры δ_1 и δ_2 равны 0, а соответствующая форма не является неотрицательной. Таким образом, условия $\delta_1 \geqslant 0$, $\delta_2 \geqslant 0$ *недостаточны* для неотрицательности формы.

13. Пусть в пространстве K_n задана невырожденная билинейная симметричная форма $A(x, y)$. Пусть $K' \subset K$ — подпространство размерности r. Доказать, что сопряженное подпространство $K'' \subset K$ имеет размерность $n-r$.

14. В пространстве R_2 задана симметричная билинейная форма $(x, y) = \xi_1\eta_1 - \xi_2\eta_2$. Найти оператор, сопряженный (относительно этой формы) к оператору поворота с матрицей

$$A = \begin{Vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{Vmatrix}.$$

15. Пусть в пространстве K_n выделена невырожденная форма (x, y). Для квадратной системы линейных уравнений

$$\sum_{k=1}^{n} a_{jk}\xi_k = b_j \qquad (j = 1, \ldots, n) \tag{43}$$

доказать «теорему Фредгольма»: система (43) имеет решение для тех и только для тех векторов $b = \{b_1, \ldots, b_n\}$, которые сопряжены ко всем решениям однородной системы

$$\sum_{k=1}^{n} a'_{jk}\eta_k = 0, \tag{44}$$

где $\| a'_{jk} \|$ — матрица, сопряженная с $\| a_{jk} \|$ относительно формы (x, y).

Отсюда вывести, что число независимых линейных условий на вектор b, необходимое и достаточное для разрешимости системы (43), равно размерности пространства решений однородной системы

$$\sum_{k=1}^{n} a_{jk}\xi_k = 0 \qquad (j = 1, \ldots, n). \tag{45}$$

Примечание. Для общей системы

$$\sum_{k=1}^{n} a_{jk}\xi_k = b_j \qquad (j = 1, \ldots, m) \tag{46}$$

две указанные характеристики уже не совпадают: их разность, называемая *индексом* системы (46), равна $m-n$.

16. Показать, что всякая неотрицательная билинейная форма ранга r в пространстве R_n может быть представлена как сумма r неотрицательных билинейных форм ранга 1.

17. Показать, что всякая билинейная форма ранга 1 в пространстве K_n имеет вид $A(x, y) = f(x) \cdot g(y)$, где $f(x)$ и $g(x)$ — линейные формы.

18. Если $A(x, y) = \sum a_{jk}\xi_j\eta_k$ и $B(x, y) = \sum b_{jk}\xi_j\eta_k$ — неотрицательные билинейные формы в пространстве R_n, то форма $C(x, y) = \sum a_{jk}b_{jk}\xi_j\eta_k$ также неотрицательна.

ЕВКЛИДОВЫ ПРОСТРАНСТВА

§ 8.1. Введение

Большое многообразие фактов, которыми богата геометрия, в значительной мере объясняется возможностью различных измерений — в основном возможностью измерения длин отрезков и углов между прямыми. В общем линейном пространстве мы еще не имеем способа производить такие измерения; и это, разумеется, суживает область нашего исследования.

Желая наиболее естественным образом распространить на общие линейные пространства методы, связанные с возможностью измерений, мы обратимся к определению скалярного произведения двух векторов, принятому в аналитической геометрии (и пригодному, конечно, только для обычных векторов — элементов пространства V_3). Это определение гласит: *скалярное произведение двух векторов есть произведение длин этих векторов и косинуса угла между ними.* Следовательно, это определение уже основано на возможности измерения длин векторов и угла между ними. Но, с другой стороны, зная скалярное произведение любой пары векторов, мы можем восстановить длины их и угол между ними; действительно, квадрат длины вектора равен скалярному произведению этого вектора с самим собой, а косинус угла между двумя векторами — отношению их скалярного произведения к произведению их длин. Следовательно, в понятии скалярного произведения потенциально заключены и возможность измерения длин, и возможность измерения углов, а вместе с ними и вся область геометрии, связанная с измерениями («метрическая геометрия»).

В общем линейном пространстве нам легче будет сначала ввести понятие скалярного произведения двух векторов и затем из имеющегося уже понятия скалярного произведения получить определения длины вектора и угла между векторами.

Посмотрим, какие свойства обычного скалярного произведения можно использовать для построения аналогичной величины в общем линейном пространстве. Вначале ограничимся случаем вещественных пространств.

В *7.95* мы уже видели, что в пространстве V_3 (*2.15 а*) скалярное произведение (x, y) есть билинейная форма от векторов x и y, симметричная и положительно определенная. Формы с такими свойствами, вообще говоря, имеются и в общем вещественном линейном пространстве. Рассмотрим в вещественном линейном пространстве произвольную фиксированную билинейную форму $\mathrm{A}(x, y)$, симметричную и положительно определенную. Назовем ее «скалярным произведением» векторов x и y. После этого определим длину каждого вектора и угол между каждыми двумя векторами по тем же правилам, по которым длина вектора и угол между векторами вычислялись через скалярные произведения в пространстве V_3. Разумеется, только дальнейшее исследование может показать, насколько удачно это определение; и мы увидим на протяжении этой и следующих глав, что приведенное определение на самом деле позволит распространить на общие линейные пространства методы метрической геометрии и тем самым значительно усилить средства исследования математических объектов, встречающихся в алгебре и анализе.

Отметим здесь одно существенное обстоятельство. Исходную билинейную и положительно определенную форму можно выбрать в данном линейном пространстве многими различными способами. Длина некоторого вектора x, вычисленная с помощью одной такой формы, будет отличаться от длины этого же вектора x, вычисленной с помощью другой формы; то же относится и к углу между двумя данными векторами. Таким образом, длина вектора и угол между векторами не определяются однозначно. Но эта неоднозначность не должна нас смущать; ведь ничего нет удивительного в том, что одному и тому же отрезку на прямой, измеренному различными масштабами, приписываются в результате этих измерений в качестве его длины различные числа. Можно сказать, что выбор исходной билинейной симметричной и положительной формы аналогичен выбору такого «масштаба» для измерения длин векторов и углов между ними.

Вещественное линейное пространство с выбранной в нем «масштабной» билинейной симметричной положительно определенной формой будем в дальнейшем называть *евклидовым*. Линейное пространство без заданной «масштабной» формы будем называть *аффинным*. Случай комплексного пространства мы рассмотрим в гл. 9.

§ 8.2. Определение евклидова пространства

8.21. Вещественное линейное пространство R называется *евклидовым*, если 1) имеется правило, которое позволяет построить для каждых двух векторов x и y из R вещественное число, называемое *скалярным произведением* векторов x и y и обозначаемое (x, y); 2) это правило удовлетворяет следующим требованиям:

а) $(x, y)=(y, x)$ (переместительный закон),

б) $(x, y+z)=(x, y)+(x, z)$ (распределительный закон),

в) $(\lambda x, y)=\lambda(x, y)$ для любого вещественного числа λ,

г) $(x, x)>0$ при $x \neq 0$ и $(x, x)=0$ при $x=0$.

Аксиомы а)—г) утверждают в совокупности, что скалярное произведение векторов x и y есть *билинейная форма* (б—в), *симметричная* (а) *и положительно определенная* (г). И обратно, *всякая форма, обладающая этими свойствами, может быть принята за скалярное произведение.*

Поскольку скалярное произведение векторов x и y является билинейной формой, для него имеет место формула *7.11* (2). В данном случае эта формула принимает следующий вид:

$$\left(\sum_{i=1}^{k}\alpha_i x_i,\ \sum_{j=1}^{m}\beta_j y_j\right)=\sum_{i=1}^{k}\sum_{j=1}^{m}\alpha_i\beta_j(x_i, y_j). \tag{1}$$

Здесь $x_1, x_2, \ldots, x_k, y_1, y_2, \ldots, y_m$ — произвольные векторы евклидова пространства R, $\alpha_1, \alpha_2, \ldots, \alpha_k$, $\beta_1, \beta_2, \ldots, \beta_m$ — любые вещественные числа.

8.22. Примеры.

а. В пространстве V_3 свободных векторов (*2.15а*) скалярное произведение вводится по правилам аналитической геометрии. Условия а)—г) выражают собой основные свойства скалярного произведения; они доказываются в векторной алгебре.

б. В пространстве R_n *(2.15б)* мы введем скалярное произведение векторов $x = (\xi_1, \xi_2, \ldots, \xi_n)$ и $y = (\eta_1, \eta_2, \ldots, \eta_n)$ так:

$$(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n. \tag{2}$$

(Это определение обобщает известную формулу выражения скалярного произведения векторов в трехмерном пространстве через координаты сомножителей в ортогональной системе координат.) Легко проверить, что требования а) — г) удовлетворяются.

Заметим, что формула (2) — не единственный способ введения скалярного произведения в R_n. Все возможные способы введения скалярного произведения (т. е. симметричной билинейной положительно определенной формы) в пространстве R_n мы описали фактически в *7.96*.

в. В пространстве $R(a, b)$ вещественных непрерывных функций на отрезке $a \leqslant t \leqslant b$ *(2.15в)* мы можем ввести скалярное произведение функций $x(t)$ и $y(t)$ по формуле

$$(x, y) = \int_a^b x(t)\, y(t)\, dt. \tag{3}$$

Легко проверить, применяя основные правила интегрирования, что требования а) — г) удовлетворяются.

В дальнейшем будем обозначать пространство $R(a, b)$ со скалярным произведением по формуле (2) через $R_2(a, b)$.

§ 8.3. Основные метрические понятия

Имея скалярное произведение, мы можем дать определение и основных метрических понятий — длины вектора и угла между двумя векторами.

8.31. Длина вектора. *Длиной* вектора x в евклидовом пространстве R мы будем называть величину

$$|x| = +\sqrt{(x, x)}. \tag{4}$$

Примеры.

а. В пространстве V_3 наше определение длины вектора приводит к обычному значению длины вектора.

б. В пространстве R_n для вектора $x=(\xi_1, \xi_2, \ldots, \xi_n)$ получается выражение длины в виде

$$|x| = +\sqrt{\xi_1^2+\xi_2^2+\ldots+\xi_n^2}.$$

в. В пространстве $R_2(a, b)$ длина вектора $x=x(t)$ оказывается равной

$$|x| = +\sqrt{(x, x)} = \sqrt{\int_a^b x^2(t)\,dt}.$$

Эту величину обозначают иногда $\|x(t)\|$ и называют *нормой* функции $x(t)$ (чтобы избежать ложных ассоциаций, связанных со словами «длина функции»).

8.32. Из аксиомы г) вытекает, что у каждого вектора x евклидова пространства R существует длина; у всякого вектора $x \neq 0$ длина положительна, у нуль-вектора длина равна нулю. Равенство

$$|\alpha x| = \sqrt{(\alpha x, \alpha x)} = \sqrt{\alpha^2(x, x)} = |\alpha|\sqrt{(x, x)} = |\alpha|\,|x| \quad (5)$$

показывает, что *абсолютную величину числового множителя можно выносить за знак длины вектора.*

Вектор x, имеющий длину 1, называется *нормированным.* Всякий ненулевой вектор y можно *нормировать*, т. е. умножить на такое число λ, чтобы в результате получился нормированный вектор. Действительно, уравнение $|\lambda y| = 1$ относительно λ имеет решение, например,

$$\lambda = \frac{1}{|y|}.$$

Множество $F \subset R$ называется *ограниченным*, если длины всех векторов $x \in F$ ограничены фиксированной константой. Примерами ограниченных множеств являются *единичный шар* пространства R — совокупность всех векторов $x \in R$ с длиной, не превышающей единицы, а также *единичная сфера* — совокупность всех векторов $x \in R$ с длиной, равной 1.

8.33. Угол между векторами. *Углом между парой векторов* x и y мы будем называть тот угол (в пределах от 0 до 180°), косинус которого равен отношению $\frac{(x, y)}{|x||y|}$.

Для обычных векторов (в пространстве V_3) наше определение согласуется с обычным выражением угла между векторами через скалярное произведение.

Чтобы это определение можно было применить в общем евклидовом пространстве, необходимо доказать, что указанное отношение по абсолютной величине не превосходит единицы, каковы бы ни были векторы x и y.

Для доказательства этого утверждения рассмотрим вектор $\lambda x - y$, где λ — вещественное число. В силу аксиомы г) при любом λ

$$(\lambda x - y, \ \lambda x - y) \geqslant 0. \tag{6}$$

Используя формулу (1), мы можем написать это неравенство в виде

$$\lambda^2 (x, x) - 2\lambda (x, y) + (y, y) \geqslant 0. \tag{7}$$

В левой части неравенства стоит квадратный трехчлен относительно λ с постоянными коэффициентами. Трехчлен этот не может иметь различных вещественных корней, так как тогда он не мог бы сохранять знака для всех значений λ. Поэтому дискриминант $(x, y)^2 - (x, x)(y, y)$ этого трехчлена не может быть положительным. Следовательно,

$$(x, y)^2 \leqslant (x, x)(y, y),$$

откуда, извлекая квадратный корень, получаем

$$|(x, y)| \leqslant |x||y|, \tag{8}$$

что и требовалось.

Неравенство (8) называют *неравенством Коши — Буняковского*.

8.34. Выясним, когда в неравенстве (8) возможен знак равенства.

Если векторы x и y коллинеарны, т. е., например, $y = \lambda x$, $\lambda \in R$, то, очевидно,

$$|(x, y)| = |(x, \lambda x)| = |\lambda|(x, x) = |\lambda||x|^2 = |x||y|.$$

Покажем, что и обратно, *если неравенство* (8) *для некоторой пары векторов x и y обратилось в равенство, то эти векторы x и y коллинеарны.*

Если имеет место равенство

$$|(x, y)| = |x||y|,$$

то дискриминант квадратного трехчлена (7) равен нулю и, следовательно, трехчлен имеет один вещественный корень λ_0. Мы получаем, таким образом,

$$\lambda_0^2(x, x) - 2\lambda_0(x, y) + (y, y) = (\lambda_0 x - y, \lambda_0 x - y) = 0,$$

откуда в силу аксиомы г) находим, что $\lambda_0 x - y = 0$, или $y = \lambda_0 x$. Итак, *абсолютная величина скалярного произведения двух векторов тогда и только тогда равна произведению их длин, когда эти векторы коллинеарны.*

Примеры.

а. В пространстве V_3 неравенство Коши — Буняковского, очевидно, вытекает из самого определения скалярного произведения как произведения длин векторов и косинуса угла между ними.

б. В пространстве R_n неравенство Коши — Буняковского имеет вид

$$\left|\sum_{j=1}^{n}\xi_j\eta_j\right| \leqslant \sqrt{\sum_{j=1}^{n}\xi_j^2}\sqrt{\sum_{j=1}^{n}\eta_j^2}; \qquad (9)$$

оно справедливо для любой пары векторов $x = (\xi_1, \xi_2, \ldots, \xi_n)$ и $y = (\eta_1, \eta_2, \ldots, \eta_n)$ или, что то же самое, для любых двух систем вещественных чисел $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$.

в. В пространстве $R_2(a, b)$ неравенство Коши — Буняковского имеет вид

$$\left|\int_a^b x(t)\,y(t)\,dt\right| \leqslant \sqrt{\int_a^b x^2(t)\,dt}\sqrt{\int_a^b y^2(t)\,dt}. \qquad (10)$$

8.35. Ортогональность. Векторы x и y называются *ортогональными*, если $(x, y) = 0$. Таким образом, понятие ортогональности векторов x и y совпадает с понятием сопряженности (*7.41а*) этих векторов относительно билинейной формы (x, y). Если $x \neq 0$ и $y \neq 0$, то это определение в соответствии с общим определением угла между двумя векторами (*8.33*) означает, что x и y образуют угол в $90°$. Нулевой вектор оказывается ортогональным к любому вектору $x \in \mathrm{R}$.

Примеры.

а. В пространстве R_n (*8.22 б*) условие ортогональности векторов $x = (\xi_1, \xi_2, \ldots, \xi_n)$ и $y = (\eta_1, \eta_2, \ldots, \eta_n)$ имеет вид

$$\xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n = 0.$$

Например, векторы $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0), \ldots$ $\ldots, e_n = (0, 0, \ldots, 1)$ попарно взаимно ортогональны.

б. В пространстве $R_2(a, b)$ условие ортогональности векторов $x = x(t)$ и $y = y(t)$ имеет вид

$$\int_a^b x(t)\, y(t)\, dt = 0.$$

Читатель легко проверит, вычислив соответствующие интегралы, что в пространстве $R_2(-\pi, \pi)$ любые два вектора «тригонометрической системы»

$$1,\ \cos t,\ \sin t,\ \cos 2t,\ \sin 2t,\ \ldots,\ \cos nt,\ \sin nt, \ldots$$

взаимно ортогональны.

8.36. Приведем несколько простых утверждений, связанных с понятием ортогональности.

а. Лемма. *Взаимно ортогональные ненулевые векторы $x_1, x_2, \ldots, x_k$ линейно независимы.*

Доказательство. Допустим, что эти векторы линейно зависимы; тогда имеет место равенство

$$C_1x_1 + C_2x_2 + \ldots + C_kx_k = 0,$$

где, например, $C_1 \neq 0$. Умножим это равенство скалярно на x_1; в силу предположения о взаимной ортогональности векторов $x_1, x_2, \ldots, x_k$ мы получим $C_1(x_1, x_1) = 0$. Отсюда $(x_1\ x_1) = 0$ и, следовательно, x_1 есть нулевой вектор в противоречие с предположением.

Результат этой леммы мы будем тогда использовать в такой форме: *если сумма взаимно ортогональных векторов равна нулю, то каждое из слагаемых равно нулю.*

6. Лемма. *Если векторы $y_1, y_2, \ldots, y_k$ ортогональны к вектору x, то любая линейная комбинация $\alpha_1 y_1 + \alpha_2 y_2 + \ldots$ $\ldots + \alpha_k y_k$ также ортогональна к вектору x.*

Действительно,

$$(\alpha_1 y_1 + \ldots + \alpha_k y_k, x) = \alpha_1 (y_1, x) + \ldots + \alpha_k (y_k, x) = 0;$$

следовательно, вектор $\alpha_1 y_1 + \ldots + \alpha_k y_k$ ортогонален к вектору x, что и утверждалось (ср. *7.41в*).

Совокупность всех линейных комбинаций $\alpha_1 y_1 + \alpha_2 y_2 + \ldots$ $\ldots + \alpha_k y_k$ образует подпространство $\mathrm{L} = \mathrm{L}(y_1, y_2, \ldots, y_k)$ — линейную оболочку векторов $y_1, y_2, \ldots, y_k$ (*2.51*). Следовательно, вектор x ортогонален к каждому вектору подпространства L. В таких случаях мы будем говорить, что *вектор x ортогонален к подпространству* L. Вообще, если $\mathrm{F} \subset \mathrm{R}$ — произвольное множество векторов в евклидовом пространстве R, то мы будем говорить, что вектор x ортогонален к множеству F, если он ортогонален к любому вектору из F.

Совокупность G всех векторов x, ортогональных к множеству F, в силу той же леммы *б* сама составляет подпространство в пространстве R. Чаще всего такая ситуация встречается в том случае, когда F есть подпространство; тогда подпространство G называется *ортогональным дополнением подпространства* F.

8.37. Теорема Пифагора и ее обобщение. Пусть векторы x и y ортогональны; тогда по аналогии с элементарной геометрией вектор $x+y$ можно называть гипотенузой прямоугольного треугольника, построенного на векторах x и y. Умножая $x+y$ скалярно на себя и используя ортогональность векторов x и y, мы получаем

$$|x+y|^2 = (x+y, x+y) = (x, x) + 2(x, y) + (y, y) = \\ = (x, x) + (y, y) = |x|^2 + |y|^2.$$

Мы доказали тем самым в общем евклидовом пространстве теорему Пифагора: *квадрат гипотенузы равен сумме квадратов катетов*. Нетрудно обобщить эту теорему на случай любого числа слагаемых. Именно, пусть векторы $x_1, x_2, \ldots$ $\ldots, x_k$ взаимно ортогональны и $z = x_1 + x_2 + \ldots + x_k$;

тогда

$$|z|^2=(x_1+x_2+\ldots+x_k,\ x_1+x_2+\ldots+x_k)= \\ =|x_1|^2+|x_2|^2+\ldots+|x_k|^2. \quad (11)$$

8.38. Неравенства треугольника. Если x и y — произвольные векторы, то по аналогии с элементарной геометрией вектор $x+y$ естественно называть третьей стороной треугольника, построенного на векторах x и y. Используя неравенство Коши — Буняковского, мы получаем

$$\begin{aligned} |x+y|^2=(x+y,\ x+y)=(x,\ x)+2(x,\ y)+(y,\ y), \\ \leqslant |x|^2+2|x||y|+|y|^2=(|x|+|y|)^2, \\ \geqslant |x|^2-2|x||y|+|y|^2=(|x|-|y|)^2, \end{aligned}$$

или

$$|x+y|\leqslant|x|+|y|, \quad (12)$$

$$|x+y|\geqslant\big||x|-|y|\big|. \quad (13)$$

Неравенства (12) — (13) называются неравенствами треугольника. Геометрически они означают, что *длина любой стороны всякого треугольника не больше, чем сумма длин двух других сторон, и не меньше, чем абсолютная величина разности длин этих сторон.*

8.39. Мы могли бы, далее, последовательно переносить на евклидово пространство остальные теоремы элементарной геометрии. Но в этом нет нужды. Введем понятие евклидова изоморфизма между евклидовыми пространствами R' и R'': именно, мы будем говорить, что R' и R'' *евклидово-изоморфны*, если они изоморфны как вещественные линейные пространства (*2.71*) и, кроме того, для любой пары соответствующих векторов x', y' в R' и x'', y'' в R'' выполняется равенство

$$(x',\ y')=(x'',\ y'').$$

Очевидно, что всякая геометрическая теорема, — так мы будем называть теорему, основанную на понятиях линейного пространства и скалярного произведения, — доказанная для пространства R', будет справедлива и для изоморфного ему пространства R''. Теперь заметим, что в силу теоремы *7.97* всякие два евклидова пространства равной размерности

n евклидово-изоморфны. Следовательно, всякая геометрическая теорема, справедливая в n-мерном евклидовом пространстве R'_n будет справедливой и в любом другом n-мерном евклидовом пространстве R''_n. Геометрические теоремы элементарной геометрии, т. е. геометрические теоремы в пространстве R_3, по доказанному, остаются справедливыми в любом трехмерном подпространстве любого евклидова пространства. Таким образом, все геометрические теоремы элементарной геометрии справедливы в любом евклидовом пространстве.

§ 8.4. Ортогональный базис

8.41. Теорема. *В n-мерном евклидовом пространстве R_n существует базис из n ненулевых взаимно ортогональных векторов.*

Доказательство. Для билинейной формы (x, y), как и для всякой симметричной билинейной формы в n-мерном пространстве, существует канонический базис (*7.43*) $y_1, y_2, \ldots, y_n$. Условие $(y_i, y_k) = 0$ каноничности базиса при $i \neq k$ есть в данном случае условие ортогональности векторов y_i и y_k; таким образом, канонический базис $y_1, y_2, \ldots, y_n$ образуется в данном случае из n взаимно ортогональных векторов. Это и доказывает теорему.

В § 8.6 мы рассмотрим способы эффективного построения ортогонального базиса.

8.42. Векторы $y_1, y_2, \ldots, y_n$ ортогонального базиса удобно нормировать, разделив каждый из них на его длину. Мы получаем после этого в пространстве R *ортогональный* и *нормированный базис* (иногда говорят «ортонормированный» или «ортонормальный» базис).

Пусть $e_1, e_2, \ldots, e_n$ — произвольный ортогональный нормированный базис в евклидовом пространстве R_n. Каждый вектор $x \in R_n$ можно представить в виде

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n, \tag{14}$$

где $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x. Мы будем также называть эти координаты *коэффициентами Фурье* вектора x относительно ортогональной и нормированной системы $e_1, e_2, \ldots, e_n$. Умножая равенство (14) скалярно на

e_i, находим выражение коэффициента ξ_i:

$$\xi_i = (x, e_i) \qquad (i = 1, 2, \ldots, n). \tag{15}$$

Если $y = \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n$ — любой другой вектор пространства R_n, то по формуле (1) мы получаем

$$(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n. \tag{16}$$

Итак, *в нормированном ортогональном базисе скалярное произведение двух векторов равно сумме произведений их соответствующих координат — коэффициентов Фурье.*

В частности, полагая $y = x$, получим

$$|x|^2 = (x, x) = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2. \tag{17}$$

§ 8.5. Задача о перпендикуляре

8.51. Рассмотрим в евклидовом пространстве R некоторое конечномерное подпространство R′ и вектор f, вообще говоря, не входящий в подпространство R′. Поставим задачу найти разложение

$$f = g + h, \tag{18}$$

где вектор g принадлежит подпространству R′, а вектор h ортогонален к этому подпространству. Вектор g, участвующий в разложении (18), называется *проекцией вектора f на* подпространство R′, а вектор h — *перпендикуляром, опущенным из конца вектора f на подпространство* R′. Такие названия связаны с привычными нам геометрическими ассоциациями*).

Решение этой задачи фактически было дано еще в *7.54*, для любой симметричной билинейной формы, невырожденной на подпространстве R′. Так как положительно определенная форма (x, y) невырождена на любом подпространстве $R' \subset R$ (*7.94*), то существование решения нашей задачи, вместе с единственностью, следует из *7.54*.

Как мы видели в *7.55*, наличие разложения (18) показывает, что все пространство R есть прямая сумма подпространства R′ и его ортогонального дополнения R″.

*) И предназначены только для того, чтобы вызывать такие ассоциации. Поскольку понятие «конец вектора» не фигурирует в нашей аксиоматике, не следует искать в этом названии логического смысла.

Прямая сумма, слагаемые которой ортогональны, называется *ортогональной прямой суммой*; мы построили, таким образом, разложение R в ортогональную прямую сумму R′ и R″. Если размерность пространства R равна n, а размерность R′ равна k, то размерность R″ равна $n-k$, поскольку размерность прямой суммы есть сумма размерностей слагаемых (*2.47a*).

Заметим, что задача решается и в том случае, когда вектор f лежит в подпространстве R′. В этом случае решение имеет вид

$$f=f+0.$$

Другого решения, очевидно, нет: если бы мы имели $f=g+h$, $g\in\mathrm{R}'$, $h\in\mathrm{R}''$, то было бы также $h=f-g\in\mathrm{R}'$, откуда $h=0$, $g=f$.

8.52. Применяя к разложению (18) теорему Пифагора (*8.37*), получаем

$$|f|^2=|g|^2+|h|^2, \tag{19}$$

откуда вытекает, что справедливо неравенство

$$0\leqslant|h|\leqslant|f|, \tag{20}$$

геометрически выражающее тот факт, что *длина перпендикуляра не превосходит длины наклонной.*

Отметим те случаи, когда в одном из неравенств (20) имеет место знак равенства. Условие $0=|h|$ равносильно условию $f=g+0=g$, которое означает, что f входит в подпространство R′. Условие $|h|=|f|$, согласно теореме Пифагора, показывает, что $g=0$ и, следовательно,

$$f=0+h=h;$$

таким образом, f ортогонален к подпространству R′. Итак, *равенство* $|h|=0$ *означает, что вектор* f *входит в подпространство* R′; *равенство* $|h|=|f|$ *означает, что вектор* f *ортогонален к этому подпространству*. При всяком ином расположении вектора f длина вектора h будет положительной величиной, меньшей чем длина вектора f.

Пусть теперь $e_1, \ldots, e_k$ — ортонормальный базис в подпространстве R′ и пусть $g = \sum_{j=1}^{k} a_j g_j$. Тогда по *8.42* (17)

$$|g|^2 = \sum_{j=1}^{k} a_j^2.$$

Подставляя это значение $|g|^2$ в равенство (19), получаем

$$|f|^2 = |h|^2 + \sum_{j=1}^{k} a_j^2.$$

В частности, для любой (конечной) ортогональной нормированной системы $e_1, e_2, \ldots, e_k$ и любого вектора f мы получаем неравенство

$$\sum_{j=1}^{k} a_j^2 \leqslant |f|^2,$$

которое называется *неравенством Бесселя*. Геометрический смысл его очевиден: *квадрат длины вектора f не меньше, чем сумма квадратов его проекций на любые k взаимно ортогональных направлений.*

8.53. В приложениях требуется иногда дать эффективное решение задачи о перпендикуляре, когда в подпространстве R′ дан некоторый (вообще говоря, не ортогональный и не нормированный) базис $\{b_1, b_2, \ldots, b_k\} = \{b\}$.

Чтобы получить это решение, разложим искомый вектор g по базису $\{b\}$,

$$g = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_k b_k,$$

и подчиним вектор $h = f - g$ условию ортогональности со всеми векторами $b_1, b_2, \ldots, b_k$. Мы получим следующую систему уравнений:

$$\begin{aligned}
&(h, b_1) = (f - g, b_1) = \\
&\qquad = (f, b_1) - \beta_1 (b_1, b_1) - \beta_2 (b_2, b_1) - \ldots - \beta_k (b_k, b_1) = 0, \\
&(h, b_2) = (f - g, b_2) = \\
&\qquad = (f, b_2) - \beta_1 (b_1, b_2) - \beta_2 (b_2, b_2) - \ldots - \beta_k (b_k, b_2) = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&(h, b_k) = (f - g, b_k) = \\
&\qquad = (f, b_k) - \beta_1 (b_1, b_k) - \beta_2 (b_2, b_k) - \ldots - \beta_k (b_k, b_k) = 0
\end{aligned}$$

с определителем

$$D = \begin{vmatrix} (b_1, b_1) & (b_2, b_1) & \dots & (b_k, b_1) \\ (b_1, b_2) & (b_2, b_2) & \dots & (b_k, b_2) \\ \dots & \dots & \dots & \dots \\ (b_1, b_k) & (b_2, b_k) & \dots & (b_k, b_k) \end{vmatrix}.$$

Определитель D, как определитель матрицы положительно определенной формы (x, y) в базисе $b_1, \dots, b_k$, отличен от нуля (*7.96*). Разрешая систему по правилу Крамера, получаем выражения для коэффициентов β_j $(j = 1, 2, \dots, n)$:

$$\beta_j = \frac{1}{D} \begin{vmatrix} (b_1, b_1) & \dots & (b_{j-1}, b_1) & (f, b_1) & (b_{j+1}, b_1) & \dots & (b_k, b_1) \\ (b_1, b_2) & \dots & (b_{j-1}, b_2) & (f, b_2) & (b_{j+1}, b_2) & \dots & (b_k, b_2) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ (b_1, b_k) & \dots & (b_{j-1}, b_k) & (f, b_k) & (b_{j+1}, b_k) & \dots & (b_k, b_k) \end{vmatrix}.$$

8.54. Задачу о перпендикуляре можно поставить не только для подпространства, но и для гиперплоскости. В этом случае она формулируется так: в евклидовом пространстве R даны гиперплоскость R″, полученная параллельным сдвигом некоторого подпространства R′, и вектор f; требуется доказать, что существует и единственно разложение

$$f = g + h, \tag{21}$$

где g принадлежит гиперплоскости R″ *), а вектор h ортогонален подпространству R′. Геометрический смысл этого разложения ясен из рис. 1, *а*. В разложении (21) слагаемые, вообще говоря, уже не ортогональны.

Эту задачу легко свести к задаче *8.51*. Действительно, если в гиперплоскости R″ фиксировать любой вектор f_0 и вычесть его из обеих частей равенства (21), то мы получим задачу о разложении вектора $f - f_0$ на слагаемые $g - f_0$ и h, первое из которых принадлежит подпространству R′, а второе ортогонально к этому подпространству (см. рис. 1, *б*). В силу результата *8.51* такое разложение существует;

*) Геометрически это означает, что конец вектора g лежит в гиперплоскости R″ (а его начало, как всегда,— в начале координат). Не следует представлять себе весь вектор g лежащим в гиперплоскости R″!

следовательно, существует и разложение (21). Остается установить единственность разложения (21). В случае наличия двух разложений указанного вида

$$f = g_1 + h_1 = g_2 + h_2$$

мы имели бы

$$0 = (g_1 - g_2) + (h_1 - h_2).$$

Здесь $g_1 - g_2$ принадлежит подпространству R′, а $h_1 - h_2$

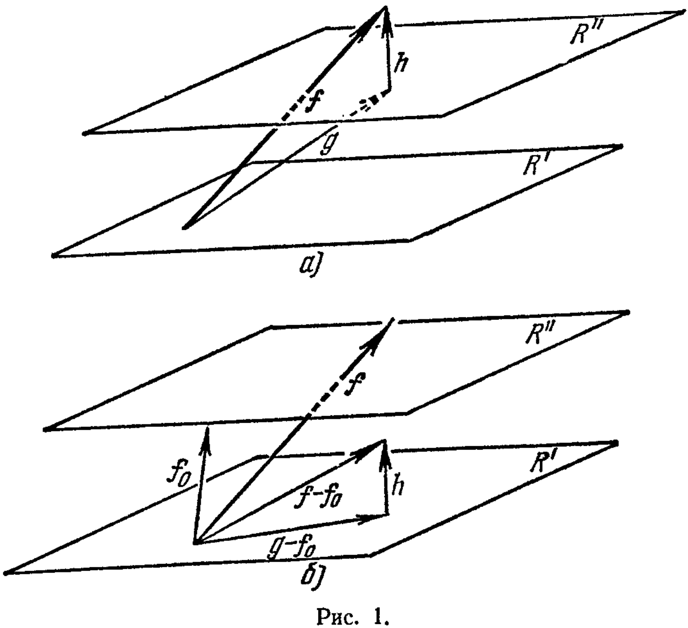

Рис. 1.

ортогонально этому подпространству. Отсюда $g_1 - g_2 = h_1 - h_2 = 0$, что и требуется.

§ 8.6. Общая теорема об ортогонализации

8.61. Для построения ортогональных систем в евклидовом пространстве основное значение имеет следующая общая теорема:

Теорема. (Теорема об ортогонализации.) *Пусть $x_1, x_2, \ldots, x_k, \ldots$ — некоторая последовательность* *векторов евклидова пространства* R (*конечная или бесконечная*). *Обозначим через* $\mathrm{L}_k = \mathrm{L}(x_1, x_2, \ldots, x_k)$ *линейную оболочку первых* k *векторов этой системы. Утверждается, что существует система векторов* $y_1, y_2, \ldots, y_k, \ldots$, *обладающая следующими свойствами:*

1) *Для любого натурального* k *линейная оболочка* L'_k *векторов* $y_1, y_2, \ldots, y_k$ *совпадает с подпространством* L_k.

2) *Для любого натурального* k *вектор* y_{k+1} *ортогонален к подпространству* L_k.

Доказательство. Положим $y_1 = x_1$. Очевидно, что при этом выполнено условие

$$\mathrm{L}'_1 = \mathrm{L}_1.$$

Далее будем доказывать теорему по индукции: предположим, что уже построено k векторов $y_1, y_2, \ldots, y_k$, удовлетворяющих поставленным условиям, и построим вектор y_{k+1} так, чтобы он также обладал требуемыми свойствами.

Пространство L_k конечномерно, и поэтому в силу *8.51* существует разложение

$$x_{k+1} = g_k + h_k, \tag{22}$$

где вектор g_k входит в подпространство L_k, а вектор h_k ортогонален к этому подпространству. Мы положим $y_{k+1} = h_k$. Проверим выполнение условий теоремы ортогонализации для определенного таким образом вектора y_{k+1}.

Подпространство L_k, по предположению индукции, содержит векторы $y_1, y_2, \ldots, y_k$; поэтому и более широкое подпространство L_{k+1} содержит эти векторы; кроме того, из формулы (22) вытекает, что L_{k+1} содержит вектор $h_k = y_{k+1}$. Таким образом, подпространство L_{k+1} содержит все векторы $y_1, \ldots, y_{k+1}$, а вместе с ними и всю их линейную оболочку L'_{k+1}. Но и обратно, подпространство L'_{k+1} содержит векторы $x_1, x_2, \ldots, x_k$ и, как показывает равенство (22), содержит и вектор x_{k+1}; отсюда вытекает, что L'_{k+1} содержит все подпространство L_{k+1}. Следовательно, $\mathrm{L}'_{k+1} = \mathrm{L}_{k+1}$, и первое условие теоремы ортогонализации выполнено. Выполнение же второго условия очевидно по самому построению вектора $y_{k+1} = h_k$.

Индукция, таким образом, проведена, и тем самым теорема полностью доказана.

с линейной оболочкой $L(1, t, t^2, \ldots, t^n)$ и, следовательно, совпадает с совокупностью L_n, что и требовалось.

б. *Вектор* $p_n(t)$ *ортогонален к подпространству* L_{n-1}. Достаточно проверить, что многочлен $p_n(t)$ ортогонален в $R_2(-1, 1)$ к функциям $1, t, \ldots, t^{n-1}$.

При доказательстве мы будем использовать формулу интегрирования по частям в определенном интеграле, предполагаемую известной из анализа. Производные, входящие в эту формулу, для многочленов суть те же самые производные, которые мы рассматривали в *6.73а* с чисто алгебраической точки зрения. В частности, многочлен

$$[(t^2-1)^n] = (t-1)^n (t+1)^n$$

в силу *6.73в* при $t = \pm 1$ имеет равные 0 производные порядков $0, 1, \ldots, n-1$.

Итак, будем вычислять скалярное произведение t^k и $p_n(t)$. Интегрируя по частям, мы получаем

$$(t^k, p_n(t)) = \int_{-1}^{+1} t^k [(t^2-1)^n]^{(n)}\, dt =$$

$$= t^k [(t^2-1)^n]^{(n-1)} \Big|_{-1}^{+1} - k \int_{-1}^{+1} t^{k-1} [(t^2-1)^n]^{(n-1)}\, dt.$$

Внеинтегральный член полученного выражения в силу сказанного выше обращается в нуль. Оставшийся интеграл снова берем по частям, и продолжаем этот процесс, пока показатель при t не снизится до нуля:

$$(t^k, p_n(t)) = -kt^{k-1} [(t^2-1)^n]^{(n-2)} \Big|_{-1}^{+1} +$$

$$+ k(k-1) \int_{-1}^{+1} t^{k-2} [(t^2-1)^n]^{(n-2)}\, dt =$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$= \pm k! \int_{-1}^{+1} [(t^2-1)^n]^{(n-k)}\, dt = \pm k! [(t^2-1)^n]^{(n-k-1)} \Big|_{-1}^{+1} = 0,$$

что и требуется.

теоремы ортогонализации дает

$$y_0(t)=1,\quad y_1(t)=t,\quad y_2(t)=t^2-\frac{1}{3},\quad y_3(t)=t^3-\frac{3}{5}t \text{ и т. д.}$$

Эти многочлены были введены в 1785 г. французским математиком Лежандром в связи с задачами теории потенциала. Общая формула для многочленов Лежандра была найдена Родригом в 1814 г. Именно, оказалось, что многочлен $y_n(t)$ с точностью до числового множителя равен многочлену

$$p_n(t)=\frac{d^n}{dt^n}[(t^2-1)^n] \qquad (n=0, 1, 2, \ldots). \tag{24}$$

Мы воспользуемся для доказательства этого предложения замечанием *8.63*. Именно, мы покажем, что многочлен $p_n(t)$ удовлетворяет условиям теоремы ортогонализации; в силу указанного замечания мы будем для каждого n иметь равенство $p_n(t)=c_n y_n(t)$, что нам и нужно.

а. *Линейная оболочка векторов $p_0(t)$, $p_1(t)$, ..., $p_n(t)$ совпадает с совокупностью всех многочленов до n-й степени.* В самом деле, как видно из формулы (24), многочлен $p_k(t)$ есть, очевидно, многочлен точно k-й степени от t; в частности,

$$\left.\begin{aligned}
&p_0(t)=a_{00},\\
&p_1(t)=a_{10}+a_{11}t,\\
&p_2(t)=a_{20}+a_{21}t+a_{22}t^2,\\
&\cdots\cdots\cdots\cdots\\
&p_k(t)=a_{k0}+a_{k1}t+\ldots+a_{kk}t^k,\\
&\cdots\cdots\cdots\cdots\cdots\\
&p_n(t)=a_{n0}+a_{n1}t+\ldots+a_{nk}t^k+\ldots+a_{nn}t^n,
\end{aligned}\right\} \tag{25}$$

причем старшие коэффициенты a_{00}, a_{11}, ..., a_{nn} отличны от нуля.

Таким образом, все многочлены $p_0(t)$, $p_1(t)$, ..., $p_n(t)$ входят в линейную оболочку функций $1, t, \ldots, t^n$, которая, очевидно, есть совокупность L_n всех многочленов от t не выше n-й степени. Так как матрица линейных соотношений (25) имеет определитель $a_{00}a_{11}\ldots a_{nn}$, отличный от нуля, то и, обратно, функции $1, t, t^2, \ldots, t^n$ могут быть линейно выражены через $p_0(t)$, $p_1(t)$, ..., $p_n(t)$; поэтому линейная оболочка $L[p_0(t), p_1(t), \ldots, p_n(t)]$ совпадает

с линейной оболочкой $L(1, t, t^2, \ldots, t^n)$ и, следовательно, совпадает с совокупностью L_n, что и требовалось.

б. *Вектор $p_n(t)$ ортогонален к подпространству L_{n-1}.* Достаточно проверить, что многочлен $p_n(t)$ ортогонален в $R_2(-1, 1)$ к функциям $1, t, \ldots, t^{n-1}$.

При доказательстве мы будем использовать формулу интегрирования по частям в определенном интеграле, предполагаемую известной из анализа. Производные, входящие в эту формулу, для многочленов суть те же самые производные, которые мы рассматривали в *6.73а* с чисто алгебраической точки зрения. В частности, многочлен

$$[(t^2-1)^n] = (t-1)^n (t+1)^n$$

в силу *6.73в* при $t = \pm 1$ имеет равные 0 производные порядков $0, 1, \ldots, n-1$.

Итак, будем вычислять скалярное произведение t^k и $p_n(t)$. Интегрируя по частям, мы получаем

$$(t^k, p_n(t)) = \int_{-1}^{+1} t^k [(t^2-1)^n]^{(n)}\, dt =$$

$$= t^k [(t^2-1)^n]^{(n-1)} \Big|_{-1}^{+1} - k \int_{-1}^{+1} t^{k-1} [(t^2-1)^n]^{(n-1)}\, dt.$$

Внеинтегральный член полученного выражения в силу сказанного выше обращается в нуль. Оставшийся интеграл снова берем по частям, и продолжаем этот процесс, пока показатель при t не снизится до нуля:

$$(t^k, p_n(t)) = -k t^{k-1} [(t^2-1)^n]^{(n-2)} \Big|_{-1}^{+1} +$$

$$+ k(k-1) \int_{-1}^{+1} t^{k-2} [(t^2-1)^n]^{(n-2)}\, dt =$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$= \pm k! \int_{-1}^{+1} [(t^2-1)^n]^{(n-k)}\, dt = \pm k! [(t^2-1)^n]^{(n-k-1)} \Big|_{-1}^{+1} = 0,$$

что и требуется.

Итак, мы доказали, что для каждого n многочлен $y_n(t)$ с точностью до числового множителя совпадает с многочленом $p_n(t)=[(t^2-1)^n]^{(n)}$.

Вычислим значение $p_n(1)$. Для этого применим к функции $(t^2-1)^n=(t+1)^n(t-1)^n$ правило n-кратного дифференцирования произведения:

$$[(t+1)^n(t-1)^n]^{(n)}=$$
$$=(t+1)^n[(t-1)^n]^{(n)}+C_n^1[(t+1)^n]'[(t-1)^n]^{(n-1)}+\ldots=$$
$$=(t+1)^n n!+C_n^1 n(t+1)^{n-1}n(n-1)\ldots 2(t-1)+\ldots$$

При подстановке $t=1$ все члены этой суммы, начиная со второго, обращаются в нуль. Мы получаем, следовательно,

$$p_n(1)=2^n n!.$$

Для вычислительных целей удобнее иметь значения наших ортогональных функций равными 1 при $t=1$. Чтобы достичь этого, мы должны ввести числовой множитель $\frac{1}{2^n n!}$. Именно полученные после этого многочлены и называются *многочленами Лежандра*; многочлен Лежандра степени n обозначается символом $P_n(t)$, так что

$$P_n(t)=\frac{1}{2^n n!}[(t^2-1)^n]^{(n)}.$$

§ 8.7. Определитель Грама

8.71. *Определителем Грама* называется определитель вида

$$G(x_1, x_2, \ldots, x_k)=\begin{vmatrix} (x_1, x_1) & (x_1, x_2) & \ldots & (x_1, x_k) \\ (x_2, x_1) & (x_2, x_2) & \ldots & (x_2, x_k) \\ \cdot & \cdot & \cdot & \cdot \\ (x_k, x_1) & (x_k, x_2) & \ldots & (x_k, x_k) \end{vmatrix},$$

где $x_1, x_2, \ldots, x_k$ — произвольные векторы евклидова пространства R. Мы знаем, что в случае линейно независимых векторов $x_1, x_2, \ldots, x_n$ этот определитель положителен (*7.96*).

Для вычисления определителя Грама применим к векторам $x_1, x_2, \ldots, x_k$ процесс ортогонализации. Пусть, например, $y_1 = x_1$ и вектор $y_2 = \alpha_1 y_1 + x_2$ ортогонален к y_1. Заменим на всех местах определителя вектор x_1 на y_1. Далее прибавим ко второму столбцу первый столбец определителя Грама, умноженный на α_1 (отнеся α_1 ко второму множителю скалярных произведений), и затем прибавим ко второй строке первую строку определителя, умноженную на α_1 (отнеся α_1 к первому множителю скалярных произведений). В результате на всех тех местах определителя, где был вектор x_2, будет находиться вектор y_2.

Далее, пусть $y_3 = \beta_1 y_1 + \beta_2 y_2 + x_3$ ортогонален к y_1 и y_2; прибавим к третьему столбцу первый, умноженный на β_1, и второй, умноженный на β_2; ту же операцию произведем со строками. В результате x_3 на всех местах будет заменен на y_3. Мы можем продолжать этот процесс далее, пока не дойдем до последнего столбца. Так как наши операции не изменяют величины определителя, то в результате мы получим

$$G(x_1, x_2, \ldots, x_k) = \begin{vmatrix} (y_1, y_1) & 0 & \ldots & 0 \\ 0 & (y_2, y_2) & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & (y_k, y_k) \end{vmatrix} = $$
$$= (y_1, y_1)(y_2, y_2) \ldots (y_k, y_k). \qquad (26)$$

В силу результата *8.62* мы получаем следующее неравенство:

$$0 \leqslant G(x_1, x_2, \ldots, x_k) \leqslant (x_1, x_1)(x_2, x_2) \ldots (x_k, x_k). \qquad (27)$$

Выясним, в каких случаях величина $G(x_1, x_2, \ldots, x_k)$ может принимать крайние значения 0 или $(x_1, x_1) \ldots (x_k, x_k)$.

Из выражения определителя Грама (26) вытекает, что он равен нулю в том и только в том случае, когда один из векторов $y_1, y_2, \ldots, y_k$ равен нулю. Но в силу *8.62* это эквивалентно линейной зависимости векторов $x_1, x_2, \ldots, x_k$. С другой стороны, равенство определителя Грама правой части неравенства (27) возможно в силу формулы (26) и *8.62* только в том случае, когда векторы $x_1, x_2, \ldots, x_k$ взаимно ортогональны. Итак, мы доказали следующую теорему:

Т е о р е м а. (Т е о р е м а о б о п р е д е л и т е л е Г р а м а.) *Определитель Грама векторов $x_1, x_2, \ldots, x_k$ равен нулю,* *если эти векторы линейно зависимы, и положителен, если они линейно независимы; он равен произведению квадратов длин векторов $x_1, x_2, \ldots, x_k$, если они взаимно ортогональны, в противном случае он меньше этой величины.*

8.72. Объем k-мерного гиперпараллелепипеда. Площадь параллелограмма, как известно из планиметрии, равна произведению его основания на высоту. Если параллелограмм построен на двух векторах x_1, x_2, то можно принять за основание длину вектора x_1, за высоту — длину перпендикуляра, опущенного из конца вектора x_2 на ось вектора x_1.

Аналогично объем параллелепипеда, построенного на векторах x_1, x_2, x_3, равен произведению площади основания на высоту; площадь основания есть площадь параллелограмма, построенного на векторах x_1, x_2, а высота есть длина перпендикуляра, опущенного из конца вектора x_3 на плоскость векторов x_1, x_2.

Эти соображения делают естественным следующее индуктивное определение объема k-мерного гиперпараллелепипеда в евклидовом пространстве.

Пусть дана система векторов $x_1, x_2, \ldots, x_k$ в евклидовом пространстве R. Обозначим через h_j перпендикуляр, опущенный из конца вектора x_{j+1} на подпространство $L(x_1, x_2, \ldots, x_j)$ $(j=1, 2, \ldots, k-1)$. Введем, далее, следующие обозначения:

$V_1 = |x_1|$ (одномерный объем — длина вектора x_1),
$V_2 = V_1 \cdot |h_1|$ (двумерный объем — площадь параллелограмма, построенного на векторах x_1, x_2),
$V_3 = V_2 \cdot |h_2|$ (трехмерный объем — объем параллелепипеда, построенного на векторах x_1, x_2, x_3),
. . . .
$V_k = V_{k-1} \cdot |h_{k-1}|$ (k-мерный объем — объем гиперпараллелепипеда, построенного на векторах $x_1, x_2, \ldots, x_k$).

Очевидно, что объем V_k может быть вычислен по формуле

$$V_k \equiv V[x_1, x_2, \ldots, x_k] = |x_1| \cdot |h_1| \ldots |h_{k-1}|.$$

Используя формулу (26), мы можем выразить величину V_k через векторы $x_1, x_2, \ldots, x_k$:

$$V_k^2 = \begin{vmatrix} (x_1, x_1) & (x_1, x_2) & \ldots & (x_1, x_k) \\ (x_2, x_1) & (x_2, x_2) & \ldots & (x_2, x_k) \\ \cdot & \cdot & \cdot & \cdot \\ (x_k, x_1) & (x_k, x_2) & \ldots & (x_k, x_k) \end{vmatrix}.$$

Таким образом, *определитель Грама от k векторов $x_1, x_2, \ldots, x_k$ равен квадрату объема k-мерного гиперпараллелепипеда, построенного на этих векторах.*

8.73. Пусть $\xi_i^{(j)}$ — координаты вектора x_j относительно ортогонального и нормированного базиса

$$e_1, e_2, \ldots, e_n \; (j = 1, 2, \ldots, k; \; i = 1, 2, \ldots, n).$$

Выражая скалярные произведения векторов через координаты, мы получим следующую формулу:

$$V_k^2 = \begin{vmatrix} \xi_1^{(1)}\xi_1^{(1)} + \ldots + \xi_n^{(1)}\xi_n^{(1)} & \ldots & \xi_1^{(1)}\xi_1^{(k)} + \ldots + \xi_n^{(1)}\xi_n^{(k)} \\ \xi_1^{(2)}\xi_1^{(1)} + \ldots + \xi_n^{(2)}\xi_n^{(1)} & \ldots & \xi_1^{(2)}\xi_1^{(k)} + \ldots + \xi_n^{(2)}\xi_n^{(k)} \\ \cdot & \cdot & \cdot \\ \xi_1^{(k)}\xi_1^{(1)} + \ldots + \xi_n^{(k)}\xi_n^{(1)} & \ldots & \xi_1^{(k)}\xi_1^{(k)} + \ldots + \xi_n^{(k)}\xi_n^{(k)} \end{vmatrix}.$$

Матрица определителя V_k^2, как легко убедиться, есть произведение $k \times n$-матрицы A из координат векторов $x_1, x_2, \ldots, x_k$

$$A = \begin{Vmatrix} \xi_1^{(1)} & \xi_2^{(1)} & \ldots & \xi_n^{(1)} \\ \xi_1^{(2)} & \xi_2^{(2)} & \ldots & \xi_n^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ \xi_1^{(k)} & \xi_2^{(k)} & \ldots & \xi_n^{(k)} \end{Vmatrix} \tag{28}$$

и транспонированной к ней $n \times k$-матрицы

$$A' = \begin{Vmatrix} \xi_1^{(1)} & \xi_1^{(2)} & \ldots & \xi_1^{(k)} \\ \xi_2^{(1)} & \xi_2^{(2)} & \ldots & \xi_2^{(k)} \\ \cdot & \cdot & \cdot & \cdot \\ \xi_n^{(1)} & \xi_n^{(2)} & \ldots & \xi_n^{(k)} \end{Vmatrix}.$$

Применяя формулу *4.54* (14) к определителю V_k^2, находим

$$V_k^2 = M_{1,\ 2,\ \ldots,\ k}^{1,\ 2,\ \ldots,\ k}(AA') = \sum_{i_1,\ \ldots,\ i_k} M_{i_1,\ \ldots,\ i_k}^{1,\ \ldots,\ k}(A)\, M_{1,\ \ldots,\ k}^{i_1,\ \ldots,\ i_k}(A'),$$

где суммирование распространяется на все наборы номеров $i_1 < \ldots < i_k$, принимающих значения от 1 до n.

Так как

$$M_{1,\ \ldots,\ k}^{i_1,\ \ldots,\ i_k}(A') = M_{i_1,\ \ldots,\ i_k}^{1,\ \ldots,\ k}(A),$$

то

$$V_k^2 = \sum_{i_1,\ \ldots,\ i_k} [M_{i_1,\ \ldots,\ i_k}^{1,\ \ldots,\ k}(A)]^2.$$

Итак, *квадрат объема k-мерного гиперпараллелепипеда, построенного на векторах x_j ($j = 1, 2, \ldots, k$), равен сумме квадратов всех миноров k-го порядка в матрице из координат векторов x_j относительно (любого) ортогонального и нормированного базиса $e_1, e_2, \ldots, e_n$.*

8.74. В случае $k = n$ матрица $\|\xi_i^{(j)}\|$ имеет только один минор k-го порядка, равный определителю матрицы $\|\xi_i^{(j)}\|$.

Поэтому *объем n-мерного гиперпараллелепипеда, построенного на векторах $x_1, x_2, \ldots, x_n$, равен (по абсолютной величине) определителю из координат векторов x_i ($i = 1, 2, \ldots, n$) относительно (любого) ортогонального и нормированного базиса.*

8.75. Неравенство Адамара. Из результата *8.74* можно получить одну важную оценку абсолютной величины произвольного определителя k-го порядка

$$D = \begin{vmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1k} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ \xi_{k1} & \xi_{k2} & \cdots & \xi_{kk} \end{vmatrix}.$$

Будем рассматривать числа $\xi_{i1}, \xi_{i2}, \ldots, \xi_{ik}$ ($i = 1, 2, \ldots, k$) как координаты вектора x_i в ортогональном нормированном базисе k-мерного евклидова пространства. Результат *8.74* дает нам возможность истолковать абсолютную величину определителя D как объем k-мерного гиперпараллелепипеда, построенного на векторах $x_1, x_2, \ldots, x_k$, и использовать выражение объема через определитель Грама $D^2 = G(x_1, x_2, \ldots, x_k)$. Применяя теорему *8.71*, получаем

$$D^2 \leqslant (x_1, x_1)(x_2, x_2) \ldots (x_k, x_k) = \prod_{i=1}^{k} \sum_{j=1}^{k} \xi_{ij}^2.$$

Это неравенство носит название *неравенства Адамара*.

Отметим, что оно обращается в равенство в силу теоремы *8.71* в том и только в том случае, когда векторы $x_1, x_2, \ldots, x_k$ взаимно ортогональны.

Неравенство Адамара имеет прозрачный геометрический смысл: *объем гиперпараллелепипеда не превосходит произведения длин его ребер; он равен этому произведению в том и только в том случае, когда его ребра взаимно ортогональны.*

§ 8.8. Несовместные системы линейных уравнений и метод наименьших квадратов

8.81. Пусть дана несовместная система линейных уравнений

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m &= b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m &= b_n. \end{aligned}\right\} \quad (29)$$

Поскольку она несовместна, ее нельзя решить, т. е. нельзя найти такие числа $c_1, c_2, \ldots, c_m$, чтобы при подстановке этих чисел вместо неизвестных $x_1, x_2, \ldots, x_m$ удовлетворялись бы все уравнения системы (29).

Если подставить какие-нибудь числа $\xi_1, \xi_2, \ldots, \xi_m$ вместо неизвестных $x_1, x_2, \ldots, x_m$ в левые части уравнений (29), то мы получим результаты $\gamma_1, \gamma_2, \ldots, \gamma_n$, отличные от чисел $b_1, b_2, \ldots, b_n$. Поставим себе задачу: при известных значениях чисел a_{jk} и b_k $(k = 1, \ldots, n;$ $j = 1, \ldots, m)$ *определить числа* $\xi_1, \xi_2, \ldots, \xi_m$ *так, чтобы* *квадратичное уклонение результатов* $\gamma_1, \gamma_2, \ldots, \gamma_n$ *от данных величин* $b_1, b_2, \ldots, b_n$, *определяемое выражением*

$$\delta^2 = \sum_{j=1}^{n} (\gamma_j - b_j)^2, \tag{30}$$

оказалось наименьшим из возможных, а также найти это минимальное уклонение.

Такая задача возникает на практике, когда, например, коэффициенты ξ_j линейной зависимости величины b от величин $a_1, a_2, \ldots, a_m$

$$b = \xi_1 a_1 + \xi_2 a_2 + \ldots + \xi_m a_m$$

должны быть найдены из результатов измерений величин a_j $(j = 1, 2, \ldots, m)$ и соответствующих значений b. Если при i-м измерении получены значения a_{ij} для величины a_j и b_i для величины b, то мы должны составить уравнение

$$\xi_1 a_{i1} + \xi_2 a_{i2} + \ldots + \xi_m a_{im} = b_i; \tag{31}$$

n измерений приводят к системе n уравнений (31), т. е. к системе вида (29). Эта система вследствие неизбежных ошибок измерений будет, вообще говоря, несовместной, и задача определения коэффициентов $\xi_1, \xi_2, \ldots, \xi_m$, таким образом, не сводится к задаче решения системы уравнений (29). Возникает задача определить коэффициенты ξ_j так, чтобы каждое уравнение удовлетворялось, хотя бы и приблизительно, но с общей наименьшей погрешностью. Если за меру погрешности взять среднее квадратичное из уклонений величин

$$\gamma_j = \sum_{i=1}^{m} a_{ij} \xi_i$$

от известных b_j, определяемое формулой (30), то мы и придем к сформулированной выше задаче. Знание величины δ^2 в этом случае также полезно: оно помогает оценить надежность измерений.

8.82. Решение получается немедленно, если истолковать задачу геометрически в вещественном пространстве R_n. Рассмотрим m векторов $a_1, a_2, \ldots, a_m$, компоненты которых выписаны в столбцах системы (29)

$$a_1 = \{a_{11}, a_{21}, \ldots, a_{n1}\}, \ \ldots, \ a_m = \{a_{1m}, a_{2m}, \ldots, a_{nm}\}.$$

Составляя линейную комбинацию $\xi_1 a_1 + \ldots + \xi_m a_m$, мы получим вектор $\gamma = \{\gamma_1, \ldots, \gamma_m\}$. Нужно определить числа $\xi_1, \ldots, \xi_m$ так, чтобы вектор γ по норме имел наименьшее возможное отклонение от заданного вектора $b = \{b_1, b_2, \ldots, b_n\}$.

Совокупность всех линейных комбинаций векторов $a_1, \ldots, a_m$ образует подпространство $L = L(a_1, a_2, \ldots, a_m)$. В этом подпространстве наименьшее расстояние до вектора b имеет проекция вектора b на подпространство L. Числа $\xi_1, \ldots, \xi_m$, следовательно, должны быть выбраны так, чтобы линейная комбинация $\xi_1 a_1 + \ldots + \xi_m a_m$ привела к проекции вектора b на подпространство L. Но решение этой задачи нам известно; оно дается формулами *8.53*, именно, мы имеем

$$\xi_j = \frac{1}{D} \begin{vmatrix} (a_1, a_1) \ldots (a_{j-1}, a_1)\,(b, a_1)\,(a_{j+1}, a_1) \ldots (a_m, b_1) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ (a_1, a_m) \ldots (a_{j-1}, a_m)\,(b, a_m)\,(a_{j+1}, a_m) \ldots (a_m, b_m) \end{vmatrix},$$

где D — определитель Грама $G(a_1, a_2, \ldots, a_m)$.

8.83. Результаты *8.72* дают нам возможность оценить и само уклонение δ. Действительно, величина δ есть высота $(m+1)$-мерного параллелепипеда, построенного на векторах $a_1, \ldots, a_m, b$, и равна поэтому отношению объемов

$$\frac{V(a_1, a_2, \ldots, a_m, b)}{V(a_1, a_2, \ldots, a_m)}.$$

Записывая каждый из объемов с помощью определителя Грама, находим окончательно

$$\delta^2 = \frac{G(a_1, a_2, \ldots, a_m, b)}{G(a_1, a_2, \ldots, a_m)}.$$

Тем самым поставленная нами задача полностью решена.

8.84. В вычислительной практике часто встречается следующая задача (интерполирование с наименьшей квадратичной погрешностью): *на отрезке $a \leqslant t \leqslant b$ дана функция $f_0(t)$; указать многочлен $P(t)$ степени k, $k < n$, для которого квадратичное уклонение от функции $f_0(t)$, измеряемое величиной*

$$\delta^2(f_0, P) = \sum_{j=0}^{n} [f_0(t_j) - P(t_j)]^2,$$

становится наименьшим. Здесь $t_0, t_1, \ldots, t_n$ — некоторые фиксированные точки отрезка $a \leqslant t \leqslant b$.

Простое решение этой задачи на основании геометрических соображений было предложено М. А. Красносельским.

Введем евклидово пространство R, образованное из функций $f(t)$, рассматриваемых только в точках $t_0, t_1, \ldots, t_n$, со скалярным произведением

$$(f, g)=\sum_{j=0}^{n} f(t_j)\, g(t_j).$$

Тогда наша задача сводится к определению проекции вектора $f_0(t)$ на подпространство всех многочленов степени, не превосходящей k. Коэффициенты искомого полинома $P_0(t)=\xi_0+\xi_1 t+\ldots+\xi_k t^k$ даются теми же формулами, что и в разобранной выше задаче:

$$\xi_j=\frac{1}{D}\begin{vmatrix} (1,\ 1) & (t,\ 1) & \ldots & (t^{j-1},\ 1) & (f_0,\ 1) & (t^{j+1},\ 1) & \ldots & (t^k,\ 1) \\ (1,\ t) & (t,\ t) & \ldots & (t^{j-1},\ t) & (f_0,\ t) & (t^{j+1},\ t) & \ldots & (t^k,\ t) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (1,\ t^k) & (t,\ t^k) & \ldots & (t^{j-1},\ t^k) & (f_0,\ t^k) & (t^{j+1},\ t^k) & \ldots & (t^k,\ t^k) \end{vmatrix},$$

где D определитель Грама $D(1,\ t,\ \ldots,\ t^k)$.

Само наименьшее квадратичное уклонение δ^2 можно вычислить по формуле

$$\delta^2(f_0,\ P)=\frac{D(1,\ t,\ \ldots,\ t^k,\ f)}{D(1,\ t,\ \ldots,\ t^k)}.$$

§ 8.9. Сопряженные операторы и изометрия

8.91. Операторы, сопряженные относительно формы (x, y).

Сформулируем здесь результаты § 7.6 о связи между линейными операторами и билинейными формами в применении к случаю, когда фиксирована форма (x, y) — скалярное произведение векторов x и y. Пусть A и B — линейные операторы в евклидовом пространстве R_n. Образуем билинейные формы $A(x, y)$ и $B(x, y)$ по формулам

$$A(x,\ y)=(Ax,\ y),\quad B(x,\ y)=(x,\ By). \tag{32}$$

Так как каноническим базисом формы (x, y) является любой ортонормальный базис пространства и канонические коэффициенты формы (x, y) в любом таком базисе все равны 1, то в силу *7.61* матрица $\|a_{jk}\|$ формы $A(x, y)$ в любом ортонормальном базисе совпадает с матрицей $\|a_k^{(j)}\|$ оператора A, а матрица $\|b_{jk}\|$ формы $B(x, y)$ *транспонирована* по отношению к матрице $\|b_j^{(k)}\|$ оператора B. Обратно (*7.62*), если в пространстве R_n заданы билинейные формы $A(x, y)$ и $B(x, y)$, то существуют и единственны такие линейные операторы A и B, что выполнены равенства (32). Далее, применяя результаты *7.63* к форме (x, y), получаем следующий результат:

Теорема. *Для каждого линейного оператора* A *в евклидовом n-мерном пространстве* R_n *существует и единствен сопряженный оператор — линейный оператор* A′, *удовлетворяющий уравнению*

$$(\mathrm{A}x,\ y) = (x,\ \mathrm{A}'y)$$

при любых x *и* y *из* R_n. *Матрица оператора* A′ *в любом ортогональном и нормированном базисе пространства* R_n *транспонирована по отношению к матрице оператора* A.

8.92. С помощью операции сопряжения в евклидовом пространстве вводятся следующие классы операторов.

а. *Симметричные операторы*, определяемые равенством

$$\mathrm{A}' = \mathrm{A}.$$

Симметричный оператор характеризуется тем, что в ортонормальном базисе его матрица не меняется при транспонировании.

б. *Антисимметричные операторы*, определяемые равенством

$$\mathrm{A}' = -\mathrm{A}.$$

Антисимметричный оператор характеризуется тем, что в ортонормальном базисе его матрица после транспонирования изменяет знак.

в. *Нормальные операторы*, определяемые равенством

$$\mathrm{A}'\mathrm{A} = \mathrm{A}\mathrm{A}'.$$

Очевидно, нормальные операторы включают в себя и симметричные и антисимметричные. Изучение введенных классов операторов будет произведено в § 9.4.

8.93. Теперь сформулируем для евклидова пространства R_n результаты *7.73—7.76* относительно инвариантных операторов. Рассматриваются линейные обратимые отображения Q пространства R_n в себя, не изменяющие скалярного произведения:

$$(\mathrm{Q}x,\ \mathrm{Q}y) = (x,\ y).$$

Эти отображения, названные в *7.73* инвариантными относительно формы (x, y), здесь будем называть *изометрическими*. Изометрический оператор Q характеризуется равенством *7.73*(33)

$$Q'Q = E,$$

где Q' — оператор, сопряженный к Q относительно формы (x, y), т. е. оператор, сопряженный к Q в смысле *8.91*. Обратный оператор $Q^{-1} = Q'$ к изометрическому оператору Q также изометричен. Произведение двух изометрических операторов является изометрическим оператором (*7.74*).

Согласно *7.75* изометрический оператор Q характеризуется тем, что любой ортонормированный базис $e_1, \ldots, e_n$ переводится этим оператором снова в ортонормированный базис $f_1 = Qe_1, \ldots, f_n = Qe_n$. Матрица $Q = \| q_i^{(j)} \|$ изометрического оператора Q в любом ортонормированном базисе называется *ортогональной матрицей*. Ортогональная матрица характеризуется условиями *7.76*(35), имеющими в данном случае вид

$$\sum_{i=1}^{n} q_j^{(i)} q_k^{(i)} = \begin{cases} 1 & \text{при} \quad j = k, \\ 0 & \text{при} \quad j \neq k, \end{cases}$$

или же условиями 7.76(35'), имеющими вид

$$\sum_{k=1}^{m} q_k^{(j)} q_k^{(m)} = \begin{cases} 1 & \text{при} \quad j = m, \\ 0 & \text{при} \quad j \neq m, \end{cases}$$

т. е. сумма квадратов элементов любой строки (любого столбца) равна 1, сумма произведений соответствующих элементов двух разных строк (столбцов) равна 0.

8.94. В силу равенства $Q^{-1} = Q'$, обращением формул перехода от ортонормального базиса $e_1, \ldots, e_n$ к другому ортонормальному базису $f_1, \ldots, f_n$

$$\left.\begin{aligned} f_1 &= q_1^{(1)} e_1 + \ldots + q_n^{(1)} e_n, \\ &\ldots\ldots\ldots\ldots \\ f_n &= q_1^{(n)} e_1 + \ldots + q_n^{(n)} e_n \end{aligned}\right\} \qquad (33)$$

служат формулы

$$\left.\begin{array}{l} e_1 = q_1^{(1)} f_1 + \ldots + q_1^{(n)} f_n, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ e_n = q_n^{(1)} f_1 + \ldots + q_n^{(n)} f_n. \end{array}\right\} \qquad (34)$$

В силу *5.31* координаты η_k вектора x в базисе $f_1, \ldots, f_n$ выражаются через координаты ξ_j этого же вектора в базисе $e_1, \ldots, e_n$ по формулам

$$\left.\begin{array}{l} \eta_1 = q_1^{(1)} \xi_1 + \ldots + q_n^{(1)} \xi_n, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \eta_n = q_1^{(n)} \xi_1 + \ldots + q_n^{(n)} \xi_n \end{array}\right\} \qquad (35)$$

с обратными формулами

$$\left.\begin{array}{l} \xi_1 = q_1^{(1)} \eta_1 + \ldots + q_1^{(n)} \eta_n, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \xi_n = q_n^{(1)} \eta_1 + \ldots + q_n^{(n)} \eta_n. \end{array}\right\} \qquad (36)$$

8.95. Рассмотрим здесь еще следующую задачу. Пусть заданы $m < n$ числовых строк $q_i^{(j)}$ $(i = 1, \ldots, n;\ j = 1, \ldots, m)$, удовлетворяющих условиям

$$\sum_{i=1}^{n} q_i^{(j)} q_i^{(k)} = \begin{cases} 1 & \text{при} \quad j = k, \\ 0 & \text{при} \quad j \neq k. \end{cases}$$

Требуется найти еще $n - m$ числовых строк $q_i^{(s)}$, $s = m+1, \ldots, n$, так, чтобы $n \times n$-матрица $\| q_i^{(j)} \|$ $(i, j = 1, \ldots, n)$ была ортогональной.

Решение получается просто из геометрических соображений. Интерпретируя заданные строки $q_i^{(j)}$ как координаты m векторов в евклидовом пространстве R_n со скалярным произведением (*8.22б*)

$$((\xi_1, \ldots, \xi_n), (\eta_1, \ldots, \eta_n)) = \sum_{j=1}^{n} \xi_j \eta_j,$$

мы видим, что нам заданы m ортогональных и нормированных векторов $q_1, \ldots, q_m$ и задача состоит в том, чтобы дополнить их до ортонормального базиса в пространстве R_n. В этой геометрической формулировке задача становится очевидно разрешимой: например, можно дополнить векторы

$q_1, \ldots, q_m$ произвольно до системы n линейно независимых векторов и далее провести ортонормализацию.

8.96. Рассмотрим еще некоторые свойства симметричных операторов.

а. *Если подпространство* $\mathrm{R}' \subset \mathrm{R}$ *инвариантно относительно оператора* A, *то ортогональное дополнение подпространства* R', *на основании 7.65, инвариантно относительно сопряженного оператора* A'.

Поэтому, если A — симметричный оператор, то из инвариантности относительно A подпространства R' следует инвариантность относительно A и ортогонального дополнения к подпространству R'.

б. Покажем, что на плоскости ($n=2$) *всякий симметричный оператор имеет собственный вектор.*

Действительно, в данном случае уравнение для собственных значений

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0$$

есть квадратное уравнение с дискриминантом

$$(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{21}a_{12}) = (a_{11}-a_{22})^2 + 4a_{12}^2 \geqslant 0,$$

поэтому его корни вещественны.

в. Пользуясь *а* и *б* и существованием инвариантной плоскости у всякого оператора в вещественном пространстве (вытекающим из выражения вещественной жордановой формы *6.63* (18)), можно доказать, что *в пространстве* R_n *всякий симметричный оператор имеет ортогональный базис из собственных векторов.* Мы в дальнейшем получим этот результат из общих соображений, не опираясь на вещественную жорданову форму (*9.45*).

ЗАДАЧИ

1. Назовем скалярным произведением двух векторов пространства V_3 произведение их длин. Будет ли пространство евклидовым?

2. А если назвать скалярным произведением двух векторов того же пространства произведение их длин на куб косинуса угла между ними?

3. А если назвать скалярным произведением удвоенное обычное скалярное произведение этих векторов?

4. Найти угол между противоположными ребрами правильного тетраэдра.

5. Найти углы в треугольнике, образованном в пространстве $R_2(-1, 1)$ векторами $x_1(t)=1$, $x_2(t)=t$ и $x_3(t)=1-t$.

6. Написать неравенства треугольника в пространстве $C_2(a, b)$.

7. Определить косинусы углов между прямой $\xi_1=\xi_2=\ldots=\xi_n$ и осями координат в пространстве R_n.

8. В пространстве R_4 разложить вектор f на сумму двух векторов, один из которых лежит в линейной оболочке векторов b_i, а другой ортогонален к этому подпространству:

а) $f=(5, 2, -2, 2)$, $b_1=(2, 1, 1, -1)$, $b_2=(1, 1, 3, 0)$;

б) $f=(-3, 5, 9, 3)$, $b_1=(1, 1, 1, 1)$, $b_2=(2, -1, 1, 1)$, $b_3=(2, -7, -1, -1)$.

9. Доказать, что из всех векторов подпространства R′(*8.51*) наименьший угол с вектором f образует вектор g.

10. Доказать, что если вектор g_0 в подпространстве R′ ортогонален к проекции g вектора f на это подпространство, то g_0 ортогонален и к самому f.

11. Показать, что перпендикуляр, опущенный из начала координат на гиперплоскость H, имеет наименьшую длину из всех векторов, соединяющих начало координат с этой гиперплоскостью.

12. В пространстве V_3 с базисом $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ задана система векторов

$$\boldsymbol{x}_1=\boldsymbol{i},\quad \boldsymbol{x}_2=2\boldsymbol{i},\quad \boldsymbol{x}_3=3\boldsymbol{i},\quad \boldsymbol{x}_4=4\boldsymbol{i}-2\boldsymbol{j},\quad \boldsymbol{x}_5=-\boldsymbol{i}+10\boldsymbol{j},\quad \boldsymbol{x}_6=\boldsymbol{i}+\boldsymbol{j}+5\boldsymbol{k}.$$

Построить ортогональные векторы $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_6$ (*8.61*).

13. В трехмерном подпространстве пространства R_4, порожденном векторами (1, 2, 1, 3), (4, 1, 1, 1), (3, 1, 1, 0), построить ортогональный базис, используя метод теоремы ортогонализации.

14. Пусть даны подпространства R′ и R″ евклидова пространства R. Рассмотрим перпендикуляры, опущенные из концов единичных векторов $e' \in$ R′ на R″, и пусть m(R′, R″) — максимум из их длин; аналогично определяем величину m(R″, R′). Неотрицательная величина $\theta=\max\{m(\mathrm{R}', \mathrm{R}''), m(\mathrm{R}'', \mathrm{R}')\}$ называется *раствором* подпространств R′ и R″. Показать, что при $\theta<1$ размерности подпространств R′ и R″ равны (М. А. Красносельский и М. Г. Крейн).

15. Найти старший коэффициент A_n многочлена Лежандра $P_n(t)$.

16. Показать, что многочлен Лежандра $P_n(t)$ есть четная функция при четном n и нечетная функция при нечетном n. Найти, в частности, $P_n(-1)$.

17. Доказать, что в разложении многочлена $tP_{n-1}(t)$ по многочленам Лежандра

$$tP_{n-1}(t)=a_0P_0(t)+a_1P_1(t)+\ldots+a_nP_n(t)$$

коэффициенты $a_0, a_1, \ldots, a_{n-3}$ и a_{n-1} равны нулю.

18. Найдя коэффициенты a_{n-2} и a_n в разложении многочлена $tP_{n-1}(t)$ из задачи 17, получить рекуррентную формулу

$$nP_n(t)=(2n-1)\,tP_{n-1}(t)-(n-1)\,P_{n-2}(t).$$

19. Найти многочлен $Q(t)=t^n+b_1t^{n-1}+\ldots+b_{n-1}t+b_n$, для которого интеграл

$$\int_{-1}^{+1} Q^2(t)\,dt$$

достигает наименьшего значения.

20. Найти норму многочлена Лежандра $P_n(t)$.

21. Показать, что для любого линейного оператора A, действующего в n-мерном евклидовом пространстве R_n, отношение

$$k(A)=\frac{V[Ax_1,\ Ax_2,\ \ldots,\ Ax_n]}{V[x_1,\ x_2,\ \ldots,\ x_n]}$$

постоянно (т. е. не зависит от выбора векторов $x_1, x_2, \ldots, x_n$), и найти его величину («коэффициент искажения»).

22. Показать, что для любых двух линейных операторов A и B имеет место равенство $k(AB)=k(A)\,k(B)$.

23. Пусть $x_1, x_2, \ldots, x_k, y, z$ — векторы евклидова пространства R. Доказать неравенство

$$\frac{V[x_1,\ x_2,\ \ldots,\ x_k,\ y,\ z]}{V[x_1,\ x_2,\ \ldots,\ x_k,\ y]}\leqslant\frac{V[x_1,\ x_2,\ \ldots,\ x_k,\ z]}{V[x_1,\ x_2,\ \ldots,\ x_k]}.$$

24. Пусть $x_1, x_2, \ldots, x_m$ — векторы евклидова пространства R. Доказать неравенство

$$V[x_1,\ x_2,\ \ldots,\ x_m]\leqslant\prod_{k=1}^{m}\{V[x_1,\ \ldots,\ x_{k-1},\ x_{k+1},\ \ldots,\ x_m]\}^{\frac{1}{m-1}}.$$

Каков его геометрический смысл?

25. (Продолжение.) Доказать неравенства, уточняющие неравенство Адамара:

$$V[x_1,\ x_2,\ \ldots,\ x_m]\leqslant\prod_{k=1}^{m}\{V[x_1,\ \ldots,\ x_{k-1},\ x_{k+1},\ \ldots,\ x_m]\}^{\frac{1}{m-1}}\leqslant$$

$$\leqslant\prod_{1\leqslant k<l\leqslant m}\{V[x_1,\ \ldots,\ x_{k-1},\ x_{k+1},\ldots,\ x_{l-1},\ x_{l+1},\ \ldots,\ x_m]\}^{\frac{1\cdot 2}{(m-1)(m-2)}}\leqslant$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\leqslant\prod_{1\leqslant s_1<s_2<\ldots<s_r\leqslant m}\{V[x_{s_1},\ x_{s_2},\ \ldots,\ x_{s_r}]\}^{\frac{1\cdot 2\ldots(n-r)}{(m-1)(m-2)\ldots r}}\leqslant$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\leqslant\prod_{1\leqslant s_1<s_2\leqslant m}\{V[x_{s_1},\ x_{s_2}]\}^{\frac{1}{m-1}}\leqslant\prod_{s=1}^{m}|x_s|$$

(М. К. Фаге).

26. Если $|a_{ik}| \leqslant M$, то по неравенству Адамара $\det \| a_{ik} \| \leqslant \leqslant M^n n^{n/2}$. Показать, что эта оценка не может быть улучшена для $n = 2^m$.

27. Если N (A) и T (A) — соответственно нуль-многообразие и область значений оператора A, то ортогональные дополнения этих подпространств суть соответственно область значений и нуль-многообразие оператора A.

28. Показать, что для каждого элемента a_{ik} ортогональной матрицы A соответствующее алгебраическое дополнение $A_{ik} = a_{ik} \det A$.

29. Показать, что сумма квадратов всех миноров k-го порядка, расположенных в k фиксированных строках ортогональной матрицы, равна 1; сумма произведений всех миноров k-го порядка в одной группе строк на соответствующие миноры k-го порядка другой группы строк равна 0.

30. Некоторый линейный оператор Q сохраняет длину каждого вектора; показать, что он является изометрическим.

31. Оператор A, сохраняющий ортогональность любых двух векторов x, y (так что из $(x, y) = 0$ вытекает $(\mathrm{A}x, \mathrm{A}y) = 0$), называется *равноугольным оператором*. Всякий изометрический оператор является равноугольным. Кроме того, равноугольными являются оператор подобия ($\mathrm{A}x = \lambda x$ для любого x) и произведение оператора подобия на изометрический оператор. Доказать, что всякий равноугольный оператор есть произведение оператора подобия и изометрического оператора.

32. Показать, что при $n \geqslant 3$ всякий линейный оператор Q, действующий в n-мерном пространстве R_n и не изменяющий площади любого параллелограмма, так что

$$V[x, y] = V[\mathrm{Q}x, \mathrm{Q}y],$$

есть изометрический оператор.

33. Показать, что при $k < n$ всякий линейный оператор Q, действующий в n-мерном пространстве R_n и не изменяющий объема лю- [redacted] места, так как в этом случае всякий оператор Q с $\det Q = \pm 1$ будет удовлетворять условию задачи.

34. В евклидовом пространстве R_n заданы конечные множества $F = \{x_1, x_2, \ldots, x_k\}$ и $G = \{y_1, y_2, \ldots, y_k\}$. Для того чтобы существовал изометрический оператор Q, переводящий одновременно каждый вектор x_i в соответствующий вектор y_i $(i = 1, 2, \ldots, k)$, необходимо и достаточно, чтобы имели место равенства

$$(x_i, x_j) = (y_i, y_j) \qquad (i, j = 1, 2, \ldots, k).$$

35. Углы между двумя подпространствами. В евклидовом пространстве R заданы два подпространства R′ и R″. Пусть нормированный вектор e' пробегает единичную сферу подпространства R′, а нормированный вектор e'' независимо от e' пробегает единичную сферу подпространства R″. Угол между e' и e'' на некоторой паре $e' = e'_1$, $e'' = e''_1$ достигает минимума, который мы обозначим через φ_1. Пусть, далее, e' меняется по своей единичной сфере, оставаясь ортогональным к e'_1, и e'' меняется по своей единичной сфере, оставаясь ортогональным к e''_1. Угол между e' и e'' при этих условиях достигнет минимума $\varphi_2 \geqslant \varphi_1$ на некоторой паре $e' = e'_2$, $e'' = e''_2$. Пусть, далее, e' меняется по своей единичной сфере, оставаясь ортогональным к e'_1 и e'_2, и e'' меняется по своей единичной сфере, оставаясь ортогональным к e''_1 и e''_2; мы получим новый минимум угла $\varphi_3 \geqslant \varphi_2$ и новую пару e'_3 и e''_3. Продолжая этот процесс далее, мы получим некоторую совокупность углов $\varphi_1, \varphi_2, \ldots, \varphi_k$, число которых равно наименьшей из размерностей подпространств R' и R''. Углы $\varphi_1, \varphi_2, \ldots$ $\ldots, \varphi_k$ называются *углами между подпространствами* R' и R''. Показать, что: а) углы $\varphi_1, \varphi_2, \ldots, \varphi_k$ определены однозначно и не зависят от выбора векторов $e'_1, e''_1, e'_2, e''_2, \ldots$, если по построению эти векторы не определены однозначно; б) углы $\varphi_1, \varphi_2, \ldots, \varphi_k$ определяют подпространства R' и R'' с точностью до положения в пространстве; иными словами, если имеются две пары подпространств R', R'' и S', S'' и углы между подпространствами R' и R'' такие же, как между подпространствами S' и S'', то существует изометрический оператор, переводящий одновременно S' в R' и S'' в R''; в) для любых наперед заданных углов $0 \leqslant \varphi_1 \leqslant \varphi_2 \leqslant \ldots \leqslant \varphi_k \leqslant \pi/2$ можно построить пару подпространств R' и R'', углы между которыми совпадают соответственно с числами $\varphi_1, \varphi_2, \ldots, \varphi_k$.

36. Пусть $y_1, y_2, \ldots, y_m$ — ортогональные проекции соответственно векторов $x_1, x_2, \ldots, x_m$ на некоторое подпространство. Доказать, что объем параллелепипеда, построенного на векторах $y_1, y_2, \ldots, y_m$, не превосходит объема параллелепипеда, построенного на векторах $x_1, x_2, \ldots, x_m$.

37. (Продолжение.) Предположим, что в задаче 36 как векторы $x_1, x_2, \ldots, x_m$, так и векторы $y_1, y_2, \ldots, y_m$ линейно независимы. Показать, что справедлива формула

$$V[y_1, y_2, \ldots, y_m] = V[x_1, x_2, \ldots, x_m] \cos\alpha_1 \cos\alpha_2 \ldots \cos\alpha_m,$$

где $\alpha_1, \alpha_2, \ldots, \alpha_m$ — углы между подпространствами

$$\mathrm{L}\{x_1, x_2, \ldots, x_m\} = \mathrm{L}_1 \text{ и } \mathrm{L}\{y_1, y_2, \ldots, y_m\} = \mathrm{L}_2$$

(задача 35).

38. Будем называть *k-вектором* совокупность k векторов евклидова пространства R. Два k-вектора $\{x_1, x_2, \ldots, x_k\}$ и $\{y_1, y_2, \ldots, y_k\}$ называются *равными*, если: 1) объем $V\{x_1, x_2, \ldots, x_k\}$ равен объему $V\{y_1, y_2, \ldots, y_k\}$; 2) линейная оболочка $\mathrm{L}(x_1, x_2, \ldots, x_k)$ совпадает с линейной оболочкой $\mathrm{L}(y_1, y_2, \ldots, y_k)$; 3) системы $x_1, x_2, \ldots, x_k$ и $y_1, y_2, \ldots, y_k$ имеют одинаковую ориентацию (т. е. оператор в пространстве $\mathrm{L}(x_1, x_2, \ldots, x_k)$, переводящий систему $\{x_1, x_2, \ldots, x_k\}$ в систему $\{y_1, y_2, \ldots, y_k\}$ имеет положительный определитель).

Показать, что k-вектор $\{x_1, x_2, \ldots, x_k\}$ в n-мерном пространстве R_n определяется однозначно, если известны величины всех миноров k-го порядка матрицы $\|\xi_i^{(j)}\|$ $(i=1, 2, \ldots, n;\ j=1, 2, \ldots, k)$, составленной из координат векторов $x_1, x_2, \ldots, x_k$ в произвольном ортогональном и нормированном базисе $e_1, e_2, \ldots, e_n$ пространства R_n.

39. Если k-вектор $\{x_1, x_2, \ldots, x_k\}$ равен k-вектору $\{y_1, y_2, \ldots, y_k\}$ (см. задачу 38), то миноры k-го порядка матрицы из координат векторов $x_1, x_2, \ldots, x_k$ равны соответствующим минорам матрицы из координат векторов $y_1, y_2, \ldots, y_k$.

40. Назовем *углами* между двумя k-векторами $\{x_1, x_2, \ldots, x_k\}$ и $\{y_1, y_2, \ldots, y_k\}$ набор углов между данными подпространствами $L_1 = L\{x_1, x_2, \ldots, x_k\}$ и $L_2 = L\{y_1, y_2, \ldots, y_k\}$ (задача 35), выбираемых, однако, с тем дополнительным условием, чтобы векторы $e_1, e_2, \ldots, e_k$, взятые в подпространстве L_1 для построения этих углов, имели бы ту же ориентацию, что и векторы $x_1, x_2, \ldots, x_k$ (это условие играет роль лишь при построении последнего вектора e_k); аналогично и в подпространстве L_2.

Показать, что углы $\beta_1, \beta_2, \ldots, \beta_k$ между k-векторами и углы $\alpha_1, \alpha_2, \ldots, \alpha_k$ между соответствующими подпространствами связаны следующими соотношениями:

$$\alpha_j = \beta_j \quad (j < k), \qquad \alpha_k = \beta_k \quad \text{или} \quad \alpha_k = \pi - \beta_k.$$

41. Назовем *скалярным произведением* двух k-векторов $X = \{x_1, x_2, \ldots, x_k\}$ и $Y = \{y_1, y_2, \ldots, y_k\}$, заданных матрицами X и Y из координат векторов x_i, y_i в некотором ортогональном и нормированном базисе пространства R_n, сумму всех произведений миноров k-го порядка матрицы X на соответствующие миноры матрицы Y. Показать, что это скалярное произведение равно

$$V[x_1, x_2, \ldots, x_k]\, V[y_1, y_2, \ldots, y_k] \cos\beta_1 \cdot \cos\beta_2 \ldots \cos\beta_k,$$

где $\beta_1, \beta_2, \ldots, \beta_k$ — углы между k-векторами X и Y.

42. Показать, что скалярное произведение двух k-векторов $X = (x_1, \ldots, x_k)$ и $Y = (y_1, \ldots, y_k)$ может быть записано в форме

$$\{X, Y\} = \begin{vmatrix} (x_1, y_1) & (x_1, y_2) & \ldots & (x_1, y_k) \\ (x_2, y_1) & (x_2, y_2) & \ldots & (x_2, y_k) \\ \ldots & \ldots & \ldots & \ldots \\ (x_k, y_1) & (x_k, y_2) & \ldots & (x_k, y_k) \end{vmatrix}.$$

43. Если многочлен $[P(t)]^k$ является аннулирующим для симметричного оператора A, то многочлен $P(t)$ также является аннулирующим для этого оператора.

КОМПЛЕКСНЫЕ ПРОСТРАНСТВА СО СКАЛЯРНЫМ ПРОИЗВЕДЕНИЕМ

§ 9.1. Эрмитовы формы

9.11. Числовая функция $A(x, y)$ от двух аргументов x, y в комплексном пространстве C называется *эрмитово-билинейной*, или, короче, *эрмитовой формой*, если она является линейной формой 1-го рода от x при каждом фиксированном значении y и линейной формой 2-го рода от y (*4.14*) при каждом фиксированном значении x.

Иными словами, $A(x, y)$ есть эрмитова форма от x и y, если для любых x, y, z из C и любых комплексных α удовлетворяются следующие условия:

$$\left.\begin{aligned} A(x+z,\ y) &= A(x,\ y)+A(z, y),\\ A(\alpha x,\ y) &= \alpha A(x, y),\\ A(x,\ y+z) &= A(x, y)+A(x,\ z),\\ A(x,\ \alpha y) &= \bar{\alpha} A(x,\ y). \end{aligned}\right\} \qquad (1)$$

Из этого определения легко получается и общая формула

$$A\left(\sum_{i=1}^{k} \alpha_i x_i,\ \sum_{j=1}^{m} \beta_j y_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{m} \alpha_i \bar{\beta}_j A(x_i, y_j), \qquad (2)$$

где $x_1, \ldots, x_k, y_1, \ldots, y_m$ — векторы пространства C, $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m$ — любые комплексные числа.

9.12. Примеры.

а. Если $f_1(x)$ — линейная форма 1-го рода, а $f_2(x)$ — линейная форма 2-го рода (*4.14*), то $A(x, y) = f_1(x)\ f_2(y)$ — эрмитова форма.

б. В n-мерном пространстве C_n с базисом $e_1, \ldots, e_n$ функция от векторов $x=\sum_1^n \xi_k e_k$ и $y=\sum_1^n \eta_k e_k$

$$A(x, y)=\sum_{i, j=1}^{n} a_{ij}\xi_i\bar{\eta}_j \tag{3}$$

является эрмитовой формой при любых комплексных a_{ij} $(i, j=1, \ldots, n)$.

В действительности выражение (3) дает общий вид эрмитовой формы в n-мерном комплексном пространстве. Это доказывается так же, как в *7.13* доказывалось аналогичное предложение для билинейных форм в пространстве K_n.

9.13. Эрмитова форма $A(x, y)$ называется *симметричной эрмитовой формой*, если для любых векторов x и y

$$A(y, x)=\overline{A(x, y)}. \tag{4}$$

Если эрмитова форма $A(x, y)$ в комплексном пространстве C_n записана через координаты формулой (3) и симметрична, то

$$a_{ik}=A(e_i, e_k)=\overline{A(e_k, e_i)}=\bar{a}_{ki}, \tag{5}$$

т. е. матрица $\|a_{ik}\|$ формы A в базисе $e_1, \ldots, e_n$ после транспонирования и замены элементов на комплексно сопряженные переходит в себя. Обратно, если в некотором базисе $e_1, \ldots, e_n$ коэффициенты эрмитовой формы $A(x, y)$ удовлетворяют условиям (5), то форма $A(x, y)$ симметрична. В самом деле,

$$A(y, x)=\sum_{i, k=1}^{n} a_{ik}\eta_i\bar{\xi}_k=$$
$$=\sum_{i, k=1}^{n} \bar{a}_{ki}\bar{\xi}_k\eta_i=\overline{\sum_{i, k=1}^{n} a_{ki}\xi_k\bar{\eta}_i}=\overline{A(x, y)}.$$

Матрицу $\|a_{ik}\|$, для которой $a_{ik}=\bar{a}_{ki}$ $(i, k=1, \ldots, n)$, будем называть в дальнейшем *эрмитово-симметричной*.

9.14а. Пусть эрмитова форма $A(x, y)$ в базисе $e_1, \ldots, e_n$ пространства C_n имеет матрицу $A_{(e)}=\|a_{ik}\|$, а в базисе $f_1, \ldots, f_n$ — матрицу $A_{(f)}=\|b_{ik}\|$, причем векторы f_j и e_j

связаны соотношениями

$$f_i = \sum_{j=1}^{n} p_j^{(i)} e_j \qquad (i = 1, \ldots, n).$$

Рассуждая так же, как в *7.15а*, приходим к следующему результату: между матрицами $A_{(f)}$ и $A_{(e)}$ имеет место соотношение

$$A_{(f)} = P^* A_{(e)} P, \tag{6}$$

где $P = \| p_j^{(i)} \|$ — матрица перехода от базиса $e_1, \ldots, e_n$ к базису $f_1, \ldots, f_n$, а P^* — матрица, получающаяся из матрицы P транспонированием и переходом к комплексно сопряженным величинам. Обозначая $P^* = \| \overset{*}{p}_j^{(i)} \|$, имеем

$$\overset{*}{p}_j^{(i)} = \overline{p_i^{(j)}} \qquad (i, j = 1, \ldots, n).$$

б. Из равенства (6), как и в *7.23*, вытекает, что ранг матрицы $A_{(e)}$ эрмитовой формы $\mathrm{A}(x, y)$ не зависит от выбора базиса $\{e\}$. Форма $\mathrm{A}(x, y)$ называется *невырожденной*, если ее *ранг* (т. е. ранг матрицы $A_{(e)}$ в любом базисе $\{e\}$) равен числу n — размерности пространства C_n. Если форма $\mathrm{A}(x, y)$ невырождена, то для любого $x_0 \neq 0$ найдется вектор $y_0 \in C_n$ такой, что $\mathrm{A}(x_0, y_0) \neq 0$ (ср. *7.15в*).

9.15а. *Эрмитово-квадратичной формой* в комплексном пространстве C называется функция от одного аргумента $x \in C$, которая получается из эрмитово-билинейной формы $\mathrm{A}(x, y)$ заменой y на x.

В n-мерном комплексном пространстве C_n с базисом $e_1, \ldots, e_n$ эрмитово-квадратичная форма через координаты $\xi_1, \ldots, \xi_n$ вектора x, в силу *9.12б*, записывается в виде

$$\mathrm{A}(x, x) = \sum_{i,k=1}^{n} a_{ik} \xi_i \bar{\xi}_k \tag{7}$$

с некоторыми комплексными коэффициентами a_{ik}. И обратно, функция $\mathrm{A}(x, x)$ вида (7) есть эрмитово-квадратичная форма, которая получается заменой y на x в эрмитово-билинейной форме

$$\mathrm{A}(x, y) = \sum_{i,k=1}^{n} a_{ik} \xi_i \bar{\eta}_k.$$

б. Если эрмитово-билинейная форма $A(x, y)$ симметрична, так что $a_{ik} = \bar{a}_{ki}$, то соответствующая эрмитово-квадратичная форма также называется *симметричной.*

Симметричная эрмитово-квадратичная форма $A(x, x)$ принимает лишь вещественные значения, поскольку из (4) следует, что

$$\overline{A(x, x)} = A(x, x).$$

В отличие от случая *7.22*, эрмитово-билинейная форма уже определяется однозначно по соответствующей ей эрмитово-квадратичной форме.

Действительно, мы имеем

$$A(x+y,\ x+y) = A(x, x) + A(x, y) + \overline{A(x, y)} + A(y, y),$$
$$A(x+iy,\ x+iy) = A(x, x) - iA(x, y) + i\overline{A(x, y)} + A(y, y).$$

Из первого уравнения мы можем найти $A(x, y) + \overline{A(x, y)} = 2\operatorname{Re} A(x, y)$ через значения квадратичных форм $A(x, x)$, $A(y, y)$ и $A(x+y, x+y)$. Из второго уравнения аналогично можно найти $-iA(x, y) + i\overline{A(x, y)} = 2\operatorname{Im} A(x, y)$. Это доказывает наше утверждение.

Если в некотором базисе $e_1, \ldots, e_n$ квадратичная форма $A(x, x)$ записана в виде

$$A(x, x) = \sum_{j,\ k=1}^{n} a_{jk} \xi_j \bar{\xi}_k,$$

то эрмитово-билинейная форма

$$A(x, y) = \sum_{j,\ k=1}^{n} a_{jk} \xi_j \bar{\eta}_k,$$

очевидно, приводится к форме $A(x, x)$ при подстановке $y = x$. По доказанному, выписанная форма $A(x, y)$ является единственной эрмитово-билинейной симметричной формой, приводящей к данной форме $A(x, x)$ при подстановке $y = x$.

9.16а. *В n-мерном пространстве C_n существует базис, в котором эрмитово-квадратичная симметричная форма запи-*

сывается в каноническом виде

$$A(x, x) = \sum_{k=1}^{n} \lambda_k \eta_k \bar{\eta}_k \equiv \sum_{k=1}^{n} \lambda_k |\eta_k|^2 \tag{8}$$

с вещественными коэффициентами $\lambda_1, \ldots, \lambda_n$.

Доказательство проводится аналогично доказательству теоремы *7.31*. Вместо равенства *7.31* (12) используется равенство ($b_{mm} \neq 0$)

$$\begin{aligned} b_{1m}\xi_1\bar{\xi}_m + \ldots + b_{m-1,m}\xi_{m-1}\bar{\xi}_m + b_{mm}\xi_m\bar{\xi}_m + \\ + \bar{b}_{1m}\bar{\xi}_1\xi_m + \ldots + \bar{b}_{m-1,m}\bar{\xi}_{m-1}\xi_m = \\ = b_{mm}\left|\frac{b_{1m}}{b_{mm}}\xi_1 + \ldots + \frac{b_{m-1,m}}{b_{mm}}\xi_{m-1} + \xi_m\right|^2 + A_1(x, x), \end{aligned}$$

где $A_1(x, x)$ — эрмитово-квадратичная симметричная форма от аргументов $\xi_1, \ldots, \xi_{m-1}$. Вместо преобразования *7.31* (14) используется преобразование

$$\begin{aligned} \xi_1 &= \xi_1' + \xi_2', \\ \xi_2 &= \xi_1' + i\xi_2', \\ \xi_3 &= \xi_3', \\ &\cdots\cdots \\ \xi_m &= \xi_m'. \end{aligned}$$

Сумму $a_{12}\xi_1\bar{\xi}_2 + \bar{a}_{12}\bar{\xi}_1\xi_2$ ($a_{12} \neq 0$) оно приводит к виду

$$(a_{12} + \bar{a}_{12})\xi_1'\bar{\xi}_1' - i(a_{12} - \bar{a}_{12})\xi_2'\bar{\xi}_2' + \ldots,$$

причем хотя бы один из двух коэффициентов $a_{12} + \bar{a}_{12}$ и $i(a_{12} - \bar{a}_{12})$ заведомо отличен от 0.

б. Теорема инерции *7.91* сохраняется (для эрмитово-квадратичных симметричных форм) и в комплексной области: *число* p *положительных и число* q *отрицательных коэффициентов в наборе* $\lambda_1, \ldots, \lambda_n$ (8) *не зависит от выбора канонического базиса формы* $A(x, x)$. Эти числа p и q называются *индексами инерции* формы $A(x, x)$; первый — *положительным индексом инерции*, второй — *отрицательным*.

Доказательство полностью повторяет доказательство теоремы *7.91*.

Заметим, что для квадратичных (не эрмитово-квадратичных) форм в комплексном пространстве C_n закон инерции не имеет места. Например, квадратичная форма

$$A(x, x) = \xi_1^2 + \xi_2^2$$

преобразованием координат

$$\xi_1 = \eta_1, \quad \xi_2 = i\eta_2$$

приводится к виду

$$A(x, x) = \eta_1^2 - \eta_2^2.$$

в. Для данной эрмитово-квадратичной симметричной формы $A(x, x)$ в пространстве C_n можно всегда найти канонический базис так, что соответствующие канонические коэффициенты будут равны $+1$ или -1. Для этого, приведя форму $A(x, x)$ к виду

$$A(x, x) = \lambda_1 |\eta_1|^2 + \ldots + \lambda_p |\eta_p|^2 - \mu_1 |\eta_{p+1}|^2 - \ldots \\ \ldots - \mu_q |\eta_{p+q}|^2,$$

где $\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q$ положительны, совершаем дальнейшее преобразование координат по формулам

$$\sqrt{\lambda_1}\eta_1 = \tau_1, \ \ldots, \qquad \sqrt{\lambda_p}\eta_p = \tau_p,$$
$$\sqrt{\mu_1}\eta_{p+1} = \tau_{p+1}, \ \ldots, \quad \sqrt{\mu_q}\eta_{p+q} = \tau_{p+q},$$

после чего форма $A(x, x)$ принимает вид

$$A(x, x) = |\tau_1|^2 + \ldots + |\tau_p|^2 - |\tau_{p+1}|^2 - \ldots - |\tau_{p+q}|^2.$$

9.17a. Вектор x_1 называется *сопряженным к вектору* y_1 *относительно эрмитово-билинейной формы* $A(x, y)$, *если* $A(x_1, y_1) = 0$.

Если векторы $x_1, \ldots, x_k$ сопряжены с вектором y_1, то также сопряжен с вектором y_1 любой вектор подпространства $L(x_1, \ldots, x_k)$ — линейной оболочки векторов $x_1, \ldots, x_k$. Вообще, если вектор y сопряжен с каждым вектором некоторого подпространства $C' \subset C$, мы будем называть этот вектор *сопряженным к подпространству* C'. Совокупность C'' всех векторов $x \in C$, сопряженных к подпространству C', очевидно, и сама является подпространством в C; это подпространство C'' мы будем называть *сопряженным к* C'.

Базис $e_1, \ldots, e_n$ пространства C_n называется *каноническим базисом* формы $A(x, y)$, если $A(e_i, e_j) = 0$ при $i \neq j$. *Всякая симметричная эрмитово-билинейная форма* $A(x, y)$

имеет канонический базис: это базис, в котором соответствующая квадратичная форма $A(x, x)$ записывается в каноническом виде (8). Действительно, по *9.15б* в этом базисе форма $A(x, y)$ при $x = \sum_{k=1}^{n} \eta_k e_k$, $y = \sum_{k=1}^{n} \tau_k e_k$ принимает вид

$$A(x, y) = \sum_{k=1}^{n} \lambda_k \eta_k \bar{\tau}_k,$$

так что

$$A(e_i, e_j) = \begin{cases} \lambda_i \text{ при } j = i, \\ 0 \text{ при } j \neq i. \end{cases}$$

б. Если угловые миноры $\delta_1, \ldots, \delta_{n-1}$ матрицы $\|a_{jk}\|$ эрмитово-квадратичной симметричной формы $A(x, x)$ отличны от нуля, то канонический базис формы $A(x, x)$ может быть построен по методу Якоби, как в *7.52*. При этом сохраняются и формулы (24) для канонических коэффициентов формы $A(x, x)$:

$$\lambda_1 = \delta_1, \quad \lambda_2 = \frac{\delta_2}{\delta_1}, \quad \ldots, \quad \lambda_n = \frac{\delta_n}{\delta_{n-1}}.$$

в. Эрмитово-билинейная симметричная форма $A(x, y)$ называется *положительно определенной*, если при любом $x \neq 0$

$$A(x, x) > 0.$$

Так же как и в вещественном случае (*7.94*), эквивалентным условием является положительность всех канонических коэффициентов формы $A(x, x)$, или, что то же, равенство $p = n$, где p — положительный индекс инерции формы $A(x, x)$.

Необходимым и достаточным условием положительной определенности формы $A(x, y)$, как и в *7.96*, является выполнение условий Сильвестра

$$\delta_1 > 0, \quad \delta_2 > 0, \quad \ldots, \quad \delta_n > 0.$$

Доказательство, приведенное в *7.96*, проходит в комплексном случае без изменений.

9.18а. Имея эрмитово-билинейную невырожденную симметричную форму (x, y), можно, аналогично тому как это было сделано в *7.61*, ввести понятия сопряженных операторов (относительно формы (x, y)). Прежде всего, если A и B — линейные операторы в пространстве C_n, то функции $A(x, y) = (Ax, y)$, $B(x, y) = (x, By)$ являются эрмитово-билинейными формами, матрицы которых связаны с матрицами операторов A и B (в любом каноническом базисе формы (x, y) с каноническими коэффициентами ε_j) формулами

$$a_{jm} = \varepsilon_m a_m^{(j)}, \quad b_{jm} = \varepsilon_j \overline{b_j^{(m)}}.$$

Обратно, по заданным эрмитово-билинейным формам $A(x, y)$ и $B(x, y)$, как в *7.62*, можно указать, причем единственным образом, операторы A и B так, что

$$A(x,y) = (Ax, y), \quad B(x, y) = (x, By).$$

б. Отсюда, как в *7.63*, следует существование для любого оператора A такого оператора A*, что при любых x и y из C_n

$$(Ax, y) = (x, A^*y).$$

Оператор A* определен единственным образом; в каноническом базисе формы (x, y) с каноническими коэффициентами ε_j матрицы $\|a_m^{(j)}\|$ и $\|a_m^{(j)*}\|$ операторов A и A* связаны соотношениями

$$a_j^{*m} = \frac{\varepsilon_m}{\varepsilon_j} \overline{a_m^{(j)}}.$$

Оператор A* называется *эрмитово-сопряженным к оператору* A *относительно формы* (x, y).

в. Имеют место следующие формулы (ср. *7.64*):

а) $(A^*)^* = A$ для любого оператора A;

б) $(A + B)^* = A^* + B^*$ для любых A и B;

в) $(\lambda A)^* = \bar{\lambda} A^*$ для любого оператора A и любого $\lambda \in C$;

г) $(AB)^* = B^*A^*$.

9.19а. Назовем два комплексных пространства C' и C'' с выделенными в них эрмитово-билинейными симметричными формами $A(x', y')$ и $A(x'', y'')$ *A-изоморфными*, если пространства C' и C'' изоморфны как комплексные пространства (*2.71*) и если для соответствующих друг другу пар элементов x', y' из C' и x'', y'' из C'' справедливо равенство

$$A(x', y') = A(x'', y'').$$

б. Теорема. *Два конечномерных комплексных пространства* C′ *и* C″ *с выделенными в них эрмитово-билинейными симметричными формами* $A(x', y')$ *и* $A(x'', y'')$ *являются* A-*изоморфными тогда и только тогда, когда размерности* C′ *и* C″ *одинаковы и индексы инерции* p' *и* q' *формы* $A(x', y')$ *совпадают с индексами инерции* p'' *и* q'' *формы* $A(x'', y'')$.

Доказательство проводится точно так же, как доказательство аналогичного предложения для вещественных пространств (*7.93*).

в. В частности, два пространства C'_n и C''_n одинаковой размерности n с выделенными в них положительно определенными формами $A(x', y')$ и $A(x'', y'')$ всегда A-изоморфны.

§ 9.2. Скалярное произведение в комплексном пространстве

9.21. В вещественном пространстве в качестве скалярного произведения двух векторов мы брали фиксированную симметричную положительно определенную билинейную форму. Соответствующая квадратичная форма положительна на каждом ненулевом векторе и позволяет тем самым определить его длину. В комплексном пространстве аналогичным свойством обладает положительно определенная эрмитово-билинейная форма (*9.17в*). В связи с этим мы принимаем следующее определение:

Комплексное линейное пространство C называется *унитарным пространством*, если в нем фиксирована некоторая положительно определенная эрмитово-билинейная форма, называемая (комплексным) *скалярным произведением*; иными словами, если каждой паре векторов x, y из C поставлено в соответствие комплексное число (x, y), удовлетворяющее условиям:

а) $(y, x) = \overline{(x, y)}$ для любых x, y из C;

б) $(x, y+z) = (x, y) + (x, z)$ для любых x, y, z из C;

в) $(\lambda x, y) = \lambda(x, y)$ для любых x, y из C и любого комплексного числа λ;

г) $(x, x) > 0$ для любого $x \neq 0$; $(0, 0) = 0$.

Из аксиом а) — в) следует общая формула

$$\left(\sum_{j=1}^{p} \alpha_j x_j, \sum_{k=1}^{q} \beta_k y_k\right) = \sum_{j=1}^{p} \sum_{k=1}^{q} \alpha_j \bar{\beta}_k (x_j, y_k)$$

для любых $x_1, \ldots, x_p, y_1, \ldots, y_q$ из C и любых комплексных $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$.

9.22. Примеры.

а. В n-мерном пространстве C_n (*2.15б*) введем скалярное произведение векторов $x=(\xi_1, \ldots, \xi_n)$, $y=(\eta_1, \ldots, \eta_n)$ по формуле

$$(x, y)=\xi_1\overline{\eta}_1+\ldots+\xi_n\overline{\eta}_n.$$

Выполнение свойств а) — г) легко проверяется.

б. В пространстве $C(a, b)$ комплексных непрерывных функций на отрезке $[a, b]$ (*2.15г*) скалярное произведение функций $x=x(t)$ и $y=y(t)$ введем по формуле

$$(x, y)=\int_a^b x(t)\overline{y(t)}\,dt.$$

Выполнение аксиом а) — г) следует из основных свойств интеграла.

9.23. Основные метрические понятия. В унитарном пространстве C можно ввести некоторые метрические понятия аналогично тому, как это делалось в вещественном евклидовом пространстве (§ 8.3).

а. Длина вектора. Как и в вещественном случае, длиной (или *нормой*) вектора x называют величину

$$|x|=+\sqrt{(x, x)}.$$

У всякого ненулевого вектора длина положительна, длина нулевого вектора равна 0. При любом комплексном α имеет место равенство

$$|\alpha x|=\sqrt{(\alpha x, \alpha x)}=\sqrt{\alpha\overline{\alpha}(x, x)}=|\alpha|\sqrt{(x, x)}=|\alpha|\,|x|,$$

показывающее, что модуль числового множителя можно выносить за знак длины вектора.

Вектор x длины 1 называется *нормированным*.

Каждый вектор x можно *нормировать* — разделив его на его длину, получить вектор того же направления (т. е. лежащий в том же одномерном подпространстве) и единичной длины. Совокупность всех $x \in C$ с $|x| \leqslant 1$ называется *единичным шаром* пространства C.

б. Неравенство Коши — Буняковского. Для любых двух векторов x, y из C имеет место неравенство

$$|(x, y)| \leqslant |x|\,|y|. \tag{9}$$

Доказательство проводится по той же схеме, что и в вещественном случае *(8.33)*, но с некоторой осторожностью обращения с комплексными числами. Если $(x, y) = 0$, неравенство (9) очевидно. При $(x, y) \neq 0$ замечаем, что

$$(\lambda x - y, \lambda x - y) \geqslant 0$$

при любом комплексном λ. Раскрывая скобки, находим

$$|\lambda|^2 (x, x) - \lambda (x, y) - \bar{\lambda}\,\overline{(x, y)} + (y, y) \geqslant 0. \tag{10}$$

Будем считать, что λ изменяется по прямой γ, симметричной относительно вещественной оси с прямой, определяемой комплексным числом (x, y), так что $\lambda = tz_0$, где t вещественно, а z_0 — единичный вектор, определяющий направление прямой γ, $z_0 = \dfrac{\overline{(x, y)}}{|(x, y)|}$. Тогда $\lambda (x, y) = t\,|(x, y)|$ есть вещественное число, так что $\bar{\lambda}\,\overline{(x, y)} = \lambda (x, y)$. Неравенство (10) преобразуется к виду

$$t^2 (x, x) - 2t\,|(x, y)| + (y, y) \geqslant 0. \tag{11}$$

Теперь та же аргументация, что и в *8.33*, приводит нас к искомому неравенству (9).

Если в неравенстве (9) стоит знак равенства, то трехчлен в левой части (11) имеет один вещественный корень t_0. Заменяя tz_0 на λ, мы получаем, что трехчлен в левой части (10) имеет корень $\lambda_0 = t_0 z_0$, откуда

$$(\lambda_0 x - y, \lambda_0 x - y) = 0$$

и $y = \lambda_0 x$, так что векторы x и y отличаются лишь (комплексным) множителем.

в. Ортогональность. В унитарном пространстве не вводят понятия угла между векторами. Рассматривают лишь случай, когда векторы x и y ортогональны; под этим, как и в вещественном случае, понимают выполнение равенства

$$(x, y) = 0.$$

При этом, очевидно, и $(y, x) = \overline{(x, y)} = 0$.

Для ортогональных векторов, как легко проверить, остаются справедливыми аналоги лемм *8.36а—б* и теорема Пифагора *8.37*.

Далее, справедлива теорема о разложении *8.51*: для конечномерного подпространства $C' \subset C$ и любого вектора $f \in C$ существует (и единственно) разложение

$$f = g + h,$$

где $g \in C'$, h ортогонален к C'. Совокупность всех векторов h, ортогональных к подпространству C', снова образует подпространство, которое называется *ортогональным дополнением к подпространству* C'. Обозначая его через C'', выводим, как в *8.51*, существование разложения $C = C' + C''$ в прямую сумму ортогональных слагаемых.

г. Неравенства треугольника. Если x и y — два вектора в унитарном пространстве C, то по неравенству Коши — Буняковского (*б*)

$$|x+y|^2 = (x+y, x+y) = (x, x) + (x, y) + \overline{(x, y)} + (y, y)$$
$$\begin{cases} \leqslant (x, x) + 2\,|(x, y)| + (y, y) \leqslant (|x| + |y|)^2, \\ \geqslant (x, x) - 2\,|(x, y)| + (y, y) \geqslant (|x| - |y|)^2, \end{cases}$$

откуда

$$|x+y| \begin{cases} \leqslant |x| + |y|, \\ \geqslant \big||x| - |y|\big|. \end{cases} \tag{12}$$

Неравенства (12), как и в вещественном случае, называются *неравенствами треугольника*.

9.24. Ортогональный базис в n-мерном унитарном пространстве C_n. В n-мерном пространстве (симметричная) эрмитово-билинейная форма (x, y) обладает каноническим базисом $e_1, \ldots, e_n$ (*9.16а*). Условие каноничности $(e_i, e_j) = 0$ $(i \neq j)$ в данном случае есть условие ортогональности. Ортогональные базисные векторы $e_1, \ldots, e_n$ можно далее считать нормированными, так что

$|e_1| = \ldots = |e_n| = 1$. Если при этом $x = \sum_1^n \xi_k e_k$, $y = \sum_1^n \eta_k e_k$ — любые два вектора из C_n, мы получаем формулу для скалярного произведения

$$(x, y) = \left(\sum_1^n \xi_k e_k, \ \sum_1^n \eta_k e_k\right) = \sum_1^n \xi_k \overline{\eta}_k. \qquad (13)$$

9.25а. В соответствии с *9.18а* устанавливается взаимно однозначное соответствие между эрмитово-билинейными формами $A(x, y)$ и линейными операторами A, действующими в пространстве C_n, по формуле

$$A(x, y) \equiv (Ax, y).$$

В любом ортонормированном базисе $e_1, \ldots, e_n$ пространства C_n матрица $\| a_{jk} \|$ формы $A(x, y)$ $(a_{jk} = A(e_j, e_k))$ и матрица $\| a_j^{(k)} \|$ оператора A $(Ae_k = \sum_j a_j^{(k)} e_j)$ связаны соотношениями

$$a_{jk} = a_j^{(k)}.$$

б. В соответствии с *9.18б* вводится понятие сопряженности операторов A и A^* относительно скалярного произведения (x, y). Именно, для любого линейного оператора A, действующего в пространстве C_n, существует и единствен *эрмитово-сопряженный* оператор — оператор A^*, удовлетворяющий уравнению

$$(Ax, y) = (x, A^*y)$$

для любых x и y из C_n.

Поскольку ортогональный и нормированный базис есть канонический базис формы (x, y) с каноническими коэффициентами $\varepsilon_j = 1$, в любом ортогональном и нормированном базисе пространства C_n матрицы $\| a_m^{(j)} \|$ и $\| a_m^{(j)*} \|$ операторов A и A^* связаны соотношениями

$$a_j^{*(m)} = \overline{a_m^{(j)}}.$$

в. Как и в *8.95а, ортогональное дополнение C'' к подпространству $C' \subset C$, инвариантному относительно оператора A, инвариантно относительно сопряженного оператора A^*.*

9.26. Линейное преобразование в n-мерном комплексном пространстве, соответствующее переходу от одного ортонормированного базиса к другому такому же, называется *унитарным* преобразованием. Унитарные преобразования аналогичны ортогональным преобразованиям вещественного пространства (*8.93*). Если $e_1, \ldots, e_n$ и $f_1, \ldots, f_n$ — ортонормированные базисы и $U = \|u_k^{(i)}\|$ есть матрица соответствующего унитарного преобразования, так что

$$f_i = \sum_{k=1}^{n} u_k^{(i)} e_k,$$

то, очевидно,

$$(f_i, f_j) = \sum_{k=1}^{n} u_k^{(i)} \overline{u_k^{(j)}} = \begin{cases} 0 & \text{при } i \neq j, \\ 1 & \text{при } i = j. \end{cases} \tag{14}$$

Обратно, если числа $u_k^{(i)}$ удовлетворяют соотношениям (14), то матрица $\|u_k^{(i)}\|$ есть матрица унитарного преобразования, или, короче, *унитарная матрица.*

Линейный оператор U, соответствующий унитарной матрице, называется *унитарным оператором*. Так же как изометрический оператор в вещественном пространстве, унитарный оператор в комплексном пространстве не изменяет метрики: если $x = \sum_{i=1}^{n} \xi_i e_i$, $y = \sum_{j=1}^{n} \eta_j e_j$, то

$$(\mathrm{U}x, \mathrm{U}y) = \sum_{i, j=1}^{n} \xi_i \bar{\eta}_j (\mathrm{U}e_i, \mathrm{U}e_j) =$$

$$= \sum_{i, j=1}^{n} \xi_i \bar{\eta}_j (f_i, f_j) = \sum_{i=1}^{n} \xi_i \bar{\eta}_i = (x, y).$$

Матрица V обратного перехода от базиса $\{f_j\}$ к базису $\{e_i\}$ обратна матрице U и также унитарна; далее, если $V = \|v_k^{(i)}\|$, мы имеем

$$u_k^{(i)} = (f_i, e_k), \quad v_k^{(i)} = (e_i, f_k) = \overline{u_i^{(k)}};$$

таким образом, матрица, обратная к унитарной, получается путем транспонирования и перехода к комплексно сопряженным элементам. Итак, для унитарного оператора U

$$\mathrm{U}^{-1} = \mathrm{U}^*,$$

или

$$\mathrm{U}^*\mathrm{U} = \mathrm{U}\mathrm{U}^* = \mathrm{E}.$$

§ 9.3. Нормальные операторы

9.31. Определение. Оператор A, действующий в n-мерном унитарном пространстве C_n, называется *нормальным*, если он коммутирует со своим сопряженным:

$$A^*A = AA^*. \tag{15}$$

Примером служит оператор A, обладающий ортогональным базисом из собственных векторов $e_1, \ldots, e_n$, так что

$$Ae_j = \lambda_j e_j \qquad (j = 1, \ldots, n).$$

Действительно, матрица этого оператора A в базисе $e_1, \ldots, e_n$ имеет вид

$$\left\| \begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{matrix} \right\|. \tag{16}$$

Согласно *9.25б* матрица оператора A^* в том же базисе $e_1, \ldots, e_n$ имеет вид

$$\left\| \begin{matrix} \bar\lambda_1 & & \\ & \ddots & \\ & & \bar\lambda_n \end{matrix} \right\|, \tag{17}$$

отсюда очевидно, что операторы A и A^* коммутируют.

9.32. Теорема. *Каждый собственный вектор x_0 нормального оператора A с собственным значением λ_0 является собственным вектором оператора A^* с собственным значением $\bar\lambda_0$.*

Доказательство. Пусть $P \subset C_n$ — подпространство, состоящее из *всех* собственных векторов оператора A с собственным значением λ_0. Покажем, что P инвариантно относительно оператора A^*. Для $x \in P$ мы имеем

$$AA^*x = A^*Ax = A^*(\lambda x) = \lambda A^*x,$$

откуда следует, что $A^*x \in P$, что и требовалось.

Далее, для любых x и y из P мы имеем

$$(A^*x, y) = (x, Ay) = (x, \lambda y) = (\bar\lambda x, y),$$

откуда для любого $x \in \mathrm{P}$

$$\mathrm{A}^* x = \bar{\lambda} x.$$

Теорема доказана.

9.33а. Теорема. *Для всякого нормального оператора* **A** *существует ортогональный и нормированный базис* $e_1, \ldots, e_n$ *из собственных векторов оператора* A.

Доказательство. Нормальный оператор A, как и всякий линейный оператор в пространстве C_n, имеет собственный вектор (*4.95б*). Пусть e — собственный вектор оператора A с собственным значением λ. Пусть $\mathrm{P} \subset C_n$ — подпространство, состоящее из *всех* собственных векторов оператора A с этим собственным значением λ. Если P есть все пространство C_n, то, произвольно дополняя в C_n вектор e_1 векторами $e_2, \ldots, e_n$ до ортогонального и нормированного базиса, получаем, что теорема доказана. Если $\mathrm{P} \neq C_n$, то пусть Q — ортогональное дополнение подпространства P в C_n. Так как в силу *9.32* каждый вектор пространства P оператором A^* переводится снова в вектор пространства P (даже в себя самого с коэффициентом $\bar{\lambda}$), то подпространство P инвариантно относительно оператора A^*. В силу равенства $\mathrm{A}^{**} = \mathrm{A}$ (*9.18в*)) и теоремы *9.25в*, подпространство Q инвариантно относительно оператора A. Теперь воспользуемся принципом индукции, считая, что теорема справедлива для пространств меньшей размерности; тогда в подпространстве Q можно выбрать ортогональный базис, удовлетворяющий требуемому условию; присоединяя любой ортогональный базис подпространства P, мы получим полный ортогональный базис в пространстве C_n, удовлетворяющий условию теоремы.

б. В силу *а*, всякий нормальный оператор A оказывается диагонализируемым (*4.72е*); в базисе из его собственных векторов, построенном в *а*, этот оператор имеет диагональную матрицу

$$A = \begin{Vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{Vmatrix}.$$

На главной диагонали матрицы стоят собственные значения оператора A; каждое из них повторяется столько раз, какова размерность соответствующего собственного подпространства.

Поэтому характеристический многочлен $\det\|A-\lambda E\|$ оператора A (который, как мы знаем, не зависит от выбора базиса) имеет вид

$$\det\|A-\lambda E\| = \prod_{j=1}^{m} (\lambda_i-\lambda)^{k_j}, \quad \sum_{j=1}^{m} k_j = n, \tag{18}$$

где $\lambda_1, \ldots, \lambda_m$ — различные собственные значения оператора A, а $k_1, \ldots, k_m$ — размерности соответствующих собственных подпространств.

в. С другой стороны, пусть известно, что у некоторого нормального оператора A характеристический многочлен имеет вид

$$\det\|A-\lambda E\| = \prod_{i=1}^{s} (\mu_i-\lambda)^{p_i}, \tag{19}$$

где $\mu_1, \ldots, \mu_s$ — различные числа, $p_1, \ldots, p_s$ — некоторые кратности. Тогда можно утверждать, что *оператор* A *имеет ортонормальный базис из собственных векторов с собственными значениями* $\mu_1, \ldots, \mu_s$, причем размерность собственного подпространства, отвечающего значению μ_j, равна числу p_j. В самом деле, в силу единственности характеристического многочлена имеет место равенство многочленов (18) и (19), откуда, применяя теорему о единственности разложения многочлена на множители, получаем нужное.

9.34. Самосопряженные операторы. Если $A^* = A$, то оператор A называется *самосопряженным*. Иными словами, оператор A самосопряжен, если соответствующая оператору A билинейная форма (x, Ay) эрмитово-симметрична:

$$(Ax, y) = (x, Ay). \tag{20}$$

Оператор A, удовлетворяющий уравнению (20) для любых двух векторов x и y, называют поэтому также эрмитово-симметричным или *эрмитовым* оператором.

В силу теоремы *9.25б* матрица самосопряженного оператора в любом ортогональном и нормированном базисе совпадает со своей эрмитово-транспонированной матрицей, иными словами, есть эрмитово-симметричная матрица. И обратно, каждый оператор A, имеющий в некотором ортогональном и нормированном базисе эрмитово-симметричную матрицу, является самосопряженным оператором.

Так как самосопряженный оператор, очевидно, нормален, то можно применить *9.32*; мы получаем в данном случае, что $\overline{\lambda_0} = \lambda_0$, откуда следует, что *каждое собственное значение сопряженного оператора вещественно*. Далее, применяя *9.33а,* получаем следующую основную теорему:

Теорема. *Для всякого самосопряженного оператора* A *в унитарном пространстве* C_n *существует ортонормальный базис из собственных векторов оператора* A *с вещественными собственными значениями.*

Обратно, *всякий линейный оператор* A *в пространстве* C_n, *обладающий указанным свойством, является самосопряженным*: действительно, по *9.31* он нормален, и, сравнивая (16) с (17), в силу вещественности чисел λ_j заключаем, что $A^* = A$.

9.35. Антисамосопряженные операторы. Если $A^* = -A$, то оператор A называется *антисамосопряженным*. Матрица антисамосопряженного оператора в любом ортонормальном базисе $e_1, \ldots, e_n$ обладает характеристическим признаком:

$$a_{jk} = (Ae_j, e_k) = (e_j, A^* e_k) = (e_j, -Ae_k) = -\overline{(Ae_k, e_j)} = -\overline{a}_{kj}$$
$$(j, k = 1, \ldots, n).$$

Антисамосопряженный оператор A, очевидно, нормален. Применяя *9.32*, мы получаем, что в данном случае $\overline{\lambda_0} = -\lambda_0$, откуда следует, что *каждое собственное значение антисамосопряженного оператора чисто мнимо*. Далее, применяя *9.33а*, получаем следующую основную теорему:

Теорема. *Для всякого антисамосопряженного оператора* A *в унитарном пространстве* C_n *существует ортонормальный базис из собственных векторов оператора* A *с чисто мнимыми собственными значениями.*

Обратно, *всякий линейный оператор* A *в пространстве* C_n, *обладающий указанным свойством, является антисамосопряженным.*

9.36. Унитарные операторы. Оператор U, действующий в пространстве C_n, называется *унитарным*, если $U^*U = UU^* = E$ (*9.26*). В частности, *унитарный оператор нормален.* Применяя *9.32*, находим, что в данном случае $\overline{\lambda}_0 \cdot \lambda_0 = 1$, или, что то же, $|\lambda_0| = 1$; таким образом, *каждое собственное значение унитарного оператора по модулю равно* 1. Далее, применяя *9.33а,* получаем следующую теорему:

Теорема. *Для всякого унитарного оператора* U *в пространстве* C_n *существует ортонормальный базис из собственных векторов оператора* U *с собственными значениями, по модулю равными* 1.

Обратно, *всякий линейный оператор* U *в пространстве* C_n, *обладающий указанным свойством, унитарен.*

§ 9.4. Применение унитарного пространства к теории операторов в евклидовом пространстве

9.41. Включение вещественного евклидова пространства в унитарное пространство. Пусть R — вещественное евклидово пространство (*8.21*) со скалярным произведением (x, y). Рассмотрим комплексное пространство C, составленное из формальных сумм $x + iy$, где $x \in R$, $y \in R$, с естественными операциями сложения и умножения на произвольные комплексные числа:

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2);$$
$$(\alpha + i\beta)(x + iy) = (\alpha x - \beta y) + i(\alpha y + \beta x).$$

Легко проверить, что здесь выполняются все аксиомы линейного комплексного пространства.

Векторы $x + i0$ мы будем отождествлять с векторами $x \in R$ и называть *вещественными векторами пространства* C. Векторы $0 + iy$ мы будем обозначать iy и называть *чисто мнимыми векторами.* Вектор $x - iy$ будем записывать также в виде $\overline{x + iy}$ и называть *комплексно сопряженным* к вектору $x + iy$.

Введем в пространство C скалярное произведение по формуле

$$(x_1+iy_1,\ x_2+iy_2)= \\ =[(x_1,\ x_2)+(y_1,\ y_2)]+i[(y_1,\ x_2)-(x_1,\ y_2)].$$

Легко проверить, что это скалярное произведение удовлетворяет условиям *9.21а—г*. В частности,

$$(x+iy,\ x+iy)=(x,\ x)+(y,\ y).$$

Пространство C содержит пространство R в качестве подмножества, допускающего операции сложения и умножения на вещественные числа и с тем же скалярным произведением.

Всякая ортогональная нормированная система $e_1, \ldots, e_n$ в пространстве R будет ортогональной нормированной системой и в пространстве C. Если $e_1, \ldots, e_n$ есть ортонормальный базис в пространстве R, то эти же векторы образуют и ортонормальный базис в пространстве C.

9.42. Всякий линейный оператор A, заданный в пространстве R продолжается на пространство C по формуле

$$\hat{\mathrm{A}}(x+iy)=\mathrm{A}x+i\mathrm{A}y, \tag{21}$$

причем оператор $\hat{\mathrm{A}}$ оказывается, очевидно, линейным оператором на пространстве C.

Матрица оператора $\hat{\mathrm{A}}$ в пространстве C в базисе $e_1, \ldots, e_n \in \mathrm{R}$ совпадает с матрицей оператора A в пространстве R в том же базисе, поскольку по (21)

$$\hat{\mathrm{A}}e_j=\mathrm{A}e_j \qquad (j=1,\ \ldots,\ n).$$

При продолжении сохраняются алгебраические соотношения между линейными операторами: если $\mathrm{A}+\mathrm{B}=\mathrm{S}$ в пространстве R, то $\hat{\mathrm{A}}+\hat{\mathrm{B}}=\hat{\mathrm{S}}$ в пространстве C; если $\mathrm{AB}=\mathrm{D}$ в пространстве R, то $\hat{\mathrm{A}}\hat{\mathrm{B}}=\hat{\mathrm{D}}$ в пространстве C. Это следует, например, из сохранения матриц при продолжении.

9.43. Пусть A' — сопряженный оператор к оператору A в вещественном пространстве R (*8.91*). Покажем, что продолжением $\hat{A}'$ оператора A' на пространство C служит оператор $\hat{A}^*$, сопряженный к продолжению $\hat{A}$ оператора A. Действительно, для любых $z = x + iy$, $w = u + iv \in C$ мы имеем

$$
\begin{aligned}
(\hat{A}'(x+iy),\ u+iv) &= \\
&= (A'x, u) + i(A'y, u) - i(A'x, v) + (A'y, v) = \\
&= (x, Au) + i(y, Au) - i(x, Av) + (y, Av) = \\
&= (x+iy,\ \hat{A}(u+iv)),
\end{aligned}
$$

что и требуется.

В частности, продолжением симметричного оператора ($A' = A$) является самосопряженный оператор ($A^* = A$), продолжением антисимметричного оператора ($A' = -A$) является антисамосопряженный оператор ($A^* = -A$) и продолжением изометрического оператора ($U' = U^{-1}$) является унитарный оператор ($U^* = U^{-1}$). Наконец, продолжением нормального оператора ($A'A = AA'$) является нормальный оператор ($A^*A = AA^*$).

9.44. Структура вещественного нормального оператора. Пусть σ и τ — вещественные числа. Легко проверяемое матричное равенство

$$
\begin{Vmatrix} \sigma & \tau \\ -\tau & \sigma \end{Vmatrix}
\begin{Vmatrix} \sigma & -\tau \\ \tau & \sigma \end{Vmatrix}
=
\begin{Vmatrix} \sigma & -\tau \\ \tau & \sigma \end{Vmatrix}
\begin{Vmatrix} \sigma & \tau \\ -\tau & \sigma \end{Vmatrix}
=
\begin{Vmatrix} \sigma^2+\tau^2 & 0 \\ 0 & \sigma^2+\tau^2 \end{Vmatrix}
\qquad (22)
$$

показывает, что матрица

$$
\begin{Vmatrix} \sigma & \tau \\ -\tau & \sigma \end{Vmatrix}
$$

перестановочна со своей сопряженной. Более общим образом, перестановочна со своей сопряженной и квазидиагональная (вещественная) матрица

$$\left\|\begin{array}{cccccccccc}
\sigma_1 & \tau_1 & & & & & & & & \\
-\tau_1 & \sigma_1 & & & & & & & & \\
 & & \sigma_2 & \tau_2 & & & & & & \\
 & & -\tau_2 & \sigma_2 & & & & & & \\
 & & & & \ddots & & & & & \\
 & & & & & \sigma_m & \tau_m & & & \\
 & & & & & -\tau_m & \sigma_m & & & \\
 & & & & & & & \lambda_{m+1} & & \\
 & & & & & & & & \ddots & \\
 & & & & & & & & & \lambda_r
\end{array}\right\| \tag{23}$$

порядка $n = 2m + r - m = m + r$.

Теорема. *Пусть* A — *нормальный оператор в вещественном евклидовом пространстве* R_n. *Тогда в этом пространстве существует ортогональный базис* $e_1, \ldots, e_n$, *в котором матрица оператора* A *принимает вид* (23). *При этом числа* $\lambda_j = \sigma_j + i\tau_j$ $(j = 1, \ldots, m)$ и $\lambda_{m+1}, \ldots, \lambda_r$ *определяются однозначно оператором* A. *Именно, они являются комплексными* $(\lambda_1, \ldots, \lambda_m)$ *или вещественными* $(\lambda_{m+1}, \ldots, \lambda_r)$ *корнями характеристического уравнения* $\det\|A - \lambda E\| = 0$ *и повторяются в матрице* (23) *столько раз, какова кратность соответствующего корня.*

Доказательство. Построим унитарное пространство C_n со скалярным произведением, продолжающим скалярное произведение (x, y) на пространство R_n. Продолжим операторы A и A′ на пространство C_n по правилу *9.42*; как мы видели в *9.43*, они продолжаются до нормального оператора $\hat{A}$ и его сопряженного $\hat{A}^*$.

Выберем в пространстве R_n, пока произвольно, ортонормальный базис $f_1, \ldots, f_n$; мы видели, что векторы $f_1, \ldots, f_n$ образуют ортонормальный базис и в пространстве C_n. Матрицу оператора A в базисе $f_1, \ldots, f_n$ обозначим $\|a_{jk}\|$; числа a_{jk} вещественны. Эта же матрица представляет оператор $\hat{A}$ в базисе $f_1, \ldots, f_n$ во всем пространстве C_n.

Так как характеристическое уравнение $\det\|A-\lambda E\|=0$ имеет вещественные коэффициенты, то вместе с каждым невещественным корнем λ_j является корнем также сопряженная величина $\bar{\lambda}_j$. С учетом этого последовательность *различных* корней запишем в следующей форме:

$$\lambda_1,\ \bar{\lambda}_1,\ \lambda_2,\ \bar{\lambda}_2,\ \ldots,\lambda_p,\ \bar{\lambda}_p;\ \lambda_{p+1},\ \ldots,\lambda_q,$$

где корни $\lambda_1,\ldots,\lambda_p$ невещественны, а корни $\lambda_{p+1},\ldots,\lambda_q$ вещественны. В силу *9.33б*, пространство C_n распадается в ортогональную сумму подпространств Λ_1, $\bar{\Lambda}_1,\ldots,\Lambda_p$, $\bar{\Lambda}_p$; $\Lambda_{p+1},\ldots,\Lambda_q$, где Λ_j состоит из собственных векторов оператора A с собственным значением λ_j, а $\bar{\Lambda}_j$ — из собственных векторов оператора A с собственным значением $\bar{\lambda}_j$; $\bar{\Lambda}_{p+1}=\Lambda_{p+1},\ldots,\bar{\Lambda}_q=\Lambda_q$.

Пусть $z=x+iy\in\Lambda_j$. Уравнение $Az=\lambda_j z$ в координатах (относительно исходного базиса $f_1,\ldots,f_n$) записывается в форме

$$\sum_{k=1}^{n} a_{jk}\zeta_k=\lambda_j\zeta_j,$$

где $z=(\zeta_1,\ldots,\zeta_n)=(\xi_1+i\eta_1,\ldots,\xi_n+i\eta_n)$. Применяя операцию комплексного сопряжения и учитывая вещественность чисел a_{jk}, получаем

$$\sum_{k=1}^{n} a_{jk}\bar{\zeta}_k=\bar{\lambda}_j\bar{\zeta}_j.$$

Это означает, что вектор $\bar{z}=(\bar{\zeta}_1,\ldots,\bar{\zeta}_n)$ также является собственным вектором оператора $\hat{A}$ с собственным значением $\bar{\lambda}$. Отсюда следует, что операция комплексного сопряжения переводит подпространство Λ_j в $\bar{\Lambda}_j$.

Пусть $\bar{\lambda}_1\neq\lambda_1$, так что $\lambda_1=\sigma_1+i\tau_1$, $\tau_1\neq 0$. Возьмем произвольно нормированный вектор $g_1\in\Lambda_1$ и найдем $\bar{g}_1\in\bar{\Lambda}_1$. Положим, далее,

$$e_1=\frac{1}{2}(g_1+\bar{g}_1),\quad e_2=\frac{1}{2i}(g_1-\bar{g}_1),$$

так что

$$g_1=e_1+ie_2,\quad \bar{g}_1=e_1-ie_2.$$

Векторы e_1 и e_2, очевидно, вещественны. Так как векторы g_1 и $\overline{g_1}$ ортогональны, то векторы e_1 и e_2 имеют длину 1. При этом $(e_1, e_2)=0$; действительно,

$$(e_1, e_2) = -\frac{1}{4i}(g_1+\overline{g_1},\ g_1-\overline{g_1}) = \\ = -\frac{1}{4i}\left[(g_1, g_1)-(\overline{g_1}, \overline{g_1})\right]=0,$$

поскольку $(g_1, \overline{g_1})=0$, $(g_1, g_1)=(\overline{g_1}, \overline{g_1})=1$.

Далее, мы имеем

$$\mathrm{A}e_1=\hat{\mathrm{A}}e_1=\frac{1}{2}(\hat{\mathrm{A}}g_1+\hat{\mathrm{A}}\overline{g_1})=\frac{1}{2}(\lambda_1 g_1+\overline{\lambda}_1\overline{g_1})= \\ =\frac{1}{2}\left[(\sigma_1+i\tau_1)(e_1+ie_2)+(\sigma_1-i\tau_1)(e_1-ie_2)\right]=\sigma_1 e_1-\tau_1 e_2,$$

$$\mathrm{A}e_2=\hat{\mathrm{A}}e_2=\frac{1}{2i}(\hat{\mathrm{A}}g_1-\hat{\mathrm{A}}g_2)=\frac{1}{2i}(\lambda_1 g_1-\overline{\lambda_1 g_1})=\sigma_1 e_2+\tau_1 e_1.$$

Таким образом, оператор A преобразует плоскость векторов e_1, e_2 в себя с матрицей (в базисе e_1, e_2)

$$\begin{Vmatrix} \sigma_1 & \tau_1 \\ -\tau_1 & \sigma_1 \end{Vmatrix}. \tag{24}$$

Если Λ_1 имеет более одного измерения, возьмем вектор $g_2 \in \Lambda_1$ ортогонально к g_1 и сопряженный вектор $\overline{g_2} \in \overline{\Lambda_1}$; последний уже автоматически будет ортогональным к g_2. Повторим для g_2 и $\overline{g_2}$ предыдущее построение; мы получим новую пару вещественных векторов e_3, e_4, линейно выражающихся через g_2 и $\overline{g_2}$ и поэтому ортогональных к векторам e_1, e_2 (которые линейно выражались через g_1 и g_2); плоскость, определяемая парой e_3, e_4, преобразуется оператором A в себя с той же матрицей (24). Продолжая построение, мы построим взаимно ортогональные вещественные векторы $e_1, e_2, \ldots, e_{2p-1}, e_{2p}$; каждую пару e_{2k-1}, e_{2k} оператор A преобразует в их плоскости с помощью матрицы, аналогичной (24).

Пусть теперь $\lambda_{p+1}=\overline{\lambda}_{p+1}$ вещественно. Операция перехода к сопряженным векторам переводит подпространство Λ_{p+1} в себя. Пусть $g \in \Lambda_{p+1}$ — любой вектор и $\overline{g}$ — ему сопряженный.

Возможно одно из двух: либо векторы g и $\bar{g}$ линейно независимы (в C_n), либо линейно зависимы.

Если g и $\bar{g}$ линейно независимы, то и векторы

$$e=\frac{1}{2}(g+\bar{g}), \quad f=\frac{1}{2i}(g-\bar{g})$$

линейно независимы. Векторы e и f вещественны, и поскольку они лежат в Λ_{p+1} вместе с g и $\bar{g}$, они являются собственными векторами оператора А с тем же собственным значением λ_{p+1}. Если g и $\bar{g}$ линейно зависимы, то, так как они равны по норме, можно написать

$$\bar{g}=e^{2i\varphi}g, \qquad 0\leqslant\varphi<\pi,$$

или

$$e^{i\varphi}g=e^{-i\varphi}\bar{g}=\overline{e^{i\varphi}g},$$

так что вектор $e^{i\varphi}g$ уже вещественный. При этом, поскольку он лежит в Λ_{p+1} вместе с g, он является собственным для оператора А с тем же собственным значением λ_{p+1}.

Таким образом, в подпространстве Λ_{p+1} можно указать базис из вещественных векторов. Применяя к ним процесс ортогонализации (*8.61*), мы получим в Λ_{p+1} уже ортогональный и, далее, ортонормальный базис. Произведя аналогичную процедуру в $\Lambda_{p+2}, \ldots, \Lambda_q$, завершаем доказательство теоремы.

Представление нормального оператора в форме (23) дает возможность выяснить его геометрический смысл.

Оператор с матрицей (24) в плоскости векторов e_1, e_2 можно истолковать как оператор поворота с растяжением. Действительно, мы имеем

$$\begin{Vmatrix} \sigma & \tau \\ -\tau & \sigma \end{Vmatrix}=\sqrt{\sigma^2+\tau^2}\begin{Vmatrix} \frac{\sigma}{\sqrt{\sigma^2+\tau^2}} & \frac{\tau}{\sqrt{\sigma^2+\tau^2}} \\ -\frac{\tau}{\sqrt{\sigma^2+\tau^2}} & \frac{\sigma}{\sqrt{\sigma^2+\tau^2}} \end{Vmatrix}=$$

$$=M\begin{Vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{Vmatrix},$$

где

$$M=\sqrt{\sigma^2+\tau^2}, \quad \cos\alpha=\frac{\sigma}{\sqrt{\sigma^2+\tau^2}}, \quad \sin\alpha=\frac{\tau}{\sqrt{\sigma^2+\tau^2}}.$$

Матрица

$$\begin{Vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{Vmatrix}$$

определяет в плоскости e_1, e_2 поворот всех векторов на угол α, а коэффициент M есть коэффициент растяжения.

В целом нормальный оператор, как видно из представления (23), в m взаимно ортогональных плоскостях осуществляет повороты с растяжениями и в $r-m$ направлениях, ортогональных друг к другу и к указанным плоскостям, только растяжения (в $\lambda_{m+1}, \ldots, \lambda_r$ раз) *).

9.45. Структура вещественного симметричного оператора. Если оператор A в пространстве R_n симметричен, $\mathrm{A}' = \mathrm{A}$, то продолжение $\hat{\mathrm{A}}$ оператора A в пространство C_n есть самосопряженный оператор, $\mathrm{A}^* = \mathrm{A}$. Все собственные значения $\lambda_1, \ldots, \lambda_n$ самосопряженного оператора вещественны (*9.34*), поэтому в представлении (23) клетки вида (24) отсутствуют и остаются лишь диагональные элементы. Мы получаем теорему:

Теорема. *Для всякого симметричного оператора* A *в пространстве* R_n *существует ортонормальный базис из собственных векторов.*

Геометрически симметричный оператор осуществляет по каждому из n ортогональных направлений $e_1, \ldots, e_n$ растяжение (соответственно в $\lambda_1, \ldots, \lambda_n$ раз). Так как числа $\lambda_1, \ldots, \lambda_n$ суть корни уравнения $\det \|A - \lambda E\| = 0$, то, в частности, для симметричной матрицы $A = \|a_{jk}\|$ характеристическое уравнение $\det \|A - \lambda E\| = 0$ всегда имеет n вещественных корней (не обязательно различных) и вовсе не имеет невещественных корней.

9.46. Структура вещественного антисимметричного оператора. Если оператор A в пространстве R_n антисимметричен, $\mathrm{A}' = -\mathrm{A}$, то продолжение $\hat{\mathrm{A}}$ оператора A в пространство C_n есть антисамосопряженный

---

*) При $0 < \lambda_k < 1$ растяжение в λ_k раз на самом деле есть сжатие. При $\lambda_k < 0$ растяжение в λ_k раз на самом деле есть растяжение, соединенное с отражением.

оператор, $A^* = -A$. Все собственные значения $\lambda_1, \ldots, \lambda_n$ антисамосопряженного оператора чисто мнимы (*9.35*). Поэтому в представлении (23) клетки вида (24) имеют специальный вид

$$\begin{Vmatrix} 0 & \tau \\ -\tau & 0 \end{Vmatrix},$$

а числа $\lambda_{m+1}, \ldots, \lambda_r$ могут быть лишь равными 0. Мы получаем теорему.

Теорема. *Для всякого антисимметричного оператора* A *в пространстве* R_n *существует ортонормальный базис, в котором матрица оператора* A *записывается в виде*

$$\begin{Vmatrix} 0 & \tau_1 & & & & & & & & & \\ -\tau_1 & 0 & & & & & & & & & \\ & & 0 & \tau_2 & & & & & & & \\ & & -\tau_2 & 0 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & 0 & \tau_m & & & & \\ & & & & & -\tau_m & 0 & & & & \\ & & & & & & & 0 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & 0 \end{Vmatrix}. \qquad (25)$$

Обратно, если матрица некоторого оператора A в пространстве R_n в некотором ортогональном базисе имеет вид (25), то оператор A антисимметричен (*8.92б*).

Геометрически антисимметричный оператор осуществляет в каждой из m взаимно ортогональных плоскостей поворот на 90° с последующим растяжением (соответственно в $\tau_1, \ldots, \tau_m$ раз), а все векторы, ортогональные к указанным плоскостям, переводит в 0.

9.47. Структура вещественного изометричного оператора. Если оператор A в пространстве R_n изометричен, $A' = A^{-1}$, то продолжение $\hat{A}$ оператора A в пространство C_n есть унитарный оператор, $\hat{A}^* = \hat{A}^{-1}$. Все собственные значения $\lambda_1, \ldots, \lambda_n$ унитарного оператора по модулю равны 1 (*9.36*). Можно написать $\lambda_j = \cos\alpha_j - i\sin\alpha_j$ ($j = 1, \ldots, m$). Поэтому в представлении (23) клетки вида (24) имеют специальный вид

$$\left\| \begin{matrix} \cos\alpha_j & \sin\alpha_j \\ -\sin\alpha_j & \cos\alpha_j \end{matrix} \right\|, \tag{26}$$

а числа $\lambda_{m+1}, \ldots, \lambda_r$ могут быть равными лишь ± 1. Мы получаем теорему:

Теорема. *Для всякого изометричного оператора* A *в пространстве* R_n *существует ортонормальный базис, в котором матрица оператора* A *записывается в виде*

$$\left\| \begin{matrix} \cos\alpha_1 & \sin\alpha_1 & & & & & & \\ -\sin\alpha_1 & \cos\alpha_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \cos\alpha_m & \sin\alpha_m & & & \\ & & & -\sin\alpha_m & \cos\alpha_m & & & \\ & & & & & \pm 1 & & \\ & & & & & & \ddots & \\ & & & & & & & \pm 1 \end{matrix} \right\|. \tag{27}$$

Геометрически изометричный оператор A в каждой из m взаимно ортогональных плоскостей осуществляет поворот на некоторый угол (без растяжений) и в $r-m$ направлениях, ортогональных к этим плоскостям и друг к другу, действует как оператор E или $-$E. Впрочем, каждые два из таких направлений с одинаковыми коэффициентами растяжения (оба $+1$ или оба -1) можно объединить в плоскость, на которой оператор A также осуществляет поворот (на $0°$ или на $180°$). Произведя все такие возможные объединения, мы получим, при n нечетном, последнее направление с коэффициентом $+1$ или -1; при n четном, возможно, два последних направления с коэффициентами $+1$ или -1. Наличие среди оставшихся коэффициентов -1 показывает, что к имеющимся поворотам присоединено отражение относительно некоторой координатной плоскости, например ортогональной к вектору e_n. В этом случае мы имеем $\det A = -1$; в случае, когда отражения нет, имеем $\det A = +1$.

ЗАДАЧИ

1. Самосопряженный оператор, действующий в унитарном пространстве C_n, называется *неотрицательным* (*положительным*), если все его собственные значения $\lambda_1, \ldots, \lambda_n$ неотрицательны (положительны). Показать, что квадрат всякого самосопряженного оператора неотрицателен.

2. Показать, что для всякого самосопряженного неотрицательного (положительного) оператора A можно найти, причем единственным образом, такой неотрицательный (положительный) оператор B, что $B^2 = A$ («квадратный корень из оператора A»).

3. Извлечь квадратный корень из оператора A, заданного в ортогональном и нормированном базисе e_1, e_2, e_3 матрицей

$$A = \begin{Vmatrix} 13 & 14 & 4 \\ 14 & 24 & 18 \\ 4 & 18 & 29 \end{Vmatrix}.$$

4. Если A — произвольный линейный оператор, действующий в унитарном пространстве C_n, A^* — ему сопряженный, то A^*A есть неотрицательный оператор. Если A невырожден, то A^*A — положительный оператор.

5. Известно, что некоторый линейный оператор A есть произведение самосопряженного оператора S и унитарного Q: $A = SQ$. Показать, что $S^2 = AA^*$.

6. Показать, что всякий линейный оператор A с $\det A \neq 0$ может быть представлен как произведение самосопряженного и унитарного операторов.

7. Доказать единственность представления оператора A в виде произведения SQ в условиях задачи 6.

8. Линейный оператор V, действующий в C_n, называется нерастягивающим, если $|Vx| \leqslant |x|$ для любого x. Показать, что любой линейный оператор A может быть представлен как произведение самосопряженного и нерастягивающего.

9. Показать, что самосопряженные операторы A и B перестановочны тогда и только тогда, когда они имеют общую систему из n взаимно ортогональных собственных векторов.

10. Для каждого линейного оператора A, действующего в пространстве C_n, указать ортонормальный базис, в котором матрица оператора A имеет треугольный вид:

$$A = \begin{Vmatrix} a_1^{(1)} & a_2^{(1)} & \ldots & a_n^{(1)} \\ 0 & a_2^{(2)} & \ldots & a_n^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_n^{(n)} \end{Vmatrix}.$$

КВАДРАТИЧНЫЕ ФОРМЫ В ЕВКЛИДОВОМ И УНИТАРНОМ ПРОСТРАНСТВАХ

§ 10.1. Основная теорема о квадратичных формах в евклидовом пространстве

10.11. Начнем со следующего предложения, касающегося симметричной билинейной формы в евклидовом пространстве:

Теорема. *В n-мерном евклидовом пространстве всякая симметричная билинейная форма имеет канонический базис из взаимно ортогональных векторов.*

Доказательство. Рассмотрим линейный оператор A, отвечающий данной симметричной билинейной форме $\mathrm{A}(x, y)$ (*8.91*). Этот оператор также симметричен. Согласно теореме о симметричном операторе (*9.45*) в пространстве R имеется ортогональный и нормированный базис из собственных векторов оператора A. В этом базисе матрица оператора A диагональна. Поскольку эта же матрица является и матрицей билинейной формы $\mathrm{A}(x, y)$, построенный базис есть канонический базис формы $\mathrm{A}(x, y)$, что и требовалось.

10.12. Этот результат мы применим теперь для изучения квадратичных форм.

Пусть дана квадратичная форма

$$\mathrm{A}(x, x)=\sum_{i, j=1}^{n} a_{ij}\xi_i\xi_j \qquad (a_{ij}=a_{ji}). \tag{1}$$

Будем считать числа $\xi_1, \xi_2, \ldots, \xi_n$ координатами вектора x в евклидовом пространстве R_n со скалярным произведением, определенным по формуле

$$(x, y)=\sum_{i=1}^{n} \xi_i\eta_i,$$

где $y=(\eta_1, \eta_2, \ldots, \eta_n)$. Базис

$e_1=\{1, 0, \ldots, 0\}$, $e_2=\{0, 1, \ldots, 0\}$, $\ldots$, $e_n=\{0, 0, \ldots, 1\}$

является ортогональным и нормированным базисом в R_n, причем

$$x = \sum_{i=1}^{n} \xi_i e_i, \quad y = \sum_{i=1}^{n} \eta_i e_i.$$

Рассмотрим билинейную форму

$$A(x, y) = \sum_{i, j=1}^{n} a_{ij} \xi_i \eta_j,$$

соответствующую квадратичной форме (1). В силу теоремы *10.11* у этой формы существует ортогональный и нормированный канонический базис $f_1, f_2, \ldots, f_n$. Если относительно этого базиса векторы x и y имеют соответственно координаты $\tau_1, \tau_2, \ldots, \tau_n$ и $\theta_1, \theta_2, \ldots, \theta_n$, то форма $A(x, y)$ будет иметь вид

$$A(x, y) = \sum_{i=1}^{n} \lambda_i \tau_i \theta_i,$$

а квадратичная форма $A(x, x)$ примет вид

$$A(x, x) = \sum_{i=1}^{n} \lambda_i \tau_i^2. \tag{2}$$

Переход от базиса $e_1, e_2, \ldots, e_n$ к базису $f_1, f_2, \ldots, f_n$ осуществляется с помощью некоторой ортогональной матрицы (*8.93*) $Q = \| q_i^{(j)} \|$ по формулам

$$f_j = \sum_{i=1}^{n} q_i^{(j)} e_i \qquad (j = 1, 2, \ldots, n).$$

Связь между координатами $\tau_1, \tau_2, \ldots, \tau_n$ и $\xi_1, \xi_2, \ldots, \xi_n$ может быть записана в силу формул (36) из *8.94* системой равенств

$$\xi_i = \sum_{j=1}^{n} q_i^{(j)} \tau_j \qquad (i = 1, 2, \ldots, n) \tag{3}$$

с использованием транспонированной матрицы Q'. Мы получили тем самым следующую важную теорему:

Теорема. (Теорема о квадратичной форме в евклидовом пространстве.) *Всякая квадратичная форма* (1) *может быть приведена к каноническому виду* (2) *с помощью изометрического преобразования координат* (3).

10.13. Последовательность действий, которые нужно произвести для построения формул перехода (3) и канонического вида (2) квадратичной формы (1), вытекает из результатов *4.94* и *9.45*; мы приведем ее здесь в окончательном виде.

а) *По квадратичной форме* (1) *строим симметричную матрицу* $A = \| a_{ij} \|$.

б) *Составляем характеристический многочлен* $\Delta(\lambda) = \det(A - \lambda E)$ *и находим его корни.* В силу *9.45* этот многочлен имеет n вещественных корней (не обязательно различных).

в) *Зная корни многочлена* $\Delta(\lambda)$, *можно написать уже квадратичную форму в каноническом виде* (2), *в частности, можно сказать, каковы ее положительный и отрицательный индексы инерции.*

г) *Корень* λ_1 *подставим в систему* (23) *из 4.94.* Для данного корня λ_1 эта система имеет ровно столько линейно независимых решений, какова кратность корня λ_1. *Найдем эти линейно независимые решения, пользуясь правилами решения однородных систем линейных уравнений.*

д) *Если кратность корня* λ_1 *больше единицы, ортогонализируем полученные линейно независимые решения по методу 8.61.*

е) *Проделав указанные операции с каждым корнем, мы получим систему из* n *взаимно ортогональных векторов. Пронормируем ее, разделив каждый вектор на его длину. Полученные векторы*

$$
\begin{aligned}
f_1 &= \{q_1^{(1)},\ q_2^{(1)},\ \ldots,\ q_n^{(1)}\},\\
f_2 &= \{q_1^{(2)},\ q_2^{(2)},\ \ldots,\ q_n^{(2)}\},\\
&\ldots\ldots\ldots\ldots\ldots\\
f_n &= \{q_1^{(n)},\ q_2^{(n)},\ \ldots,\ q_n^{(n)}\}
\end{aligned}
$$

образуют уже ортогональную и нормированную систему.

ж) *Используя числа* $q_i^{(j)}$, *можно написать формулы перехода* (3).

з) *Если требуется дать выражения новых координат* $\{\tau\}$ *через старые координаты* $\{\xi\}$, *то, поскольку матрица, обратная к ортогональной, получается транспонированием, мы* *можем написать искомые выражения в виде*

$$\tau_i = \sum_{j=1}^{n} q_j^{(i)} \xi_j \qquad (i = 1, 2, \ldots, n).$$

10.14. Мы видели в *7.33а*, что в аффинном пространстве ни канонический базис, ни канонический вид квадратичной формы не определены однозначно; вообще говоря, можно было включить в канонический базис формы любой наперед заданный вектор. В евклидовом пространстве и при условии, что рассматриваются только ортогональные и нормированные базисы, положение иное. Дело в том, что вместе с матрицей квадратичной формы, как мы видели, преобразуется и матрица соответствующего симметричного линейного оператора; если найден канонический базис квадратичной формы, то одновременно найден базис из собственных векторов симметричного оператора. При этом коэффициенты квадратичной формы в каноническом базисе («канонические коэффициенты») совпадают с соответствующими собственными значениями оператора. Но собственные значения оператора A суть корни уравнения $\det(\mathrm{A} - \lambda \mathrm{E}) = 0$, которое не зависит от выбора базиса и инвариантно связано с оператором A. Следовательно, *совокупность канонических коэффициентов формы* $(\mathrm{A}x, x)$ *определена однозначно*. Что же касается канонического базиса квадратичной формы $(\mathrm{A}x, x)$, то он определен с той же степенью произвола, с какой определена полная ортогональная и нормированная система собственных векторов оператора A: не считая взаимных перестановок этих векторов, можно любой из них умножить на -1; более общим является любое изометрическое преобразование в собственном подпространстве, отвечающем фиксированному собственному значению λ.

§ 10.2. Экстремальные свойства квадратичной формы

10.21. Пусть в евклидовом пространстве R_n задана квадратичная форма $\mathrm{A}(x, x)$. Будем рассматривать ее значения на единичной сфере пространства R_n, т. е. при $(x, x) = 1$, и поставим следующий вопрос: в каких точках единичной сферы значения формы стационарны? Напомним, что дифференцируемая числовая функция $f(x)$, определенная для точек некоторой поверхности U, принимает по определению в точке $x_0 \in U$ *стационарное значение*, если в точке x_0 производная функции $f(x)$ по любому направлению на поверхности U равна нулю. В частности, функция $f(x)$ стационарна в тех точках, где она достигает максимума или минимума.

Задача об определении стационарных значений есть задача на условный экстремум; одним из методов ее решения является метод Лагранжа, который мы сейчас и используем *). Возьмем в пространстве R_n ортогональный и нормированный базис и обозначим через $\xi_1, \xi_2, \ldots, \xi_n$ координаты вектора x в этом базисе. В этих координатах квадратичная форма будет иметь вид $A(x, x) = \sum_{i, j=1}^{n} a_{ij}\xi_i\xi_j$, а условие $(x, x) = 1$ запишется равенством $\sum_{i=1}^{n} \xi_i^2 = 1$. Следуя методу Лагранжа, построим функцию

$$F(\xi_1, \xi_2, \ldots, \xi_n) = \sum_{i, j=1}^{n} a_{ij}\xi_i\xi_j - \lambda \sum_{i=1}^{n} \xi_i^2$$

и приравняем нулю ее первую частную производную по ξ_i $(i = 1, 2, \ldots, n)$:

$$2 \sum_{j=1}^{n} a_{ij}\xi_j - 2\lambda\xi_i = 0 \qquad (i = 1, 2, \ldots, n).$$

После сокращения на 2 мы получаем уже знакомую нам систему **) (23) из *4.94*:

$$\begin{aligned}
(a_{11} - \lambda)\,\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n &= 0, \\
a_{21}\xi_1 + (a_{22} - \lambda)\,\xi_2 + \ldots + a_{2n}\xi_n &= 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + (a_{nn} - \lambda)\,\xi_n &= 0,
\end{aligned}$$

которая служила для определения собственных векторов симметричного оператора, отвечающего квадратичной форме $A(x, x)$. Отсюда вытекает следующее предложение:

---

*) См., например, В. И. Смирнов, Курс высшей математики, т. I, стр. 392, Гостехиздат, 1951.

**) Напомним, что $a_{ij} = a_{ji}$ $(i, j = 1, 2, \ldots, n)$.

Квадратичная форма $\mathrm{A}(x, x)$ *принимает стационарные значения на тех векторах единичной сферы, которые являются собственными векторами симметричного оператора* A, *соответствующего форме* $\mathrm{A}(x, x)$.

10.22. Вычислим значения, которые принимает форма в стационарных точках. Для этого введем соответствующий симметричный оператор A и представим квадратичную форму в виде $\mathrm{A}(x, x) = (\mathrm{A}x, x)$. Так как, по доказанному, форма $\mathrm{A}(x, x)$ принимает стационарное значение на собственном векторе e_i оператора A, то мы имеем $\mathrm{A}e_i = \lambda_i e_i$; отсюда

$$\mathrm{A}(e_i, e_i) = (\mathrm{A}e_i, e_i) = \lambda_i (e_i, e_i) = \lambda_i.$$

Итак, *стационарное значение формы* $\mathrm{A}(x, x)$ *при* $x = e_i$ *равно соответствующему собственному значению оператора* A. Так как собственные значения оператора A совпадают с каноническими коэффициентами формы $\mathrm{A}(x, x)$, то мы можем, далее, заключить, что *стационарные значения формы* $\mathrm{A}(x, x)$ *совпадают с ее каноническими коэффициентами*.

В частности, максимум формы $\mathrm{A}(x, x)$ на единичной сфере равен наибольшему из ее канонических коэффициентов, минимум — наименьшему.

10.23. Как билинейную, так и квадратичную форму $\mathrm{A}(x, x)$ можно рассматривать не во всем n-мерном пространстве R_n, а в некотором k-мерном подпространстве $\mathrm{R}_k \subset \mathrm{R}_n$, и разыскивать в этом подпространстве ортогональный и нормированный канонический базис. Пусть форма $\mathrm{A}(x, x)$ во всем пространстве R_n имеет канонический вид

$$\mathrm{A}(x, x) = \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \ldots + \lambda_n \xi_n^2, \tag{4}$$

а в подпространстве R_k — канонический вид

$$\mathrm{A}(x, x) = \mu_1 \tau_1^2 + \mu_2 \tau_2^2 + \ldots + \mu_k \tau_k^2.$$

Выясним, как связаны коэффициенты $\mu_1, \mu_2, \ldots, \mu_k$ с коэффициентами $\lambda_1, \lambda_2, \ldots, \lambda_n$. Для удобства предположим, что нумерация канонических коэффициентов произведена в порядке их убывания, так что

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n, \quad \mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_k.$$

Величина λ_1, как мы уже знаем, есть максимальное значение квадратичной формы $\mathrm{A}(x, x)$ на единичной сфере пространства R_n; аналогично μ_1 есть максимальное значение квадратичной формы $\mathrm{A}(x, x)$ на единичной сфере подпространства R_k, и потому $\mu_1 \leqslant \lambda_1$. Покажем, что $\mu_1 \geqslant \lambda_{n-k-1}$.

Пусть $e_1, e_2, \ldots, e_n$ — канонический базис формы $\mathrm{A}(x, x)$, в котором она записывается в виде (4). Рассмотрим $(n-k+1)$-мерное подпространство R', порожденное векторами $e_1, e_2, \ldots, e_{n-k+1}$. Так как $k+(n-k+1) > n$, то подпространства R' и R_k в силу *2.47в* имеют хотя бы один общий ненулевой вектор. Пусть это вектор $x_0 = \{\xi_1^{(0)}, \ldots, \xi_{n-k+1}^{(0)}, 0, \ldots, 0\}$; предположим, что x_0 нормирован, т. е. что $|x_0| = 1$. Для вектора x_0 по формуле (4) имеем

$$\mathrm{A}(x_0, x_0) = \lambda_1 \xi_1^{(0)^2} + \ldots + \lambda_{n-k+1} \xi_{n-k+1}^{(0)^2} \geqslant \\ \geqslant \lambda_{n-k+1} \left(\xi_1^{(0)^2} + \ldots + \xi_{n-k+1}^{(0)^2}\right) = \lambda_{n-k+1}.$$

Отсюда вытекает, что μ_1 как максимальное значение квадратичной формы $\mathrm{A}(x, x)$ на единичной сфере подпространства R_k не может быть меньше, чем λ_{n-k+1}, что и требуется.

Таким образом, величина μ_1 заключена в следующих границах:

$$\lambda_1 \geqslant \mu_1 \geqslant \lambda_{n-k+1}. \tag{5}$$

10.24. Для различных k-мерных подпространств величина μ_1 принимает, естественно, различные значения. Покажем, что *существуют такие k-мерные подпространства, для которых μ_1 принимает крайние значения, указанные в неравенстве* (5).

Рассмотрим подпространство R', порожденное первыми k векторами $e_1, e_2, \ldots, e_k$ канонического базиса формы $\mathrm{A}(x, x)$. В подпространстве R' в базисе $e_1, e_2, \ldots, e_k$ форма $\mathrm{A}(x, x)$ имеет вид

$$\mathrm{A}(x, x) = \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \ldots + \lambda_k \xi_k^2.$$

В частности,

$$\mathrm{A}(e_1, e_1) = \lambda_1 = \max \mathrm{A}(x, x) \qquad (|x| = 1,\ x \in \mathrm{R}').$$

Таким образом, на подпространстве R' величина

$$\mu_1 = \max A(x, x) \qquad (|x| = 1, \ x \in R')$$

достигает наибольшего возможного значения λ_1.

Рассмотрим теперь подпространство R'', порожденное последними k векторами $e_{n-k+1}, e_{n-k+2}, \ldots, e_n$ канонического базиса формы $A(x, x)$. В подпространстве R'' в базисе $e_{n-k+1}, \ldots, e_n$ форма $A(x, x)$ имеет вид

$$A(x, x) = \lambda_{n-k+1}\xi_{n-k+1}^2 + \ldots + \lambda_n \xi_n^2.$$

В частности,

$$A(e_{n-k-1}, e_{n-k-1}) = \lambda_{n-k+1} = \max A(x, x) \ \ (|x| = 1, \ x \in R'').$$

Поскольку теперь $\mu_1 = \max A(x, x)$ $(|x| = 1, \ x \in R'')$, мы заключаем, так же как и выше, что $\mu_1 = \lambda_{n-k+1}$. Следовательно, на подпространстве R'' величина μ_1 достигает своего наименьшего значения λ_{n-k+1}.

Мы получаем, таким образом, новое определение коэффициента λ_{n-k+1}; *коэффициент λ_{n-k+1} в канонической записи формы $A(x, x)$ равен наименьшему значению максимума квадратичной формы $A(x, x)$ на единичных сферах всех возможных k-мерных подпространств пространства R_n.*

10.25. Используя этот результат, мы можем дать оценки для остальных канонических коэффициентов формы $A(x, x)$ на подпространстве R_k. Например, если фиксировано подпространство R_k, то μ_2 есть наименьший из максимумов квадратичной формы $A(x, x)$ на единичных сферах $(k-1)$-мерных подпространств пространства R_k. В то же время λ_{n-k+2} есть наименьший из максимумов квадратичной формы $A(x, x)$ на единичных сферах *всех* $(k-1)$-мерных подпространств пространства R_n; поэтому $\mu_2 \geqslant \lambda_{n-k+2}$. Аналогично $\mu_3 \geqslant \lambda_{n-k+3}$, $\mu_4 \geqslant \lambda_{n-k+4}$, $\ldots$, $\mu_k \geqslant \lambda_n$. С другой стороны, λ_2 есть наименьший из максимумов квадратичной формы $A(x, x)$ на единичных сферах $(n-1)$-мерных подпространств пространства R_n; но каждое $(n-1)$-мерное подпространство пересекается с подпространством R_k согласно нашей лемме по подпространству, имеющему не менее $(n-1)+k-n=k-1$ измерений; поэтому число λ_2 не менее чем наименьший из максимумов формы $A(x, x)$ на единичных сферах этих подпространств и, в частности, не менее чем число μ_2 — наименьший из максимумов формы $A(x, x)$ на единичных сферах $(k-1)$-мерных подпространств пространства R_k. Следовательно, $\lambda_2 \geqslant \mu_2$. Аналогично $\lambda_3 \geqslant \mu_3, \ldots, \lambda_k \geqslant \mu_k$. Итак, канонические коэффициенты $\mu_1, \mu_2, \ldots, \mu_k$ удовлетворяют неравенствам

$$\left.\begin{array}{l} \lambda_1 \geqslant \mu_1 \geqslant \lambda_{n-k+1}, \\ \lambda_2 \geqslant \mu_2 \geqslant \lambda_{n-k+2}, \\ \cdots\cdots\cdots \\ \lambda_k \geqslant \mu_k \geqslant \lambda_n. \end{array}\right\} \tag{6}$$

При $k=n-1$ неравенства (6) приобретают следующий вид:

$$\left.\begin{array}{c} \lambda_1 \geqslant \mu_1 \geqslant \lambda_2, \\ \lambda_2 \geqslant \mu_2 \geqslant \lambda_3, \\ \cdots\cdots\cdots \\ \lambda_{n-1} \geqslant \mu_{n-1} \geqslant \lambda_n. \end{array}\right\} \tag{7}$$

10.26*. Если $(n-1)$-мерное подпространство R_{n-1}, на котором рассматривается квадратичная форма $A(x, x)$, задано уравнением

$$\alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_n\xi_n = 0 \qquad (\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2 = 1), \tag{8}$$

то коэффициенты $\mu_1, \mu_2, \ldots, \mu_{n-1}$ можно вычислить эффективно.

В предположении, что все числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ различны, приведем метод вычисления этих коэффициентов, предложенный М. Г. Крейном.

Из коэффициентов $\alpha_1, \alpha_2, \ldots, \alpha_n$ по крайней мере один отличен от нуля. Пусть, например, $\alpha_n \neq 0$. Тогда из уравнения (8) мы получаем

$$\xi_n = -\frac{1}{\alpha_n} \sum_{j=1}^{n-1} \alpha_j \xi_j.$$

Подставляя ξ_n в выражение квадратичной формы

$$A(x, x) = \sum_{k=1}^{n} \lambda_k \xi_k^2,$$

мы получаем, что в подпространстве R_{n-1} в координатах $\xi_1, \xi_2, \ldots, \xi_{n-1}$ форма $A(x, x)$ имеет вид

$$A(x, x) = \lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \ldots + \lambda_{n-1}\xi_{n-1}^2 + \frac{\lambda_n}{\alpha_n^2}\left(\sum_{j=1}^{n-1}\alpha_j\xi_j\right)^2.$$

Канонические коэффициенты этой квадратичной формы определяются как ее стационарные значения на единичной сфере подпространства R_{n-1} (*10.22*); эта последняя в координатах $\xi_1, \xi_2, \ldots, \xi_{n-1}$ имеет уравнение

$$B(x, x) = \xi_1^2 + \xi_2^2 + \ldots + \xi_{n-1}^2 + \frac{1}{\alpha_n^2}\left(\sum_{j=1}^{n-1}\alpha_j\xi_j\right)^2 = 1.$$

Для определения стационарных значений действуем, как и ранее, по методу Лагранжа: составляем функцию

$$A(x, x) - \lambda B(x, x) = \sum_{j=1}^{n-1}(\lambda_j - \lambda)\xi_j^2 + \frac{\lambda_n - \lambda}{\alpha_n^2}\left(\sum_{j=1}^{n-1}\alpha_j\xi_j\right)^2$$

и приравниваем нулю ее частные производные по ξ_k $(k = 1, 2, \ldots, n-1)$:

$$\xi_k(\lambda_k - \lambda) + \frac{\lambda_n - \lambda}{\alpha_n^2}\left(\sum_{j=1}^{n-1}\alpha_j\xi_j\right)\alpha_k = 0. \tag{9}$$

Искомые коэффициенты $\mu_1, \mu_2, \ldots, \mu_{n-1}$ являются корнями того уравнения, которое получается, когда мы приравняем нулю определитель $D(\lambda)$ системы линейных уравнений (9). Матрица из коэффициентов этой системы есть, очевидно, сумма двух матриц, первая из которых диагональна, с числами $\lambda_k - \lambda$ по диагонали $(k = 1, 2, \ldots, n-1)$, а вторая имеет вид

$$\frac{\lambda_n - \lambda}{\alpha_n^2}\begin{Vmatrix} \alpha_1\alpha_1 & \alpha_2\alpha_1 & \ldots & \alpha_{n-1}\alpha_1 \\ \alpha_1\alpha_2 & \alpha_2\alpha_2 & \ldots & \alpha_{n-1}\alpha_2 \\ \cdot & \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \alpha_1\alpha_{n-1} & \alpha_2\alpha_{n-1} & \ldots & \alpha_{n-1}\alpha_{n-1} \end{Vmatrix}.$$

В силу линейного свойства определителей (*1.44*) искомый определитель равен сумме определителя первой матрицы и всех определителей, полученных заменой некоторых столбцов определителя первой матрицы на соответствующие столбцы второй матрицы $\left(\text{с учетом множителя } \frac{\lambda_n-\lambda}{\alpha_n^2}\right)$. Поскольку всякие два столбца второй матрицы пропорциональны, достаточно рассматривать только те случаи, когда один из столбцов определителя первой матрицы заменен на соответствующий столбец второй матрицы.

Если, в частности, k-й столбец первой матрицы заменен на k-й столбец второй матрицы, то соответствующий определитель имеет следующее выражение:

$$\frac{\lambda_n-\lambda}{\alpha_n^2}\begin{vmatrix} \lambda_1-\lambda & 0 & \dots & 0 & \alpha_k\alpha_1 & 0 & \dots & 0 \\ 0 & \lambda_2-\lambda & \dots & 0 & \alpha_k\alpha_2 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_{k-1}-\lambda & \alpha_k\alpha_{k-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \alpha_k\alpha_k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \alpha_k\alpha_{k+1} & \lambda_{k+1}-\lambda & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & \alpha_k\alpha_{n-1} & 0 & \dots & \lambda_{n-1}-\lambda \end{vmatrix} =$$

$$= \frac{\alpha_k^2}{\alpha_n^2}\,\frac{\prod\limits_{j=1}^{n}(\lambda_j-\lambda)}{\lambda_k-\lambda}.$$

Введем обозначения

$$F(\lambda) = \prod_{k=1}^{n-1}(\lambda_k-\lambda) \quad \text{(определитель первой матрицы)},$$

$$G(\lambda) = \prod_{k=1}^{n}(\lambda_k-\lambda).$$

Тогда искомый определитель $D(\lambda)$ примет вид

$$D(\lambda) = F(\lambda) + \frac{1}{\alpha_n^2}G(\lambda)\sum_{k=1}^{n-1}\frac{\alpha_k^2}{\lambda_k-\lambda}. \tag{10}$$

Решая уравнение $D(\lambda)=0$, мы и найдем интересующие нас величины $\mu_1, \mu_2, \dots, \mu_{n-1}$. Заметим, что они зависят не от самих чисел α_j, а от квадратов этих чисел; таким образом, если у одного или нескольких коэффициентов уравнения (8) изменить знак, то искомые канонические коэффициенты формы $A(x, x)$ на подпространстве R_{n-1} не изменяются.

10.27*. Формула (10) интересна еще и тем, что *она позволяет построить по данным числам* $\mu_1, \mu_2, \ldots, \mu_{n-1}$, *удовлетворяющим неравенствам* (8), *подпространство* R_{n-1}, *на котором форма* $\mathrm{A}(x, x)$ *имеет канонические коэффициенты* $\mu_1, \mu_2, \ldots, \mu_{n-1}$ (в предположении, что числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ различны). Покажем, как решается эта задача.

Заметим, что формулу (10) мы можем записать в виде

$$\alpha_n^2 \frac{D(\lambda)}{G(\lambda)} = \alpha_n^2 \frac{E(\lambda)}{G(\lambda)} + \sum_{k=1}^{n-1} \frac{\alpha_k^2}{\lambda_k - \lambda} = \sum_{k=1}^{n} \frac{\alpha_k^2}{\lambda_k - \lambda}; \tag{11}$$

таким образом, числа $\alpha_1^2, \alpha_2^2, \ldots, \alpha_n^2$ пропорциональны коэффициентам разложения рациональной функции $\frac{D(\lambda)}{G(\lambda)}$ на простейшие дроби.

Пусть заданы числа $\mu_1, \mu_2, \ldots, \mu_{n-1}$, удовлетворяющие неравенствам

$$\left.\begin{aligned} \lambda_1 &> \mu_1 > \lambda_2, \\ \lambda_2 &> \mu_2 > \lambda_3, \\ &\ldots\ldots\ldots \\ \lambda_{n-1} &> \mu_{n-1} > \lambda_n. \end{aligned}\right\} \tag{12}$$

Положим $D_1(\lambda) = \prod_{k=1}^{n-1} (\mu_k - \lambda)$ и разложим рациональную функцию $\frac{D_1(\lambda)}{G(\lambda)}$ на простейшие дроби:

$$\frac{D_1(\lambda)}{G(\lambda)} = \frac{c_1}{\lambda_1 - \lambda} + \frac{c_2}{\lambda_2 - \lambda} + \ldots + \frac{c_n}{\lambda_n - \lambda}. \tag{13}$$

Покажем, что коэффициенты $c_1, c_2, \ldots, c_n$ — одного знака. Известно, что эти коэффициенты вычисляются по формулам *)

$$c_k = \frac{D_1(\lambda_k)}{(\lambda_k - \lambda_1)\ldots(\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1})\ldots(\lambda_k - \lambda_n)} = -\frac{D_1(\lambda_k)}{G'(\lambda_k)}.$$

Числа $D_1(\lambda_1), D_1(\lambda_2), \ldots, D_1(\lambda_n)$ имеют попеременно противоположные знаки, поскольку корни многочлена $D_1(\lambda)$ по условию перемежаются с корнями многочлена $G(\lambda)$. Числа $G'(\lambda_1), G'(\lambda_2), \ldots, G'(\lambda_n)$ также имеют попеременно противоположные знаки, поскольку $\lambda_1, \lambda_2, \ldots, \lambda_n$ — простые

*) См., например, М. Гребенча, С. Новоселов, Курс математического анализа, стр. 405, Учпедгиз, 1951.

корни многочлена $G(\lambda)$. Поэтому отношения $\frac{D_1(\lambda_k)}{G'(\lambda_k)}$, а с ними и коэффициенты c_k имеют одинаковые знаки. С точностью до множителя можно все коэффициенты c_k считать положительными, а их сумму — равной единице, и тогда можно определить числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ из условий

$$\alpha_1^2 = c_1, \ \alpha_2^2 = c_2, \ \ldots, \ \alpha_n^2 = c_n. \tag{14}$$

Числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ можно взять любого знака. Покажем, что подпространство R_{n-1}, определяемое уравнением

$$\alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_n\xi_n = 0,$$

и будет искомым.

Действительно, многочлен $D(\lambda)$, корни которого суть канонические коэффициенты формы $\mathrm{A}(x, x)$ на подпространстве R_{n-1}, по доказанному, выражается с помощью формулы (10) или эквивалентной ей (11). Сравнивая (11) и (12) и учитывая (14), мы получаем, что многочлен $D(\lambda)$ отличается только числовым множителем от построенного нами многочлена $D_1(\lambda)$. Но тогда корни многочлена $D(\lambda)$ совпадают с числами $\mu_1, \mu_2, \ldots, \mu_{n-1}$, что и утверждалось.

Замечание. Можно показать, что полученные числа $\alpha_1, \ldots, \alpha_n$ зависят непрерывно от величин $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_{n-1}$. Используя этот факт, можно проверить, что задача имеет решение и для чисел $\mu_1, \ldots, \mu_{n-1}$, удовлетворяющих неравенствам (7), а также и без предположения, что числа $\lambda_1, \ldots, \lambda_n$ различны.

§ 10.3. Задача о паре квадратичных форм

10.31. В некоторых вопросах математики и физики существенную роль играет решение следующей задачи: *для двух квадратичных форм* $\mathrm{A}(x, x)$ *и* $\mathrm{B}(x, x)$, *заданных в n-мерном аффинном пространстве* R_n, *указать базис, в котором обе эти формы записываются в каноническом виде* (т. е. в виде суммы квадратов координат с некоторыми коэффициентами).

Следующий пример на плоскости ($n=2$) показывает, что эта задача не всегда допускает решение.

Рассмотрим следующие две формы от двух переменных ξ_1, ξ_2:

$$\mathrm{A}(x, x) = \xi_1^2 - \xi_2^2,$$
$$\mathrm{B}(x, x) = \xi_1\xi_2.$$

Найти общий канонический базис для этих форм означает найти общую пару взаимно сопряженных векторов для гипербол $A(x, x)=1$ и $B(x, x)=1$ (см. *7.44*). Эти гиперболы равносторонние; из аналитической геометрии известно, что сопряженные направления таких гипербол симметричны относительно их асимптот. Поэтому полярные углы φ_1 и φ_2, отвечающие паре сопряженных направлений, для первой гиперболы связаны равенством

$$\varphi_1+\varphi_2=\pi/2,$$

а для второй гиперболы — равенством

$$\varphi_1+\varphi_2=0$$

(оба равенства с точностью до слагаемого, кратного π).

Так как эти равенства исключают одно другое, то в данном случае общих взаимно сопряженных направлений не существует.

Оказывается, что задача имеет решение, если дополнительно допустить, что *одна из этих форм, например* $B(x, x)$, *положительно определенная* (т. е. $B(x, x)>0$ при $x\neq 0$).

Существование решения легко установить следующим путем. Пусть $B(x, y)$ — симметричная билинейная форма, соответствующая квадратичной форме $B(x, x)$. Введем в аффинном пространстве R_n евклидову метрику, полагая

$$(x, y)=B(x, y).$$

Выполнение аксиом скалярного произведения обеспечивается симметричностью и положительной определенностью формы $B(x, x)$.

В силу *10.11* существует ортогональный и нормированный (относительно введенной нами метрики) базис $e_1, \ldots, e_n$, в котором форма $A(x, x)$ принимает канонический вид

$$A(x, x)=\lambda_1\eta_1^2+\lambda_2\eta_2^2+\ldots+\lambda_n\eta_n^2 \qquad (15)$$

($\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора x в построенном базисе). В этом же базисе вторая квадратичная форма $B(x, x)$ в силу формулы *8.42* (17) имеет вид

$$B(x, x)=(x, x)=\eta_1^2+\eta_2^2+\ldots+\eta_n^2.$$

Итак, базис, в котором обе формы имеют канонический вид, существует.

10.32. Используем для вычисления координат векторов $e_1, \ldots, e_n$ искомого базиса экстремальные свойства квадратичных форм. Как было показано в *10.21*, векторы $e_1, \ldots, e_n$

суть те векторы, которые подчинены условию

$$(x, x) = \mathrm{B}(x, x) = 1$$

и для которых форма $\mathrm{A}(x, x)$ принимает стационарные значения. Предположим, что в исходном базисе формы $\mathrm{A}(x, x)$ и $\mathrm{B}(x, x)$ имели следующие выражения:

$$\mathrm{A}(x, x) = \sum_{i,\, k=1}^{n} a_{ik}\xi_i\xi_k, \quad \mathrm{B}(x, x) = \sum_{i,\, k=1}^{n} b_{ik}\xi_i\xi_k.$$

Действуя по методу Лагранжа, мы должны составить функцию

$$F(\xi_1, \xi_2, \ldots, \xi_n) = \sum_{i,\, k=1}^{n} a_{ik}\xi_i\xi_k - \mu \sum_{i,\, k=1}^{n} b_{ik}\xi_i\xi_k$$

и приравнять нулю ее частные производные по всем координатам:

$$\sum_{k=1}^{n} a_{ik}\xi_k - \mu \sum_{k=1}^{n} b_{ik}\xi_k = 0 \qquad (i = 1, 2, \ldots, n). \tag{16}$$

Полученная система однородных уравнений

$$\left.\begin{array}{l} (a_{11}-\mu b_{11})\,\xi_1 + (a_{12}-\mu b_{12})\,\xi_2 + \ldots + (a_{1n}-\mu b_{1n})\,\xi_n = 0, \\ (a_{21}-\mu b_{21})\,\xi_1 + (a_{22}-\mu b_{22})\,\xi_2 + \ldots + (a_{2n}-\mu b_{2n})\,\xi_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (a_{n1}-\mu b_{n1})\,\xi_1 + (a_{n2}-\mu b_{n2})\,\xi_2 + \ldots + (a_{nn}-\mu b_{nn})\,\xi_n = 0 \end{array}\right\} \tag{17}$$

допускает ненулевые решения тогда и только тогда, когда ее определитель обращается в нуль:

$$\begin{vmatrix} a_{11}-\mu b_{11} & a_{12}-\mu b_{12} & \ldots & a_{1n}-\mu b_{1n} \\ a_{21}-\mu b_{21} & a_{22}-\mu b_{22} & \ldots & a_{2n}-\mu b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1}-\mu b_{n1} & a_{n2}-\mu b_{n2} & \ldots & a_{nn}-\mu b_{nn} \end{vmatrix} = 0. \tag{18}$$

Решая уравнение (18), мы находим n возможных значений $\mu = \mu_k$ $(k = 1, 2, \ldots, n)$; подставляя μ_k в систему (17), мы сможем найти координаты $\xi_1^{(k)}, \xi_2^{(k)}, \ldots, \xi_n^{(k)}$ соответствующего искомого базисного вектора. Теорема, доказанная в *10.31*, обеспечивает существование вещественных корней определителя системы (17) и для каждого кратного корня наличие соответствующего числа линейно независимых решений этой системы.

10.33. Переходим к вычислению канонических коэффициентов. Покажем, что коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_n$ в канонической записи (15) формы $A(x, x)$ совпадают с соответствующими корнями $\mu_1, \mu_2, \ldots, \mu_n$ определителя системы (17). Здесь можно было бы использовать рассуждение, аналогичное проведенному в *10.22*; мы теперь предпочитаем провести непосредственное вычисление. Если для заданного корня μ_m умножить i-е уравнение системы (16) на i-ю координату решения $\xi_i^{(m)}$ $(i=1, 2, \ldots, n)$ и все эти уравнения сложить, то получим равенство

$$A(e_m, e_m) = \sum_{i, k=1}^{n} a_{ik}\xi_i^{(m)}\xi_k^{(m)} =$$

$$= \mu_m \sum_{i, k=1}^{n} b_{ik}\xi_i^{(m)}\xi_k^{(m)} = \mu_m B(e_m, e_m) = \mu_m,$$

так как $B(e_m, e_m) = 1$.

С другой стороны, канонические координаты $\eta_1, \eta_2, \ldots, \eta_m$ для вектора e_m имеют, очевидно, значения $\eta_i = 0$ при $i \neq m$, $\eta_m = 1$, и форма $A(x, x) = \sum_{i=1}^{n} \lambda_i \eta_i^2$ при $x = e_m$ становится равной λ_m. Отсюда $\mu_m = \lambda_m$, что и утверждалось.

Этот результат дает возможность написать форму $A(x, x)$ в искомом каноническом виде, минуя вычисление канонического базиса.

10.34. Поставленная в *10.31* задача об одновременном приведении к каноническому виду двух квадратичных форм $A(x, x)$ и $B(x, x)$, из которых, например, $B(x, x)$ положительно определена, решена нами в несколько усиленной форме; именно, форма $B(x, x)$ приведена к виду суммы квадратов координат с коэффициентами, равными 1. Вообще говоря, это не требуется, и поэтому коэффициенты преобразованных форм заведомо не определяются однозначно. Мы покажем все же, что *отношения соответствующих канонических коэффициентов не зависят от способа одновременного приведения форм $A(x, x)$ и $B(x, x)$ к каноническому виду.*

Пусть формы $A(x, x)$ и $B(x, x)$ двумя способами приведены к каноническому виду: в координатах $\xi_1, \xi_2, \ldots, \xi_n$

$$A(x, x) = \sum_{i=1}^{n} \lambda_i \xi_i^2, \qquad B(x, x) = \sum_{i=1}^{n} \nu_i \xi_i^2,$$

а в координатах $\eta_1, \eta_2, \ldots, \eta_n$

$$A(x, x) = \sum_{i=1}^{n} \rho_i \eta_i^2, \quad B(x, x) = \sum_{i=1}^{n} \tau_i \eta_i^2.$$

Так как форма $B(x, x)$ положительно определенная, числа ν_i и τ_i $(i = 1, 2, \ldots, n)$ все положительны. Рассмотрим новое преобразование координат

$$\sqrt{\nu_i}\, \xi_i = \overline{\xi}_i, \quad \sqrt{\tau_i}\, \eta_i = \overline{\eta}_i.$$

Тогда формы $A(x, x)$ и $B(x, x)$ преобразуются к виду

а) в координатах $\overline{\xi}_i$:

$$A(x, x) = \sum_{i=1}^{n} \frac{\lambda_i}{\nu_i} \overline{\xi}_i^2, \quad B(x, x) = \sum_{i=1}^{n} \overline{\xi}_i^2;$$

б) в координатах $\overline{\eta}_i$:

$$A(x, x) = \sum_{i=1}^{n} \frac{\rho_i}{\tau_i} \overline{\eta}_i^2, \quad B(x, x) = \sum_{i=1}^{n} \overline{\eta}_i^2.$$

Пусть $e_1, e_2, \ldots, e_n$ — базис, отвечающий координатам $\overline{\xi}_i$, и $f_1, f_2, \ldots, f_n$ — базис, отвечающий координатам $\overline{\eta}_i$. В метрике, определяемой формой $B(x, x)$, оба эти базиса ортогональны и нормированы. Но тогда (*10.14*) совокупность канонических коэффициентов формы $A(x, x)$ определена однозначно; таким образом, последовательность чисел $\frac{\lambda_1}{\nu_1}, \frac{\lambda_2}{\nu_2}, \ldots, \frac{\lambda_n}{\nu_n}$ должна совпадать с последовательностью $\frac{\rho_1}{\tau_1}, \frac{\rho_2}{\tau_2}, \ldots, \frac{\rho_n}{\tau_n}$ с точностью до порядка.

Теорема доказана.

§ 10.4. Приведение общего уравнения поверхности 2-го порядка к каноническому виду

10.41. В этом и следующих пунктах мы будем называть элементы линейного пространства R_n не векторами, а точками (*2.17*), что будет более соответствовать геометрическому представлению. *Поверхностью 2-го порядка в n-мерном пространстве мы будем называть геометрическое место точек*

$x=(\xi_1, \xi_2, \ldots, \xi_n)$, *удовлетворяющих уравнению вида*

$$\sum_{i,\,k=1}^{n} a_{ik}\xi_i\xi_k+2\sum_{i=1}^{n} b_i\xi_i+c=0, \tag{19}$$

или

$$\mathrm{A}(x,\ x)+2L(x)+c=0,$$

где $\mathrm{A}(x,\ x)=\sum_{i,\,k=1}^{n} a_{ik}\xi_i\xi_k$ — *квадратичная форма от радиуса-вектора точки* x, $L(x)=\sum_{i=1}^{n} b_i\xi_i$ — *линейная форма,* c — *постоянная* *).

Пространство R_n будем считать евклидовым и числа $\xi_1, \xi_2, \ldots, \xi_n$ — координатами вектора x в ортогональном и нормированном базисе. Задачей настоящего пункта является выбор в пространстве R_n нового ортогонального и нормированного базиса и нового начала координат так, чтобы наша поверхность 2-го порядка определялась некоторым специальным и особенно простым уравнением, которое называется *каноническим*. В дальнейшем по каноническому уравнению мы изучим свойства поверхности.

10.42. Совершим прежде всего в пространстве R_n ортогональное преобразование координат

$$\xi_i=\sum_{j=1}^{n} q_{ij}\eta_j \qquad (i=1,\ 2,\ \ldots,\ n), \tag{20}$$

как указано в *10.12*, с тем чтобы в новых переменных квадратичная форма $\mathrm{A}(x, x)$ приняла канонический вид

$$\mathrm{A}(x,\ x)=\sum_{i=1}^{n} \lambda_i\eta_i^2.$$

Уравнение (19) будет после подстановки (20) иметь вид

$$\sum_{i=1}^{n} \lambda_i\eta_i^2+2\sum_{i=1}^{n} l_i\eta_i+c=0, \tag{21}$$

---

*) В случае $n=2$ геометрический образ, определяемый уравнением (19), называется кривой 2-го порядка. Однако в дальнейшем мы всюду употребляем слово «поверхность», не оговаривая каждый раз, что при $n=2$ его нужно заменить словом «кривая».

где l_i $(i=1, 2, \ldots, n)$ — новые коэффициенты линейной формы $L(x)$.

Если в полученном уравнении $\lambda_i \neq 0$ для некоторого i, то переносом начала координат можно добиться исчезновения соответствующего члена первого измерения. Пусть, например, $\lambda_1 \neq 0$; тогда, очевидно,

$$\lambda_1 \eta_1^2 + 2l_1\eta_1 = \lambda_1\left(\eta_1 + \frac{l_1}{\lambda_1}\right)^2 - \frac{l_1^2}{\lambda_1}.$$

Положим $\eta_1' = \eta_1 + \frac{l_1}{\lambda_1}$; это равносильно переносу начала координат в точку $\left(-\frac{l_1}{\lambda_1}, 0, 0, \ldots, 0\right)$. В результате подстановки группа членов $\lambda_1\eta_1^2 + 2l_1\eta_1$ заменится на $\lambda_1\eta_1'^2 - \frac{l_1^2}{\lambda_1}$; таким образом, член второго измерения останется с тем же самым коэффициентом λ_1, член первого измерения пропадает, свободный член получит добавок $-\frac{l_1^2}{\lambda_1}$. После всех таких преобразований уравнение поверхности примет вид

$$\lambda_1\eta_1^2 + \lambda_2\eta_2^2 + \ldots + \lambda_r\eta_r^2 + 2l_{r+1}\eta_{r+1} + \ldots + 2l_n\eta_n + c = 0.$$

Здесь опущены для простоты записи штрихи у координат, а сами координаты заново перенумерованы так, чтобы вначале шли координаты, участвующие в квадратичной форме, так что $\lambda_1, \lambda_2, \ldots, \lambda_r$ отличны от нуля, $\lambda_k = 0$ при $k > r$. Если при этом $r = n$ или же $r < n$, но числа $l_{r+1}, l_{r+2}, \ldots, l_n$ оказались равными нулю, то мы получаем *каноническое уравнение центральной поверхности*

$$\lambda_1\eta_1^2 + \lambda_2\eta_2^2 + \ldots + \lambda_r\eta_r^2 + c = 0. \qquad (22)$$

Если $r = n$, то эта поверхность при $c \neq 0$ называется *истинной*, а при $c = 0$ — *конической*.

Допустим, что среди чисел $l_{r+1}, \ldots, l_n$ имеется хотя бы одно отличное от нуля. Тогда мы совершим новое ортогональное преобразование координат по формулам

$$\begin{aligned} \tau_1 &= \eta_1, \\ \tau_2 &= \eta_2, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \tau_r &= \eta_r, \\ \tau_{r+1} &= -(l_{r+1}\eta_{r+1} + \ldots + l_n\eta_n)\frac{1}{M}, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{aligned}$$

где M — положительный множитель, обеспечивающий ортогональность матрицы преобразования: так как у ортогональной матрицы сумма квадратов элементов каждой строки должна быть равна 1, то

$$M^2 = l_{r+1}^2 + l_{r+2}^2 + \ldots + l_n^2.$$

Остальные строки (следующие за $(r+1)$-й) могут быть произвольными, лишь бы полученная матрица была ортогональной (*8.95*).

В результате этого преобразования уравнение поверхности приобретает вид

$$\lambda_1\tau_1^2 + \ldots + \lambda_r\tau_r^2 = 2M\tau_{r+1} - c.$$

Если $c \neq 0$, еще один перенос начала координат по формуле

$$\tau'_{r+1} = \tau_{r+1} - \frac{c}{2M}.$$

позволяет освободиться от свободного члена; уравнение (опять с опущенным штрихом у последней координаты) получает вид

$$\lambda_1\tau_1^2 + \ldots + \lambda_r\tau_r^2 = 2M\tau_{r+1}; \tag{23}$$

это — *каноническое уравнение нецентральной поверхности.*

Всякую поверхность 2-го порядка будем называть *невырожденной,* если в ее каноническом уравнении участвуют все n координат, и *вырожденной,* если в ее каноническом уравнении участвует менее чем n координат. Все введенные названия будут разъяснены в дальнейшем.

§ 10.5. Геометрические свойства поверхностей 2-го порядка

10.51. Центр поверхности. *Центром* поверхности называется точка $x_0 = (\xi_1^0, \xi_2^0, \ldots, \xi_n^0)$, обладающая следующим свойством: если точка $(\xi_1^0 + \xi_1, \xi_2^0 + \xi_2, \ldots, \xi_n^0 + \xi_n)$ лежит на поверхности, то симметричная с ней относительно x_0 точка $(\xi_1^0 - \xi_1, \xi_2^0 - \xi_2, \ldots, \xi_n^0 - \xi_n)$ также лежит на поверхности. У поверхности с каноническим уравнением (22) существуют центры; всякая точка, для которой $\eta_1 = \eta_2 = \ldots = \eta_r = 0$, является, очевидно, центром. Этим объясняется название этого класса: центральные поверхности.

Покажем (это будет использовано в дальнейшем), что *никаких других центров у поверхности с уравнением* (22) *не существует*. Действительно, пусть $(\xi_1^0, \ldots, \xi_n^0)$ — центр этой поверхности. Тогда из условия

$$\lambda_1 (\xi_1^0 + \xi_1)^2 + \lambda_2 (\xi_2^0 + \xi_2)^2 + \ldots + \lambda_r (\xi_r^0 + \xi_r)^2 + c = 0$$

вытекает

$$\lambda_1 (\xi_1^0 - \xi_1)^2 + \lambda_2 (\xi_2^0 - \xi_2)^2 + \ldots + \lambda_r (\xi_r^0 - \xi_r)^2 + c = 0.$$

Вычитая второе равенство из первого, мы получаем

$$\lambda_1 \xi_1^0 \xi_1 + \lambda_2 \xi_2^0 \xi_2 + \ldots + \lambda_r \xi_r^0 \xi_r = 0.$$

Возьмем на поверхности точку, в которой $\xi_2 = \xi_3 = \ldots = \xi_r = 0$, $\xi_1 \neq 0$. (Очевидно, что уравнению (22) можно удовлетворить такими значениями.) Тогда мы получим

$$\lambda_1 \xi_1^0 \xi_1 = 0,$$

откуда $\xi_1^0 = 0$. Аналогично показываем, что $\xi_2^0 = \ldots = \xi_r^0 = 0$, что нам и требуется.

10.52. Истинные центральные поверхности. Рассмотрим сначала истинную центральную поверхность, т. е. предположим, что $r = n$ и $c \neq 0$. Тогда уравнение (22) легко преобразовать к виду

$$\pm \frac{\eta_1^2}{a_1^2} \pm \frac{\eta_2^2}{a_2^2} \pm \ldots \pm \frac{\eta_n^2}{a_n^2} = 1,$$

где числа a_k определяются формулами

$$a_k = +\sqrt{\left|\frac{c}{\lambda_k}\right|} \qquad (k = 1, 2, \ldots, n);$$

они называются *полуосями* поверхности.

Перенумеруем координаты заново так, чтобы сначала шли слагаемые с положительными знаками:

$$\frac{\eta_1^2}{a_1^2} + \frac{\eta_2^2}{a_2^2} + \ldots + \frac{\eta_k^2}{a_k^2} - \frac{\eta_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{\eta_n^2}{a_n^2} = 1. \qquad (24)$$

Случай $k = 0$ естественно исключить из рассмотрения, так как при $k = 0$ никакие вещественные значения $\eta_1, \eta_2, \ldots, \eta_n$ не могут удовлетворить уравнению (24); в этом случае говорят иногда, что уравнение (24) определяет «мнимую» поверхность.

Остаю́тся n различных типов истинных центральных поверхностей, отвечающих значениям $k = 1, 2, \ldots, n$.

а. В двумерном случае ($n = 2$) уравнение (24) определяет при $k = 1$ и $k = 2$ две известные из аналитической геометрии кривые:

$$k = 1: \qquad \frac{\eta_1^2}{a_1^2} - \frac{\eta_2^2}{a_2^2} = 1 \ \text{(гипербола)},$$

$$k = 2: \qquad \frac{\eta_1^2}{a_1^2} + \frac{\eta_2^2}{a_2^2} = 1 \ \text{(эллипс)}.$$

б. При $n = 3$ имеем $k = 1$, $k = 2$, $k = 3$ и соответственно три невырожденные центральные поверхности в трехмерном пространстве, определяемые следующими уравнениями:

$$k = 1: \qquad \frac{\eta_1^2}{a_1^2} - \frac{\eta_2^2}{a_2^2} - \frac{\eta_3^2}{a_3^2} = 1 \ \text{(двуполостный гиперболоид)},$$

$$k = 2: \qquad \frac{\eta_1^2}{a_1^2} + \frac{\eta_2^2}{a_2^2} - \frac{\eta_3^2}{a_3^2} = 1 \ \text{(однополостный гиперболоид)},$$

$$k = 3: \qquad \frac{\eta_1^2}{a_1^2} + \frac{\eta_2^2}{a_2^2} + \frac{\eta_3^2}{a_3^2} = 1 \ \text{(эллипсоид)}.$$

Напомним читателю построение каждой из этих поверхностей.

Рассмотрим сечения каждой из них горизонтальными плоскостями $\eta_3 = ca_3$ ($-\infty < c < \infty$). Эти сечения представляют собой соответственно: гиперболы с вещественной осью η_1:

$$k=1: \qquad \frac{\eta_1^2}{a_1^2}-\frac{\eta_2^2}{a_2^2}=1+c^2;$$

эллипсы, определенные для всех значений c:

$$k=2: \qquad \frac{\eta_1^2}{a_1^2}+\frac{\eta_2^2}{a_2^2}=1+c^2;$$

эллипсы, определенные только для $|c|\leqslant 1$:

$$k=3: \qquad \frac{\eta_1^2}{a_1^2}+\frac{\eta_2^2}{a_2^2}=1-c^2.$$

Чтобы определить положения вершин этих сечений, построим сечения поверхности координатными плоскостями $\eta_1=0$,

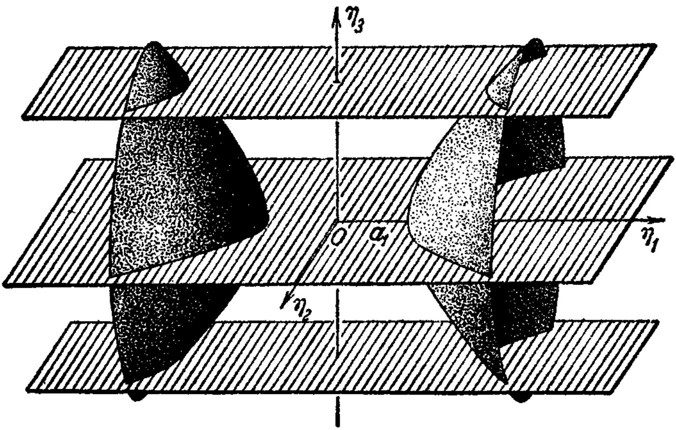

Рис. 2.

$\eta_2=0$. Для случая $k=1$ мы получаем при этом действительное сечение только координатной плоскостью $\eta_2=0$, которое будет представлять собой гиперболу:

$$\frac{\eta_1^2}{a_1^2}-\frac{\eta_3^2}{a_3^2}=1.$$

Вершины гипербол горизонтальных сечений будут располагаться на этой кривой; в результате построения получаем двуполостный гиперболоид (рис. 2).

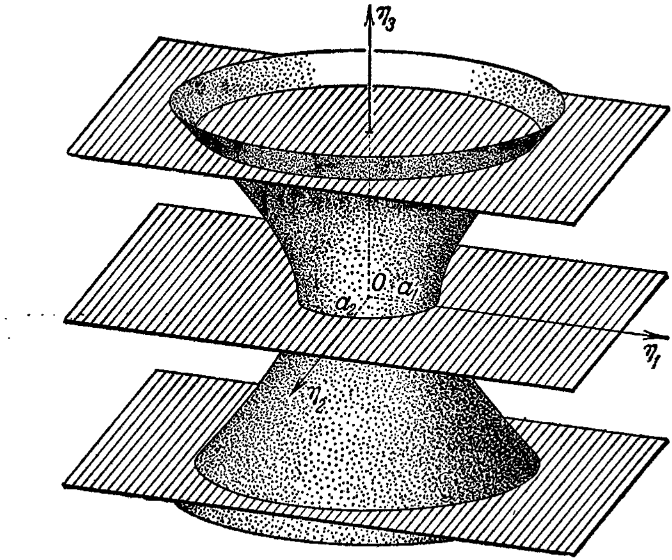

Рис 3.

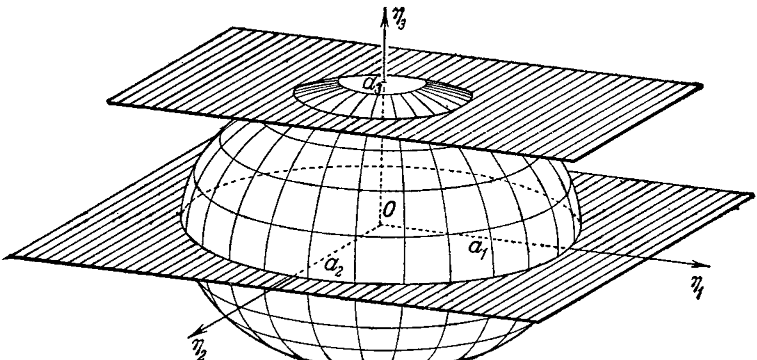

Рис. 4.

Для случая $k=2$ сечения обеими координатными плоскостями $\eta_1=0$, $\eta_2=0$ представляют собой гиперболы с мнимой осью η_3:

$$\frac{\eta_2^2}{a_2^2}-\frac{\eta_3^2}{a_3^2}=1, \qquad \frac{\eta_1^2}{a_1^2}-\frac{\eta_3^2}{a_3^2}=1.$$

Совокупность эллипсов горизонтальных сечений с вершинами на этих гиперболах и составляет однополостный гиперболоид (рис. 3). Наконец, в случае $k=3$ сечения координатными плоскостями $\eta_1=0$, $\eta_2=0$ — эллипсы; проводя эллипсы горизонтальных сечений, получаем эллипсоид (рис. 4).

в. Поверхности 2-го порядка в пространстве более чем трех измерений уже не поддаются наглядному геометрическому представлению. Тем не менее мы можем указать и в многомерном случае существенные различия между типами истинных поверхностей, отвечающих различным значениям $k=1, 2, \ldots, n$. Будем исходить из геометрически очевидных различий в трехмерном пространстве. На двуполостном гиперболоиде ($k=1$) существует пара точек, которые нельзя путем непрерывного передвижения по поверхности привести к совпадению: достаточно взять одну из точек пары на одной полости, а вторую точку — на другой полости, чтобы получить такую пару. На однополостном гиперболоиде ($k=2$) уже всякие две точки можно привести к совпадению с помощью непрерывного передвижения по поверхности; но есть замкнутая линия (например, горловая линия гиперболоида), которую нельзя непрерывной деформацией свести в одну точку. На эллипсоиде ($k=2$) уже всякая замкнутая линия может быть сведена в одну точку. Эти факты могут служить исходным пунктом при формулировке геометрических различий между центральными поверхностями в n-мерном пространстве.

Введем следующие определения. Фигура A называется *гомеоморфной* фигуре B, если существует взаимно однозначное и взаимно непрерывное отображение множества точек фигуры A на множество точек фигуры B. Фигура A, расположенная на поверхности S, называется *гомотопной* фигуре B, расположенной на этой же поверхности, если фигура A может быть переведена в фигуру B с помощью непрерывной деформации, в процессе которой фигура A остается все время на поверхности S.

Геометрические различия между центральными поверхностями с помощью этих определений формулируются следующим образом. Для $k=1$ можно указать на поверхности пару точек, не гомотопных друг другу. Для $k=2$ всякая точка на поверхности гомотопна всякой другой точке; но существует линия, гомеоморфная окружности, которая не гомотопна точке. Для $k=3$ всякая линия, гомеоморфная окружности, гомотопна точке; но существует не гомотопная точке часть поверхности, гомеоморфная сфере (точнее, двумерной сфере, т. е. сфере в трехмерном пространстве). Продолжая таким образом, мы сможем сформулировать для каждого k отличительное свойство соответствующей центральной поверхности: всякая ее часть, гомеоморфная $(k-1)$-мерной сфере, гомотопна точке, но существует часть, гомеоморфная k-мерной сфере, которая не гомотопна точке. Из этого результата, в частности, вытекает, что центральные поверхности в n-мерном пространстве, очевидно, гомеоморфные друг другу при равных значениях k, не гомеоморфны друг другу при различных k. На доказательствах этих интересных предложений мы останавливаться не можем *).

10.53. Конические поверхности. Рассмотрим теперь случай конической поверхности: в уравнении (22) $c=0$. Уравнение (22) становится однородным: вместе с точкой $(\eta_1, \eta_2, \ldots, \eta_n)$ ему удовлетворяет и точка $(t\eta_1, t\eta_2, \ldots, t\eta_n)$ при любом t. Это означает, что поверхность образована из прямых линий, проходящих через начало координат **). Аналогично предыдущему каноническое уравнение конической поверхности мы сможем записать в виде

$$\frac{\eta_1^2}{a_1^2}+\ldots+\frac{\eta_k^2}{a_k^2}-\frac{\eta_{k+1}^2}{a_{k+1}^2}-\ldots-\frac{\eta_n^2}{a_n^2}=0.$$

Оценим число различных типов конических поверхностей при заданном n. Если число $m=n-k$ отрицательных коэффициентов в каноническом уравнении (22) больше $n/2$, то, умножая уравнение на -1, мы получаем уравнение той же поверхности, но с числом отрицательных коэффициентов, уже меньшим $n/2$. Следовательно, достаточно рассмотреть случаи, отвечающие значениям $m \leqslant n/2$. Если n четное, то, исключая случай точки $(k=0)$, получаем $n/2$ различных типов конических поверхностей, отвечающих значениям $m=1, 2, \ldots, n/2$; если n нечетное, то различных типов оказывается $(n-1)/2$, именно, они отвечают значениям $m=1, 2, \ldots, (n-1)/2$.

а. На плоскости $(n=2)$, кроме точки, имеется один такой тип $(m=1)$ с каноническим уравнением

$$\frac{\eta_1^2}{a_1^2}-\frac{\eta_2^2}{a_2^2}=0.$$

Соответствующий геометрический образ — пара пересекающихся прямых с уравнениями $\frac{\eta_1}{a_1}=\pm\frac{\eta_2}{a_2}$.

---

*) См. Зейферт и Трелфал, Топология, ГОНТИ, 1938.

**) За единственным исключением, когда все слагаемые в сумме (22) одного знака и уравнение определяет одну точку — начало координат.

б. В трехмерном пространстве ($n=3$), кроме точки, имеется также только один тип конической поверхности $\left(\frac{n-1}{2}=\frac{3-1}{2}=1\right)$ с каноническим уравнением

$$\frac{\eta_1^2}{a_1^2}+\frac{\eta_2^2}{a_2^2}-\frac{\eta_3^2}{a_3^2}=0.$$

Соответствующий геометрический образ — конус (рис. 5; в частном случае, при $a_1=a_2$, прямой круговой конус).

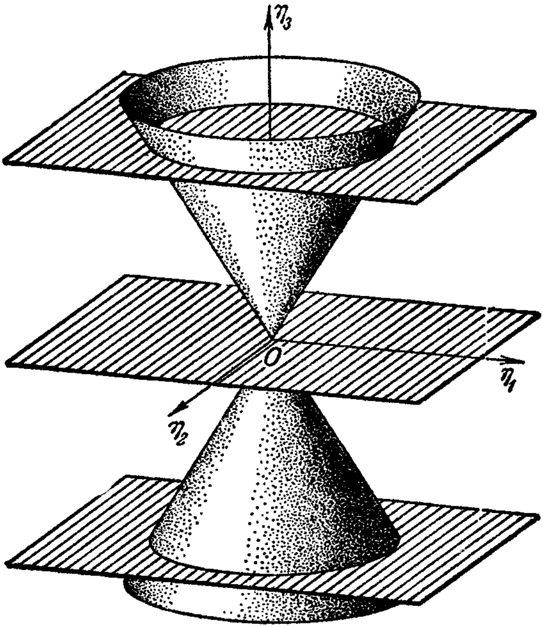

Рис. 5.

в. Чтобы представить себе форму конической поверхности в общем случае, рассмотрим ее сечение гиперплоскостью

$$\eta_n = ca_n \qquad (-\infty < c < \infty)$$

$$\frac{\eta_1^2}{a_1^2}+\ldots+\frac{\eta_k^2}{a_k^2}-\frac{\eta_{k+1}^2}{a_{k+1}^2}-\ldots-\frac{\eta_{n-1}^2}{a_{n-1}^2}=c^2.$$

Это уравнение соответствует центральной поверхности в $(n-1)$-мерном пространстве. Все эти поверхности (при различных значениях c) геометрически подобны друг другу, соответствующие размеры полуосей пропорциональны величине c.

Таким образом, каждая коническая поверхность в n-мерном пространстве может быть получена из некоторой центральной поверхности в $(n-1)$-мерном пространстве R_{n-1} при помощи перемещения этой центральной поверхности вдоль оси, перпендикулярной к R_{n-1}, с одновременным пропорциональным растяжением во всех направлениях. Чтобы получить при этом все возможные типы конических поверхностей, достаточно использовать лишь те центральные поверхности в $(n-1)$-мерном пространстве, для которых число отрицательных слагаемых в каноническом уравнении не превосходит $(n-1)/2$.

10.54. Невырожденные нецентральные поверхности (параболоиды). Тем же путем, как и в *10.52*, мы можем привести каноническое уравнение нецентральной невырожденной поверхности к виду

$$\frac{\eta_1^2}{a_1^2}+\ldots+\frac{\eta_k^2}{a_k^2}-\frac{\eta_{k+1}^2}{a_{k+1}^2}-\ldots-\frac{\eta_{n-1}^2}{a_{n-1}^2}=2\eta_n. \qquad (25)$$

Оценим число различных типов невырожденных нецентральных поверхностей. Если число отрицательных слагаемых в левой части уравнения (25) больше $(n-1)/2$, то, умножая уравнение (25) на -1, мы получаем уравнение той же поверхности, но с числом отрицательных слагаемых в левой части, меньшим $(n-1)/2$, и с измененным знаком правой части. После зеркального отражения $\eta_n'=-\eta_n$ знак у правой части восстанавливается. Таким образом, число различных типов невырожденных нецентральных поверхностей (если не причислять к различным типам поверхности, получающиеся друг из друга зеркальным отражением) определяется количеством целых чисел m, удовлетворяющих неравенству $0\leqslant m\leqslant(n-1)/2$; это количество равно $n/2$ при четном n и $(n+1)/2$ при нечетном n.

а. На плоскости $(n=2)$ существует единственная невырожденная нецентральная кривая (парабола) с каноническим уравнением

$$\eta_1^2 = 2a_1^2\eta_2 \qquad (m=0).$$

б. В трехмерном пространстве имеются две невырожденные нецентральные поверхности $\left(n=3;\ \frac{n+1}{2}=2\right)$:

1) $\frac{\eta_1^2}{a_1^2}+\frac{\eta_2^2}{a_2^2}=2\eta_3$ $(m=0)$ (эллиптический параболоид),

2) $\frac{\eta_1^2}{a_1^2}-\frac{\eta_2^2}{a_2^2}=2\eta_3$ $(m=1)$ (гиперболический параболоид).

В первом случае сечение поверхности плоскостью $\eta_3 = C > 0$ представляет собой эллипс; чтобы определить положение вершин эллипса, построим сечения поверхности координатными плоскостями $\eta_1=0$ и $\eta_2=0$. В каждом из этих сечений мы получим параболу; следы этих парабол на плоскости $\eta_3=C$ укажут положение вершин эллипса.

Получающаяся поверхность (рис. 6) и есть эллиптический параболоид (в частном случае, при $a_1=a_2$, круговой параболоид).

Во втором случае сечение поверхности плоскостью $\eta_3 = C > 0$ представляет собой гиперболу с вещественной осью η_1. Чтобы определить положение вершин, рассмотрим сечение поверхности координатной плоскостью $\eta_2=0$; в сечении получится парабола $\eta_1^2=2a_1^2\eta_3$, след которой на плоскости $\eta_3=C$ укажет положение вершин гиперболы. Сечение плоскостью $\eta_3=C<0$ представляет собой гиперболу с вещественной осью η_2; вершины этой гиперболы лежат на параболе $\eta_2^2=-2a_2^2\eta_3$ в плоскости $\eta_1=0$.

В сечении $\eta_3=0$ получаем пару прямых, которые служат асимптотами для проекций на плоскость $\eta_3=0$ всех рассмотренных нами гипербол в горизонтальных сечениях поверхности. Сама поверхность и есть гиперболический параболоид (рис. 7).

в. Чтобы представить себе форму поверхности (25) в общем случае, будем следить за изменением формы ее сечения гиперплоскостью $\eta_n=C$ при изменении C от 0 до $+\infty$. В каждом таком сечении получается центральная поверхность в $(n-1)$-мерном пространстве. Все эти поверхности подобны друг другу; соответствующие размеры полуосей (в отличие от конической поверхности) изменяются по параболическому закону (пропорционально корню квадратному из C). При $C=0$ центральная поверхность становится конической. При $C<0$ центральная поверхность переходит в *сопряженную поверхность* (т. е. коэффициенты канонического уравнения с отрицательными и положительными знаками меняются ролями). В частном случае, когда все коэффициенты уравнения (25) одного знака, — для определенности положительного, — поверхность существует только в полупространстве $\eta_n \geqslant 0$.

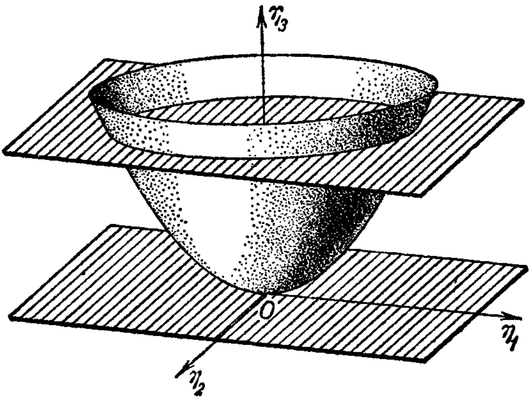

Рис. 6.

Название класса рассматриваемых невырожденных поверхностей объясняется тем, что они действительно не обладают центром. При $n=3$ это очевидно из рис. 6 и 7. В общем случае для доказательства допустим противное: пусть поверхность (25) имеет центр $(\eta_1^0, \eta_2^0, \ldots, \eta_n^0)$. Так как этот центр должен быть, в частности, центром симметрии сечения $\eta_n=\eta_n^0$, представляющего собой невырожденную центральную поверхность в $(n-1)$-мерном пространстве, то необходимо

$$\eta_1^0=\eta_2^0=\ldots=\eta_{n-1}^0=0.$$

Таким образом, центр должен находиться на оси η_n. Перейдем из произвольной точки $(\eta_1, \eta_2, \ldots, \eta_{n-1}, \eta_n^0+\delta)$, лежащей на поверхности, в симметричную точку $(-\eta_1, \ldots$ $\ldots, -\eta_{n-1}, \eta_n^0-\delta)$. При этом уравнение (25) не должно нарушиться. Но левая его часть остается неизменной при указанном переходе; следовательно, не меняется и правая часть, откуда вытекает, что $\delta=0$. Мы получаем, что на

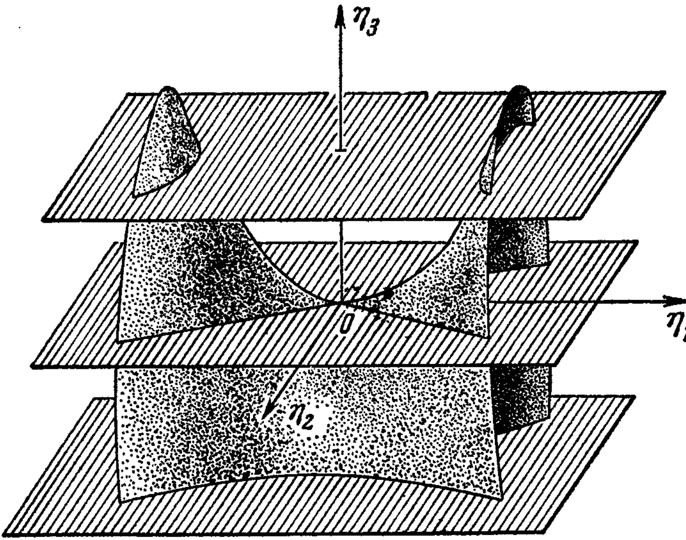

Рис. 7.

поверхности вовсе нет точек с $\eta_n \neq \eta_n^0$. Но, очевидно, уравнение (25) допускает решение $(\eta_1, \eta_2, \ldots, \eta_n)$ с $\eta_n \neq \eta_n^0$. Полученное противоречие показывает, что наша поверхность не может иметь центра.

10.55. Вырожденные поверхности. Вырожденными мы назвали те поверхности, в канонических уравнениях которых участвует меньше чем n координат. Пусть, например, в каноническом уравнении отсутствует координата η_n. Тогда все сечения поверхности $(n-1)$-мерными гиперплоскостями $\eta_n = C$ $(-\infty < C < \infty)$ представляют собой одну и ту же поверхность в $(n-1)$-мерном пространстве. Следовательно, *всякая вырожденная поверхность образуется* *параллельным переносом некоторой поверхности 2-го порядка в $(n-1)$-мерном пространстве R_{n-1} вдоль перпендикуляра к этому $(n-1)$-мерному пространству.*

а. Найдем соответствующие линии на плоскости $(n=2)$; так как в каноническом уравнении в данном случае может

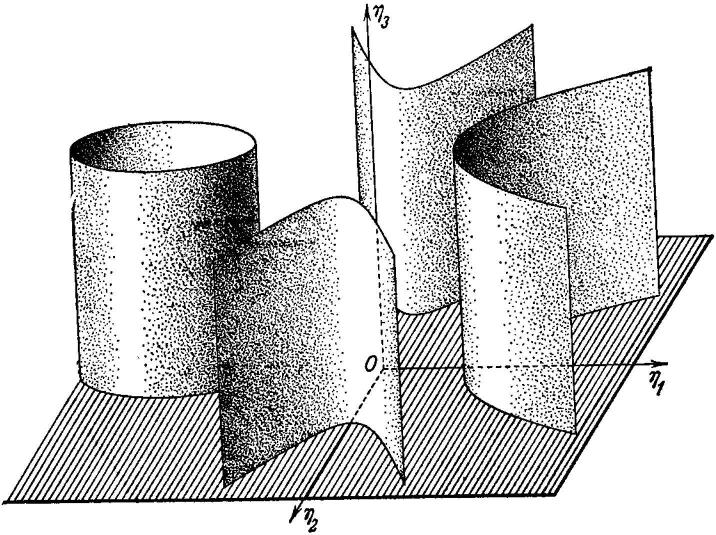

Рис. 8.

участвовать только одна координата, то это уравнение имеет вид

$$\frac{\eta_1^2}{a_1^2}=C.$$

При $C>0$ мы получаем пару параллельных прямых, при $C=0$ — пару слившихся прямых, при $C<0$ — мнимую линию.

б. Чтобы построить вырожденные поверхности в трехмерном пространстве $(n=3)$, нужно подвергнуть параллельному переносу вдоль оси η_3 все кривые 2-го порядка на плоскости (η_1, η_2). При этом эллипс, гипербола, парабола соответственно дают эллиптический, гиперболический, параболический цилиндры (рис. 8). Пара прямых, пересекающихся, параллельных или слившихся, приводит соответственно

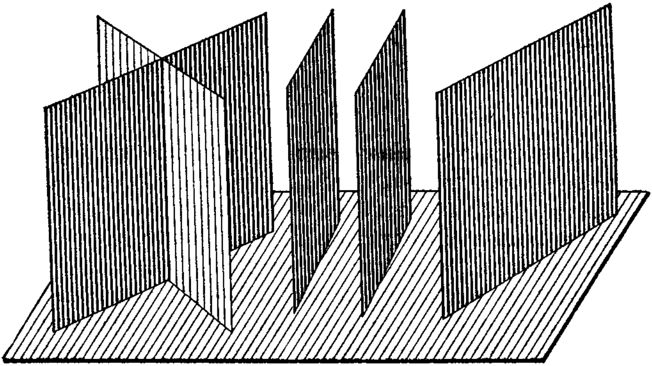

Рис. 9.

к паре плоскостей — пересекающихся, параллельных или слившихся (рис. 9).

§ 10.6. Анализ поверхности по ее общему уравнению

10.61. Мы описали все возможные типы поверхностей 2-го порядка в n-мерном евклидовом пространстве. Тип поверхности определяется по ее каноническому уравнению. Но часто поверхность задается не каноническим, а общим уравнением (19) и бывает существенно определить тип поверхности, иными словами, построить ее каноническое уравнение, не производя всех преобразований, описанных в *10.42.*

Оказывается, чтобы написать каноническое уравнение поверхности, заданной уравнением (19), достаточно знать следующие величины:

а. Корни многочлена n-й степени

$$\Delta(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

б. Коэффициенты многочлена n-й степени

$$\Delta_1(\lambda)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda & b_n \\ b_1 & b_2 & \dots & b_n & c \end{vmatrix}.$$

Чтобы получить явные выражения для коэффициентов этого многочлена, используем линейное свойство определителя (*1.44*). Каждый столбец определителя $\Delta_1(\lambda)$ (кроме последнего) можно представить в виде суммы двух столбцов, первый из которых состоит из числа a_{ij} ($i=1, 2, \dots, n$; j фиксировано) и числа b_j, а второй — из n нулей и числа $-\lambda$. Соответственно определитель $\Delta_1(\lambda)$ представляется в виде суммы некоторого числа определителей, каждый из которых получается заменой некоторых столбцов (кроме последнего) в матрице

$$A_1=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \\ b_1 & b_2 & \dots & b_n & c \end{Vmatrix} \tag{26}$$

на столбцы, состоящие из n нулей и одного элемента $-\lambda$, причем так, что число $-\lambda$ оказывается на главной диагонали матрицы. Каждый из этих определителей после разложения по тем столбцам, в которых стоят числа $-\lambda$, приводится к виду

$$(-\lambda)^k M_{n+1-k},$$

где k — число столбцов, содержащих элементы $-\lambda$, а M_{n+1-k} — некоторый минор $(n+1-k)$-го порядка матрицы A_1. Этот минор характерен тем, что в нем вместе с каждой строкой матрицы A_1 используется столбец этой матрицы с таким же номером и заведомо используются последняя строка и последний столбец этой матрицы. Миноры, обладающие этим свойством, мы будем называть *окаймляющими*. Очевидно, что к а ж д ы й окаймляющий минор матрицы A_1 появится в разложении определителя $\Delta_1(\lambda)$. Отсюда мы непосредственно получаем, что *коэффициент при* $(-\lambda)^k$ *в разложении определителя* $\Delta_1(\lambda)$ *по степеням* $-\lambda$ *равен сумме всех окаймляющих миноров порядка* $n+1-k$. Разложение определителя $\Delta_1(\lambda)$ удобно записать в виде

$$\Delta_1(\lambda)=\alpha_{n+1}-\alpha_n\lambda+\alpha_{n-1}\lambda^2-\ldots+\alpha_1(-\lambda)^n;$$

при такой записи коэффициент α_k будет равен сумме всех окаймляющих миноров матрицы A_1, имеющих порядок k.

10.62. Корни характеристического многочлена $\Delta(\lambda)$, как мы уже знаем, дают нам коэффициенты при квадратах переменных в каноническом уравнении. Чтобы найти оставшийся член (нулевой степени, если каноническое уравнение имеет вид (22), или первой степени, если оно имеет вид (23)), необходимо выяснить поведение коэффициентов многочлена $\Delta_1(\lambda)$ при преобразованиях координат.

Рассмотрим в $(n+1)$-мерном евклидовом пространстве R_{n+1} квадратичную форму

$$\mathrm{A}_1(x,\ x)=\sum_{i,\ k=1}^{n} a_{ik}\xi_i\xi_k+2\sum_{i=1}^{n} b_i\xi_i\xi_{n+1}+c\xi_{n+1}^2, \tag{27}$$

где $\xi_1, \xi_2, \ldots, \xi_n, \xi_{n+1}$ — координаты вектора $x\in\mathrm{R}_{n+1}$ в некотором ортогональном нормированном базисе $e_1, e_2, \ldots$ $\ldots, e_n, e_{n+1}$. Этой квадратичной форме отвечает симметричный оператор A_1, имеющий в базисе $\{e\}$ матрицу (26); мы будем обозначать эту матрицу также через $A_{(e)}$.

Наряду с этим оператором рассмотрим оператор E_1, определяемый равенствами $\mathrm{E}_1e_k=e_k$ при $k\leqslant n$, $\mathrm{E}_1e_{n+1}=0$. Ему отвечает следующая матрица (в том же базисе $e_1, e_2, \ldots, e_{n+1}$):

$$E_1=\begin{Vmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & 1 & \\ & & & & & 0 \end{Vmatrix}. \tag{28}$$

Обозначим через R_n подпространство с базисом из векторов $e_1, e_2, \ldots, e_n$; оператор E_1 в этом подпространстве, очевидно, является тождественным оператором.

Пусть дано некоторое изометрическое преобразование Q в пространстве R_n; оно переводит ортогональный нормированный базис $e_1, e_2, \ldots, e_n$ в некоторый базис $f_1, f_2, \ldots, f_n$, также ортогональный и нормированный. Построим изометрическое преобразование Q_1 в пространстве R_{n+1}, полагая

$$Q_1 e_1 = f_1, \quad Q_1 e_2 = f_2, \quad \ldots, \quad Q_1 e_n = f_n, \quad Q_1 e_{n+1} = e_{n+1} = f_{n+1}.$$

Если матрица оператора Q в пространстве R_n имела вид

$$Q = \left\| \begin{matrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{matrix} \right\|,$$

то в пространстве R_{n+1} матрица построенного оператора Q_1 будет иметь вид

$$Q_1 = \left\| \begin{matrix} q_{11} & q_{12} & \cdots & q_{1n} & 0 \\ q_{21} & q_{22} & \cdots & q_{2n} & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ q_{n1} & q_{n2} & \cdots & q_{nn} & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{matrix} \right\|.$$

Эта матрица отвечает следующим формулам преобразования координат (*8.94*):

$$\left. \begin{aligned} \xi_1 &= q_{11}\eta_1 + q_{21}\eta_2 + \ldots + q_{n1}\eta_n, \\ \xi_2 &= q_{12}\eta_1 + q_{22}\eta_2 + \ldots + q_{n2}\eta_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \xi_n &= q_{1n}\eta_1 + q_{2n}\eta_2 + \ldots + q_{nn}\eta_n, \\ \xi_{n+1} &= \qquad\qquad\qquad\qquad\qquad \eta_{n+1}. \end{aligned} \right\} \tag{29}$$

Оператор A_1 в новом базисе $f_1, f_2, \ldots, f_{n+1}$ имеет матрицу $A_{(f)} = Q^{-1} A_{(e)} Q$ (*5.51*); оператор E_1 — ту же матрицу (28), что и раньше. Согласно *5.52* справедливо равенство

$$\det (A_{(f)} - \lambda E_1) = \det (A_{(e)} - \lambda E_1).$$

Допустим теперь, что в качестве преобразования Q было выбрано то самое, которое в *10.42* приводило квадратичную форму $\mathrm{A}(x, x)=\sum_{i,\ k=1}^{n} a_{ik}\xi_i\xi_k$ к каноническому виду

$$\mathrm{A}(x, x)=\sum_{i=1}^{n}\lambda_i\eta_i^2.$$

Из формул (29) вытекает, что преобразование Q переводит квадратичную форму (27) от $n+1$ переменного в форму

$$\sum_{i=1}^{n}\lambda_i\eta_i^2+2\sum_{i=1}^{n} l_i\eta_i\eta_{n+1}+c\eta_{n+1}^2.$$

Матрица оператора A_1, которая, как мы знаем, преобразуется одинаково с матрицей квадратичной формы, приобретает после этого преобразования следующий вид:

$$A_{(f)}=\left\|\begin{matrix} \lambda_1 & & & & & & & & l_1 \\ & \lambda_2 & & & 0 & & & & l_2 \\ & & \ddots & & & & & & \vdots \\ & & & \lambda_r & & & & & l_r \\ & 0 & & & 0 & & & & l_{r+1} \\ & & & & & \ddots & & & \vdots \\ & & & & & & & 0 & l_n \\ l_1 & l_2 & \ldots & l_r & l_{r+1} & \ldots & & l_n & c \end{matrix}\right\|.$$

Многочлен $\Delta_1(\lambda)=\det(A_{(f)}-\lambda E_1)$ будет равен определителю

$$\left|\begin{matrix} \lambda_1-\lambda & 0 & \ldots & 0 & 0 & \ldots & 0 & l_1 \\ 0 & \lambda_2-\lambda & \ldots & 0 & 0 & \ldots & 0 & l_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_r-\lambda & 0 & \ldots & 0 & l_r \\ 0 & 0 & \ldots & 0 & -\lambda & \ldots & 0 & l_{r+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & -\lambda & l_n \\ l_1 & l_2 & \ldots & l_r & l_{r+1} & \ldots & l_n & c \end{matrix}\right|.$$

Коэффициенты этого многочлена можно вычислить с помощью окаймляющих миноров матрицы $A_{(f)}$ так же, как они вычислялись раньше через окаймляющие миноры матрицы $A_{(e)} = A_1$.

Заметим, что при $r < n$ все окаймляющие миноры матрицы $A_{(f)}$ выше $(r+2)$-го порядка заведомо обращаются в нуль, так как содержат два пропорциональных столбца. Таким образом, при $r < n$ коэффициенты $\alpha_{r+3}, \alpha_{r+4}, \ldots$ равны нулю.

Кроме того, при $r < n$ в окаймляющих минорах $(r+2)$-го порядка, кроме заведомо равных нулю, необходимо используются первые r строк и r столбцов матрицы $A_{(f)}$.

В окаймляющих минорах $(r+1)$-го порядка могут и не использоваться эти r строк и столбцов; отметим все же два случая, когда это использование заведомо имеет место:

1) $r = n$; очевидно, что у матрицы $A_{(f)}$ имеется единственный минор $(r+1)$-го, т. е. $(n+1)$-го порядка, совпадающий с ее определителем; он содержит все строки и столбцы матрицы $A_{(f)}$.

2) $r < n$, $l_{r+1} = l_{r+2} = \ldots = l_n = 0$; кроме заведомо равных нулю, имеется один окаймляющий минор $(r+1)$-го порядка; он лежит в строках и столбцах с номерами $1, 2, \ldots$ $\ldots, r, n+1$.

10.63. Посмотрим, далее, как отразится на матрице оператора A_1 следующий этап преобразований уравнения (19) § 10.4, имеющий целью аннулирование величин $l_1, l_2, \ldots, l_r$. После преобразования

$$\eta_1' = \eta_1 + \frac{l_1}{\lambda_1}\eta_{n+1}, \quad \eta_k' = \eta_k \qquad (k = 2, 3, \ldots, n+1)$$

матрица $A_{(f)}$ переходит в матрицу

$$A_{(f)}^{(1)} = \begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 & \ldots & 0 & l_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_r & 0 & \ldots & 0 & l_r \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & l_{r+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & l_n \\ 0 & l_1 & \ldots & l_r & l_{r+1} & \ldots & l_n & c - \frac{l_n}{\lambda_1} \end{Vmatrix}.$$

Операция, произведенная с матрицей $A_{(f)}$, может быть описана еще так: из последнего столбца вычтен первый столбец, умноженный на $\frac{l_1}{\lambda_1}$, а затем из последней строки вычтена первая строка, также умноженная на $\frac{l_1}{\lambda_1}$. Аналогично можно описать дальнейшие преобразования, имеющие целью аннулировать величины $l_2, l_3, \ldots, l_r$; в результате всех этих преобразований матрица $A_{(f)}$ переходит в матрицу

$$A_{(f)}^{(r)} = \left\| \begin{matrix} \lambda_1 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_r & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & l_{r+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & l_n \\ 0 & 0 & \ldots & 0 & l_{r+1} & \ldots & l_n & c' \end{matrix} \right\|.$$

При этих преобразованиях заведомо не изменяют свою величину те окаймляющие миноры матрицы $A_{(f)}$, в которых участвуют первые r строк и r столбцов этой матрицы. Рассмотрим многочлен

$$\Delta_1^{(r)}(\lambda) = \begin{vmatrix} \lambda_1 - \lambda & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 - \lambda & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_r - \lambda & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & -\lambda & \ldots & 0 & l_{r+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & -\lambda & l_n \\ 0 & 0 & \ldots & 0 & l_{r+1} & \ldots & l_n & c \end{vmatrix} =$$
$$= \alpha'_{n+1} - \alpha'_n \lambda + \ldots + \alpha'_1 (-\lambda)^n.$$

Коэффициенты этого многочлена вычисляются через окаймляющие миноры матрицы $A_{(f)}^{(r)}$ по тем же правилам, что и коэффициенты многочлена $\Delta_1(\lambda)$ через окаймляющие миноры матрицы $A_{(f)}$. В силу доказанного выше свойства неизменности миноров $(r+2)$-го порядка (при $r < n$) мы получаем, что $\alpha'_{r+2}=\alpha_{r+2}$; так же в двух указанных особых случаях мы будем иметь $\alpha'_{r+1}=\alpha_{r+1}$.

10.64. Рассмотрим сначала особый случай $r=n$. Тогда коэффициент α'_{n+1} многочлена $\Delta_1^{(r)}(\lambda)$ равен, очевидно, произведению $\lambda_1\lambda_2\ldots\lambda_n c$; тем самым величина c в каноническом уравнении (22) оказывается равной

$$c=\frac{\alpha'_{n+1}}{\lambda_1\lambda_2\ldots\lambda_n}=\frac{\alpha_{n+1}}{\lambda_1\lambda_2\ldots\lambda_n}.$$

10.65. Пусть теперь $r<n$. Определим нужный для дальнейшего коэффициент α_{r+1} у многочлена $\Delta_1^{(r)}(\lambda)$ *). Окаймляющие миноры матрицы порядка $r+2$, за исключением заведомо равных нулю, имеют вид

$$\begin{vmatrix} \lambda_1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & \ldots & \lambda_r & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & l_m \\ 0 & 0 & \ldots & 0 & l_m & c \end{vmatrix}=-\lambda_1\lambda_2\ldots\lambda_r l_m^2$$

$(m=k+1,\ \ldots,\ n)$, и их сумма, равная коэффициенту $\alpha'_{r+2}=\alpha_{r+2}$, выражается в виде

$$-\lambda_1\lambda_2\ldots\lambda_r(l_{k+1}^2+\ldots+l_n^2).$$

Вспомним, что условием приводимости уравнения (19) к канонической форме (23) было наличие среди коэффициентов $l_{k+1},\ \ldots,\ l_n$ хотя бы одного отличного от нуля. Теперь мы можем эквивалентное условие сформулировать в виде неравенства

$$\alpha_{r+2}\neq 0$$

и одновременно указать формулу для вычисления коэффициента M канонической формы (23)

$$M^2=l_{r+1}^2+\ldots+l_n^2=-\frac{\alpha_{r+2}}{\lambda_1\lambda_2\ldots\lambda_r}.$$

---

*) Нетрудно проверить, что и у многочлена $\Delta_1^{(r)}(\lambda)$ все коэффициенты α_m при $m>r+2$ в этом случае заведомо обращаются в нуль.

Если же $\alpha_{r+2}=0$, то $l_{r+1}=l_{r+2}=\ldots=l_n=0$ и уравнение приводится к канонической форме (22). Мы приходим здесь ко второму особому случаю. Коэффициент $\alpha'_{r+1}=\alpha_{r+1}$ в этом случае равен, очевидно, произведению

$$\lambda_1\lambda_2\ \ldots\ \lambda_r c,$$

откуда коэффициент c канонической формы (22) получается равным

$$\frac{\alpha'_{r+1}}{\lambda_1\lambda_2\ldots\lambda_r}=\frac{\alpha_{r+1}}{\lambda_1\lambda_2\ \ldots\ \lambda_r}.$$

10.66. Приведем сводку полученных результатов. При этом корни $\lambda_1, \lambda_2, \ldots, \lambda_n$ характеристического многочлена $\Delta(\lambda)$ условимся, как и ранее, выписывать в таком порядке, чтобы сначала шли корни, отличные от нуля. Произведение $\lambda_1\ \ldots\ \lambda_r$ обозначим через Λ_r.

Данные		Каноническое уравнение
$\lambda_n \neq 0$		$\lambda_1\eta_1^2+\lambda_2\eta_2^2+\ldots+\lambda_n\eta_n^2+\frac{\alpha_{n+1}}{\Lambda}=0$
$\lambda_n=0$	$\alpha_{n+1}\neq 0$	$\lambda_1\eta_1^2+\lambda_2\eta_2^2+\ldots+\lambda_{n-1}\eta_{n-1}^2+2\sqrt{-\frac{\alpha_{n+1}}{\Lambda_{n-1}}}\,\eta_n=0$
$\lambda_{n-1}\neq 0$	$\alpha_{n+1}=0$	$\lambda_1\eta_1^2+\lambda_2\eta_2^2+\ldots+\lambda_{n-1}\eta_{n-1}^2+\frac{\alpha_n}{\Lambda_{n-1}}=0$
$\lambda_{n-1}=0$	$\alpha_n\neq 0$	$\lambda_1\eta_1^2+\lambda_2\eta_2^2+\ldots+\lambda_{n-2}\eta_{n-2}^2+2\sqrt{-\frac{\alpha_n}{\Lambda_{n-2}}}\,\eta_{n-1}=0$
$\lambda_{n-2}\neq 0$	$\alpha_n=0$	$\lambda_1\eta_1^2+\lambda_2\eta_2^2+\ldots+\lambda_{n-2}\eta_{n-2}^2+\frac{\alpha_{n-1}}{\Lambda_{n-2}}=0$
. . . .		. . . .
$\lambda_2=0$	$\alpha_3\neq 0$	$\lambda_1\eta_1^2+2\sqrt{-\frac{\alpha_3}{\lambda_1}}\,\eta_2=0$
$\lambda_1\neq 0$	$\alpha_3=0$	$\lambda_1\eta_1^2+\frac{\alpha_2}{\lambda_1}=0$

§ 10.7. Эрмитово-квадратичные формы

10.71. Многие из теорем предыдущих параграфов переносятся на квадратичные формы в комплексном пространстве. Начнем со следующей основной теоремы:

Теорема. *В n-мерном унитарном пространстве C_n всякая эрмитово-билинейная симметричная форма имеет канонический базис из n взаимно ортогональных векторов.*

Доказательство. Линейный оператор A, отвечающий данной эрмитово-билинейной симметричной форме $A(x, y)$ (*9.18 а*) по формуле

$$A(x, y) \equiv (Ax, y),$$

имеет в любом ортонормированном базисе пространства C_n ту же матрицу, что и форма $A(x, y)$, и, следовательно, эрмитово-симметричен ($a_j^{(k)} = \overline{a_k^{(j)}}$, $j, k = 1, \ldots, n$). В силу теоремы *9.34* в пространстве C_n имеется ортонормированный базис $e_1, \ldots, e_n$ из собственных векторов оператора A. В этом базисе матрица оператора A диагональна; следовательно, в этом базисе диагональна и матрица формы $A(x, y)$; таким образом, базис $e_1, \ldots, e_n$ есть канонический базис формы $A(x, y)$. Теорема доказана.

10.72. Из теоремы *10.71.* следует, что всякая симметричная эрмитово-квадратичная форма $A(x, x)$ может быть приведена к каноническому виду

$$A(x, x) = \sum_{j=1}^{n} \lambda_j |\xi_j|^2$$

унитарным преобразованием.

Последовательность действий, которые приводят к определению координат векторов искомого канонического базиса и чисел λ_j, такова же, что в вещественном случае (*10.13*).

10.73. Найдем стационарные значения на единичной сфере $\sum_{j=1}^{n} |\xi_j|^2 = 1$ эрмитово-квадратичной симметричной формы $A(x, x)$. Как мы помним из *9,15б*, форма $A(x, x)$ принимает лишь вещественные значения. Пусть $e_1, \ldots, e_n$ — ортонормальный канонический базис формы $A(x, x)$. В этом базисе мы имеем

$$\mathrm{A}(x, x) = \sum_{j=1}^{n} \lambda_j |\xi_j|^2 = \sum_{j=1}^{n} \lambda_j (\sigma_j^2 + \tau_j^2); \quad \xi_j = \sigma_j + i\tau_j;$$

$$(x, x) = \sum_{j=1}^{n} |\xi_j|^2 = \sum_{j=1}^{n} (\sigma_j^2 + \tau_j^2).$$

Действуя по методу Лагранжа, приравниваем нулю частные производные функции $\mathrm{A}(x, x) - \lambda(x, x)$ по каждой из $2n$ вещественных координат σ_j и τ_j ($j = 1, \ldots, n$). При этом получаются уравнения

$$\left.\begin{aligned} 2\lambda_j\sigma_j - 2\lambda\sigma_j = 0, \\ 2\lambda_j\tau_j - 2\lambda\tau_j = 0, \end{aligned}\right\} j = 1, \ldots, n.$$

Эти уравнения удовлетворяются на векторе x с $|x| = 1$, лишь если λ совпадает с одним из значений $\lambda_1, \ldots, \lambda_n$. Пусть $\lambda = \lambda_k$; тогда решением служит вектор x с координатами $\xi_j = \sigma_j + i\tau_j = 0$ при $j \neq k$ и $|\xi_k| = 1$. Следовательно, как и в вещественном случае (*10.21*), эрмитово-квадратичная форма $\mathrm{A}(x, x)$ принимает стационарные значения на векторах ее ортонормального канонического базиса e_j, иными словами, на собственных векторах соответствующего эрмитово-симметричного оператора A. Сами значения формы в этих точках совпадают с соответствующими каноническими коэффициентами; в частности, максимум формы $\mathrm{A}(x, x)$ есть наибольший из коэффициентов λ_j, а минимум — наименьший из этих коэффициентов.

10.74. К теореме *10.72* сводится и проблема об одновременном приведении к каноническому виду двух симметричных эрмитово-квадратичных форм $\mathrm{A}(x, x)$ и $\mathrm{B}(x, x)$, из которых одна, например $\mathrm{B}(x, x)$, положительно определена. Именно, мы можем принять за скалярное произведение эрмитово-билинейную форму $\mathrm{B}(x, y)$. По теореме *10.72* существует ортонормированный (в смысле введенного скалярного произведения) канонический базис формы $\mathrm{A}(x, x)$. В этом базисе мы имеем

$$\mathrm{A}(x, x) = \sum_{j=1}^{n} \lambda_j |\xi_j|^2, \qquad \mathrm{B}(x, x) = \sum_{j=1}^{n} |\xi_j|^2,$$

что и требуется.

Вычисление координат векторов искомого базиса (относительно произвольного исходного базиса) и коэффициентов $\lambda_1, \ldots, \lambda_n$ производится по тем же правилам, что и в вещественном случае (*10.32*). Для обоснования их следует, введя выражения $\xi_j = \sigma_j + i\tau_j$, записать формы $A(x, x)$ и $B(x, x)$ в виде вещественных функций от вещественных переменных σ_j и τ_j $(j = 1, \ldots, n)$. Мы можем предоставить читателю детали этого вывода.

ЗАДАЧИ

1. Привести ортогональным преобразованием к каноническому виду следующие квадратичные формы:

а) $2\xi_1^2 + \xi_2^2 - 4\xi_1\xi_2 - 4\xi_2\xi_3$;

б) $2\xi_1^2 + 5\xi_2^2 + 5\xi_3^2 + 4\xi_1\xi_2 - 4\xi_1\xi_3 - 8\xi_2\xi_3$;

в) $2\xi_1^2 + 2\xi_2^2 + 2\xi_4^2 - 4\xi_1\xi_2 + 2\xi_1\xi_4 + 2\xi_2\xi_3 - 4\xi_3\xi_4$;

г) $2\xi_1\xi_2 + 2\xi_1\xi_3 - 2\xi_1\xi_4 - 2\xi_2\xi_3 + 2\xi_2\xi_4 + 2\xi_3\xi_4$.

2. Где располагаются при $|x| = 1$ стационарные значения квадратичной формы

$$A(x, x) = x_1^2 + \frac{x_2^2}{2} + \frac{x_3^2}{3} \qquad (x = (x_1, x_2, x_3))$$

и каковы они (минимум, максимум)?

3. Показать, что каждая из величин $\mu_1, \mu_2, \ldots, \mu_k$ может достигать крайних границ, указанных в неравенствах *10.25* (6).

4. Квадратичные формы $A(x, x)$ и $B(x, x)$ называются *сравнимыми*, если для любого $x \in R$ выполняется неравенство $A(x, x) \leqslant B(x, x)$. Пусть $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ — канонические коэффициенты формы $A(x, x)$, $\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant \mu_n$ — канонические коэффициенты формы $B(x, x)$. Показать, что для любого k $(1 \leqslant k \leqslant n)$ имеет место неравенство

$$\lambda_k \leqslant \mu_k \text{ *)}.$$

5. Найти общую пару сопряженных направлений у кривых

$$\frac{x^2}{4} + \frac{y^2}{1} = 1; \quad 2xy = 1.$$

6. Построить линейное преобразование, переводящее две квадратичные формы

$$A(x, x) = \xi_1^2 + 2\xi_1\xi_2 + 2\xi_2^2 - 2\xi_1\xi_3 + 3\xi_3^2,$$
$$B(x, x) = \xi_1^2 + 2\xi_1\xi_2 + 3\xi_2\xi_3 - 2\xi_1\xi_3 + 6\xi_3^2$$

к каноническому виду; указать канонический вид.

---

*) Очевидное в случае, когда формы $A(x, x)$ и $B(x, x)$ имеют общий канонический базис.

7. Показать, что базис, в котором формы $A(x, x)$ и $B(x, x)$ записываются в каноническом виде с каноническими коэффициентами соответственно $\lambda_1, \ldots, \lambda_n$ и $\nu_1, \ldots, \nu_n$, определен однозначно с точностью до числовых коэффициентов, если все отношения

$$\frac{\lambda_1}{\nu_1}, \frac{\lambda_2}{\nu_2}, \ldots, \frac{\lambda_n}{\nu_n}$$

попарно различны.

8. Доказать, что середины хорд поверхности 2-го порядка, параллельных вектору $y = \{\xi_1, \xi_2, \ldots, \xi_n\}$, расположены на $(n-1)$-мерной гиперплоскости (диаметральной гиперплоскости, сопряженной с вектором y).

9. Какие поверхности 2-го порядка в трехмерном пространстве (x, y, z) задаются уравнениями

а) $\frac{x^2}{4} - \frac{y^2}{9} + \frac{z^2}{1} = 1$; б) $\frac{x^2}{4} - \frac{y^2}{9} - \frac{z^2}{1} = -1$;

в) $x = y^2 + z^2$; г) $y = x^2 + z^2 + 1$; д) $y = xz$?

10. Упростить уравнения поверхностей 2-го порядка в трехмерном пространстве (x, y, z) и указать соответствующие формулы преобразования координат:

а) $5x^2 + 6y^2 + 7z^2 - 4xy + 4yz - 10x + 8y + 14z - 6 = 0$;

б) $x^2 + 2y^2 - z^2 + 12xy - 4xz - 8yz + 14x + 16y - 12z - 3 = 0$;

в) $4x^2 + y^2 + 4z^2 - 4xy + 8xz - 4yz - 12x - 12y + 6z = 0$.

11. Показать, что в пересечении эллипсоида с полуосями $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n$ и k-мерной гиперплоскости, проходящей через центр этого эллипсоида, получается эллипсоид с полуосями $b_1 \geqslant b_2 \geqslant \ldots \geqslant b_k$, где

$$\begin{gathered} a_1 \geqslant b_1 \geqslant a_{n-k+1}, \\ a_2 \geqslant b_2 \geqslant a_{n-k+2}, \\ \ldots\ldots\ldots\ldots \\ a_k \geqslant b_k \geqslant a_n. \end{gathered}$$

КОНЕЧНОМЕРНЫЕ АЛГЕБРЫ И АЛГЕБРЫ МАТРИЦ

§ 11.1. Еще об алгебрах

11.11. Мы ввели понятие алгебры в *6.21*; такое название было присвоено линейному пространству над некоторым полем K с введенной в нем операцией умножения элементов (подчиненной аксиомам *6.21, 1)—3)*), коммутативной или некоммутативной.

Алгебры, которые мы рассматривали в гл. 6, были в основном коммутативными. Важный пример конечномерной некоммутативной алгебры — алгебры $\mathrm{B}(\mathrm{K}_n)$ всех линейных операторов в пространстве K_n — мы там оставили в стороне. В этой главе мы ставим своей целью изучение алгебры $\mathrm{B}(\mathrm{K}_n)$ и ее подалгебр. Но будет удобнее рассмотреть вначале абстрактные конечномерные алгебры.

11.12. Не всякая алгебра обладает единицей (примером чему служит любая тривиальная алгебра, т. е. алгебра с нулевым умножением, *6.22 а*). Тем не менее каждую алгебру можно дополнить до алгебры с единицей следующим стандартным образом.

Пусть A — произвольная алгебра. Рассмотрим множество A^+, состоящее из формальных сумм вида $a+\lambda$, где $a \in \mathrm{A}$, а λ — число из поля K. Очевидно, A^+ является линейным пространством с операциями

$$(a+\lambda)+(b+\mu)=(a+b)+(\lambda+\mu)$$

и

$$\mu(a+\lambda)=\mu a+\lambda\mu,$$

и, более того, алгеброй относительно умножения

$$(a+\lambda)(b+\mu)=(ab+\lambda b+\mu a)+\lambda\mu$$

(a, $b \in \mathrm{A}$; λ, $\mu \in K$). Алгебра A^+ заведомо обладает единицей: ею является формальная сумма нулевого элемента в А и числа 1. При этом исходную алгебру А можно рассматривать как лежащую в A^+, если каждый элемент $a \in \mathrm{A}$ отождествить с формальной суммой $a+0 \in \mathrm{A}^+$.

§ 11.2. Представления абстрактных алгебр

11.21. Пусть А — абстрактная алгебра над полем K и B(K) — алгебра всех линейных операторов, действующих в линейном пространстве K над тем же полем K. Будем рассматривать морфизмы алгебры А в алгебру B(K), которые в дальнейшем будем обозначать символом вида $\mathrm{T}\colon \mathrm{A} \to \mathrm{B}(\mathrm{K})$.

а. Определение. Морфизм $\mathrm{T}\colon \mathrm{A} \to \mathrm{B}(\mathrm{K})$ называется *представлением алгебры* А *в пространстве* K. Представление называется *тривиальным*, если $\mathrm{T}a = 0$ для каждого $a \in \mathrm{A}$, и *точным*, если морфизм T есть мономорфизм, т. е. для различных элементов a и b алгебры А соответствующие операторы T_a и T_b различны в алгебре B(K).

Совокупность всех элементов $a \in \mathrm{A}$, которые представлением T переводятся в нулевой оператор, называется *ядром* представления T. Ядро тривиального представления совпадает со всей алгеброй А, ядро точного представления состоит из одного нулевого элемента алгебры. В общем случае ядро любого представления есть двусторонний идеал алгебры А (*6.25 г*).

б. Определение. Два представления $\mathrm{T}'\colon \mathrm{A} \to \mathrm{B}(\mathrm{K}')$ и $\mathrm{T}''\colon \mathrm{A} \to \mathrm{B}(\mathrm{K}'')$ алгебры А называются *эквивалентными*, если между линейными пространствами K′ и K″ можно установить изоморфизм $\mathrm{U}\colon \mathrm{K}' \to \mathrm{K}''$ такой, что для любого $a \in \mathrm{A}$

$$\mathrm{U}\mathrm{T}'_a = \mathrm{T}''_a\mathrm{U}.$$

Очевидно, в случае конечномерных пространств K′ и K″ эквивалентность представлений T′ и T″ означает, что в некоторых базисах пространств K′ и K″ операторы T'_a и T''_a ($a \in \mathrm{A}$) записываются одинаковыми матрицами.

в. Пусть $\mathrm{T}\colon \mathrm{A} \to \mathrm{B}(\mathrm{K})$ — представление алгебры А. Подпространство K′ в K мы будем называть *инвариантным подпространством представления* T, если оно инвариантно относительно всех операторов T_a, $a \in \mathrm{A}$. Рассматривая операторы T_a лишь на K′, мы, очевидно, получим некоторое новое представление $T^{K'}: A \to B(K')$, которое будем называть *сужением* представления T на K'.

г. Пусть $T: A \to B(K)$ — представление алгебры A такое, что K разлагается в прямую сумму подпространств K_k $(1 \leqslant k \leqslant n)$, инвариантных относительно представления T. Обозначим через T^k сужение представления T на K_k $(1 \leqslant k \leqslant n)$. В такой ситуации мы будем говорить, что *представление* T *разлагается в прямую сумму представлений* T^k $(1 \leqslant k \leqslant n)$.

11.22. С каждой алгеброй A естественным образом связано ее некоторое представление $\tilde{T}: A \to B(A)$ в самом линейном пространстве A, которое ставит в соответствие каждому элементу $a \in A$ оператор левого умножения на a, т. е. оператор $\tilde{T}_a$, определенный равенством $\tilde{T}_a b = ab$ для любого $b \in A$. Такое представление называется *левым регулярным представлением алгебры* A. Очевидно, инвариантные подпространства левого регулярного представления являются левыми идеалами в A (*6.23 а*). С помощью этого понятия установим следующий важный результат:

Теорема. *Любая алгебра изоморфна некоторой подалгебре алгебры* B(K) (при надлежащем выборе K).

Доказательство. Как легко видеть, утверждение теоремы равносильно следующему: каждая алгебра обладает точным представлением.

Пусть A — заданная алгебра. Как показано в *11.12*, существует алгебра A^+, обладающая единицей e и содержащая A в качестве подалгебры. Рассмотрим ее левое регулярное представление $\tilde{T}: A^+ \to B(A^+)$. Поскольку для любого $a \in A^+$, $a \neq 0$, $\tilde{T}_a e = ae = a \neq 0$, это представление является точным. Следовательно, сужение морфизма $\tilde{T}$ на подалгебру A в A^+ является точным представлением алгебры A в пространстве $K = A^+$. Тем самым теорема доказана.

§ 11.3. Неприводимые представления и лемма Шура

11.31. Среди всех представлений заданной алгебры можно выделить некоторые, устроенные в определенном смысле наиболее просто.

Каждое представление $T: A \to B(K)$ некоторой алгебры A обладает по крайней мере двумя инвариантными подпространствами: самим K и его нулевым подпространством. Остальные инвариантные подпространства мы будем называть *собственными*. Те из них, которые не содержат ни одного другого подпространства того же вида, будем называть *минимальными инвариантными подпространствами представления* T.

Определение. Нетривиальное представление T: A $\to$ $\to$ B (K) называется *неприводимым*, если у него отсутствуют собственные инвариантные подпространства.

11.32. Пусть $z \in$ K — некоторый вектор; тогда, как легко видеть, множество $K_z = \{T_a z \in K: a \in A\}$ является инвариантным подпространством представления T. Будем называть вектор $z \in$ K *циклическим* (относительно представления T), если $K_z =$ K. Из этого определения и определения неприводимости немедленно вытекает следующая теорема:

Теорема. *Представление, действующее в пространстве* K, *неприводимо тогда и только тогда, когда каждый отличный от нуля вектор пространства* K *является циклическим*.

Несмотря на свою простоту, этот результат окажется весьма полезным в дальнейшем.

11.33. Неприводимые представления алгебр над полем C комплексных чисел обладают следующим важным свойством:

Теорема (лемма Шура). *Пусть* T: A $\to$ B (C) — *неприводимое представление алгебры* A *над полем* C. *Тогда любой оператор в пространстве* C, *перестановочный со всеми операторами* T_a, $a \in$ A, *кратен тождественному оператору* E.

Доказательство. Пусть S — оператор, перестановочный со всеми T_a, $a \in$ A, и x — его собственный вектор (*4.95б*). Тогда $Sx = \lambda x$ для некоторого комплексного λ и, следовательно, $ST_a x = T_a Sx = \lambda T_a x$ для любого $a \in$ A. Но представление T неприводимо, а значит, по теореме *11.32*, любой вектор $y \in$ C представим в виде $y = T_a x$, $a \in$ A. Отсюда $Sy = \lambda y$ для любого $y \in$ C, что и требовалось доказать.

Заметим, что при доказательстве существенно использован тот факт, что каждый линейный оператор в комплексном линейном пространстве обладает собственным вектором. В силу решающей роли леммы Шура, в дальнейшем мы ограничиваемся рассмотрением линейных пространств и алгебр только над полем комплексных чисел.

§ 11.4. Основные типы конечномерных алгебр

Начиная с этого параграфа, если специально не оговорено обратное, будут рассматриваться исключительно конечномерные алгебры (т. е. алгебры, имеющие как линейные пространства конечную размерность) над полем C комплексных чисел.

Как устроены конечномерные алгебры и их представления? Основная часть настоящей главы будет посвящена результатам именно в этом направлении. В частности, будут выделены некоторые классы алгебр, строение которых удается изучить полностью, т. е. удается описать все такие алгебры с точностью до изоморфизма, а все их представления — с точностью до эквивалентности; именно, речь будет идти о классах простых и полупростых алгебр.

Различные классы алгебр возникают при рассмотрении специфических свойств их идеалов и представлений.

11.41. Определение. Нетривиальная алгебра (*6.22а*) называется *простой*, если она не содержит собственных двусторонних идеалов. Примером простой алгебры является алгебра $\mathrm{B}(C_n)$ всех линейных операторов в конечномерном пространстве. Действительно, пусть J — двусторонний идеал в алгебре $\mathrm{B}(\mathrm{K}_n)$ и пусть $A_0 = \|a_{jk}\| \in \mathrm{J}$ — ненулевая матрица, так что, например, $a_{sq} \neq 0$. Тогда, как было показано в *4.44в*, операциями умножения матрицы A справа и слева на некоторые матрицы, т. е. операциями, не выводящими за пределы идеала J, можно получить матрицу E_{rt} с единственным ненулевым элементом 1 в пересечении r-й строки и t-го столбца при любых r и t от 1 до n. Линейные комбинации матриц E_{rt} дают любую матрицу $A \in \mathrm{B}(\mathrm{K}_n)$, откуда и следует, что $\mathrm{J} = \mathrm{B}(\mathrm{K}_n)$. Как мы увидим далее, этот пример в классе всех конечномерных алгебр над полем комплексных чисел — единственный (*11.64*).

Теорема. *Простая алгебра обладает точным неприводимым представлением.*

Доказательство. Пусть A — простая алгебра. Рассмотрим ее левое регулярное представление $\tilde{\mathrm{T}}: \mathrm{A} \to \mathrm{B}(\mathrm{A})$. Из конечномерности A немедленно вытекает, что среди инвариантных подпространств представления $\tilde{\mathrm{T}}$ имеется минимальное подпространство A'. Сужение T представления $\tilde{\mathrm{T}}$ на A'

не является тривиальным представлением; покажем это. Для этого достаточно показать, что для любого $b \in A'$ множество $Ab = \{ab: a \in A\} \neq \{0\}$ *).

Пусть, напротив, $Ab = 0$. Тогда, как легко видеть, множество $bA = \{ba: a \in A\}$ — двусторонний идеал в A и ввиду простоты A либо $bA = A$, либо $bA = \{0\}$. Но в случае $bA = A$ из $Ab = 0$ вытекает, что любое произведение в A равно нулю. В случае же $bA = \{0\}$ множество $\{\lambda b: \lambda \in C\}$ является ввиду $Ab = \{0\}$ двусторонним идеалом в A и вследствие простоты A должно совпадать со всей алгеброй. Таким образом, в обоих случаях алгебра A оказывается тривиальной, а значит, не может быть простой.

Итак, представление $T: A \to B(A')$ не является тривиальным. Но тогда, во-первых, оно ввиду минимальности неприводимо, а во-вторых, его ядро, будучи отличным от всей простой алгебры ее двусторонним идеалом, состоит из одного нуля. Поэтому всякое неприводимое представление является одновременно и точным, что и доказывает теорему.

На самом деле имеет место и обратная теорема, т. е. конечномерная алгебра, имеющая точное неприводимое представление, проста. Это будет показано в конце *11.64*.

11.42. Произвольная алгебра может и не иметь точных неприводимых представлений. Однако естественно выделить те алгебры, свойства которых могут быть описаны в терминах их неприводимых представлений. Мы приходим к следующему классу алгебр, более широкому чем класс простых алгебр:

Определение. Алгебра называется *полупростой*, если для каждого ее элемента, отличного от нуля, существует неприводимое представление, отображающее этот элемент в отличный от нуля оператор.

Иными словами, в полупростой алгебре пересечение всех ядер ее неприводимых представлений состоит из одного нуля.

В силу теоремы *11.41* всякая простая алгебра является и полупростой. С другой стороны, рассмотрим коммута-

*) Следующее ниже простое доказательство этого факта предложено А. С. Немировским.

тивную n-мерную ($n>1$) алгебру C_n из элементов $a=(\alpha_1, \ldots, \alpha_n)$, где $\alpha_j \in C$, с покоординатным умножением. Эта алгебра коммутативна; множество тех $a=(\alpha_1, \ldots, \alpha_n)$, для которых, например, $\alpha_1=0$, представляет ее (двусторонний) идеал, так что алгебра C_n не простая. Приводя в соответствие элементу $a=(\alpha_1, \ldots, \alpha_n)^n$ комплексное число α_k $(1 \leqslant k \leqslant n)$ (или, что то же, оператор умножения на число α_k в одномерном пространстве C_1), мы получаем неприводимое представление алгебры C_n, которое переводит не в нуль любой элемент алгебры C_n с $\alpha_k \neq 0$. Так как у любого элемента алгебры C_n имеется хотя бы одна ненулевая координата, то для любого элемента алгебры C_n имеется и неприводимое представление, переводящее этот элемент в ненулевой оператор. Таким образом, алгебра C_n полупростая.

В этом примере алгебра C_n есть прямая сумма простых (одномерных) коммутативных алгебр. Легко можно его обобщить, рассмотрев прямую сумму простых некоммутативных алгебр; и тогда, как мы покажем ниже (*11.77*), мы получим общий вид полупростой конечномерной алгебры над полем комплексных чисел.

11.43. Алгебры, в определенном смысле противоположные по своим свойствам полупростым алгебрам, выделяются следующим образом.

Определение. Алгебра называется *радикальной*, если каждое ее нетривиальное представление обладает собственным инвариантным подпространством.

Иными словами, у радикальных алгебр вообще отсутствуют неприводимые представления.

В качестве примера рассмотрим алгебру A из многочленов $p(z)=d_1z+\ldots+d_nz^n$ с обычными операциями, но с условием $z^{n+1}=0$. Тогда и каждый элемент алгебры A, возведенный в $(n+1)$-ую степень, равен 0, так что все ее элементы необратимы. Алгебра A не имеет нетривиального одномерного представления, поскольку всякий ненулевой линейный оператор в одномерном пространстве обратим. Пусть T — нетривиальное (следовательно, неодномерное) представление алгебры A и Z — оператор, соответствующий элементу z. Так как Z (вместе с z) необратим, то имеется вектор e, для которого $Ze=0$. Но тогда и $p(Z)e=0$ при любом $p(z) \in \mathrm{A}$. Таким образом, у представления T нашлось нетривиальное инвариантное подпространство (прямая, определяемая вектором e). Мы видим, что алгебра A радикальная.

11.44. Определение. Пусть A — произвольная алгебра. Тогда *радикалом* алгебры A называется пересечение всех ядер неприводимых представлений алгебры A, если последние имеются, или вся алгебра A, если они отсутствуют.

Так как ядро каждого представления есть двусторонний идеал алгебры A (*11.21а*), то и радикал, как пересечение некоторых двусторонних идеалов, также, очевидно, является двусторонним идеалом алгебры A.

Изучение алгебр с нетривиальным радикалом (в частности, радикальных) вызывает существенные трудности, и результаты, как правило, не носят окончательного характера; некоторые из них мы приведем в конце главы. Напротив, полупростые алгебры и их представления удается исследовать полностью; их изучение, как мы увидим в дальнейшем, сводится к изучению простых алгебр.

К подробному рассмотрению простых алгебр, а также их представлений мы сейчас и переходим.

§ 11.5. Строение левого регулярного представления простой алгебры

11.51. Итак, пусть A — простая алгебра. Зафиксируем ее некоторое точное неприводимое представление $\mathrm{T}\colon \mathrm{A} \to \mathrm{B}(\mathrm{X})$ (оно существует согласно теореме *11.41*). В дальнейшем это представление мы будем называть *стандартным*.

Теорема. *Пусть* $\tilde{\mathrm{T}}\colon \mathrm{A} \to \mathrm{B}(\mathrm{A})$ — *левое регулярное представление простой алгебры* A, I — *минимальное инвариантное подпространство представления* $\tilde{\mathrm{T}}$. *Тогда*

а) *сужение* $\tilde{\mathrm{T}}^{\mathrm{I}}$ *представления* $\tilde{\mathrm{T}}$ *на* I *эквивалентно* T,

б) *подпространство* I, *рассматриваемое как подалгебра в* A, *обладает правой единицей.*

Зафиксируем некоторый элемент $a \in \mathrm{I}$, $a \neq 0$. Поскольку представление T точно, $\mathrm{T}_a x \neq 0$ для некоторого $x \in \mathrm{X}$. Рассмотрим линейный оператор $\mathrm{U}\colon \mathrm{I} \to \mathrm{X}$, определенный равенством $\mathrm{U}b = \mathrm{T}_b x$ для любого $b \in \mathrm{I}$.

Ядро оператора U является, как легко видеть, левым идеалом в A или, что то же самое, инвариантным подпространством представления $\tilde{\mathrm{T}}$, содержащимся в I и не совпадающим с ним. Следовательно, ядро U состоит лишь из нулевого элемента. С другой стороны, образ U является, очевидно, отличным от нуля инвариантным подпространством неприводимого представления T и поэтому совпадает со всем X.

Итак, U — изоморфизм I на X. Кроме того, для любых $b \in \mathrm{I}$ и $c \in \mathrm{A}$

$$\mathrm{U}\tilde{\mathrm{T}}_c^{\mathrm{I}} b = \mathrm{U}(cb) = \mathrm{T}_{cb}x = \mathrm{T}_c(\mathrm{T}_b x) = \mathrm{T}_c\mathrm{U}_b;$$

следовательно, $\mathrm{U}\tilde{\mathrm{T}}_c^{\mathrm{I}} = \mathrm{T}_c\mathrm{U}$. Тем самым доказано, что представления $\tilde{\mathrm{T}}^{\mathrm{I}}$ и T эквивалентны.

Далее, поскольку U отображает I на все X, существует $e \in \mathrm{I}$ такой, что $\mathrm{U}e = \mathrm{T}_e x = x$. Отсюда для любого $b \in \mathrm{I}$

$$\mathrm{U}(be) = \mathrm{T}_{be}x = \mathrm{T}_b(\mathrm{T}_e x)\,\mathrm{T}_b x = \mathrm{U}b.$$

Но U — взаимно однозначное отображение; следовательно, $be = b$. Таким образом, e — правая единица в алгебре I, что и завершает доказательство теоремы.

Заметим, что в качестве стандартного представления мы могли бы взять любое неприводимое представление простой алгебры. Поэтому автоматическим следствием доказанной теоремы является тот факт, что все неприводимые представления простой алгебры эквивалентны.

11.52. Лемма. *Пусть* A — *произвольная алгебра,* I_1 *и* I_2 — *ее левые идеалы, обладающие правыми единицами соответственно* e_1 *и* e_2', *причем* $ae_1 = 0$ *для любого* $a \in \mathrm{I}_2$. *Тогда в* I_2 *существует правая единица* e_2 *такая, что* $be_2 = 0$ *для любого* $b \in \mathrm{I}_1$.

Положим $e_2 = e_2' - e_1e_2'$. Тогда для любого $a \in \mathrm{I}_2$ мы будем иметь $ae_2 = ae_2' - ae_1e_2' = a$ ввиду соотношений $ae_2' = a$ и $ae_1 = 0$. Далее, для любого $b \in \mathrm{I}_1$ имеют место равенства $be_2 = be_2' - be_1e_2' = be_2' - be_2' = 0$, что и доказывает лемму.

11.53. Теорема. *Левое регулярное представление простой алгебры разлагается в прямую сумму ее неприводимых представлений.*

Искомый набор минимальных инвариантных подпространств представления $\tilde{\mathrm{T}}$: $\mathrm{A} \to \mathrm{B}(\mathrm{A})$ будем строить индуктивно, доказывая на каждом шаге, что прямая сумма уже найденных подпространств обладает, как алгебра, правой единицей.

В качестве первого подпространства зафиксируем любое минимальное инвариантное подпространство I_1 представления $\tilde{\mathrm{T}}$. Согласно теореме *11.52* I_1 обладает правой единицей e_1.

Пусть теперь найдены минимальные инвариантные подпространства $\mathrm{I}_1, \ldots, \mathrm{I}_k$ такие, что левый идеал $\mathrm{J}'_k = \mathrm{I}_1 + \ldots + \mathrm{I}_k$ обладает правой единицей e_k. В случае $\mathrm{J}'_k = \mathrm{A}$ искомые инвариантные подпространства уже построены. В противном случае положим $\mathrm{J}''_k = \{a \in \mathrm{A}\colon ae_k = 0\}$. Как легко видеть, J''_k — инвариантное подпространство представления $\tilde{\mathrm{T}}$, имеющее нулевое пересечение с J'_k. Более того, поскольку любой элемент $a \in \mathrm{A}$ представим в виде $a = ae_k + (a - ae_k)$, где $ae_k \in \mathrm{J}'_k$ и $(a - ae_k) \in \mathrm{J}''_k$, алгебра A разлагается в прямую сумму J'_k и J''_k.

Конечномерное инвариантное подпространство J''_k содержит некоторое минимальное инвариантное подпространство; обозначим его I_{k+1}. Идеал I_{k+1} обладает согласно теореме *11.52* правой единицей e'_{k+1}; при этом, поскольку $\mathrm{I}_{k+1} \subseteq \mathrm{J}''_k$, $ae_k = 0$ для любого $a \in \mathrm{I}_{k+1}$. Следовательно, по лемме *11.52* в I_{k+1} существует такая правая единица e''_k, что $be''_k = 0$ для любого $a \in \mathrm{J}'_k$.

Положим $e_{k+1} = e_k + e''_k$; тогда, как легко видеть, e_{k+1} — правая единица в идеале $\mathrm{J}'_{k+1} = \mathrm{I}_1 + \ldots + \mathrm{I}_k + \mathrm{I}_{k+1}$. Тем самым доказана законность индуктивного перехода от k к $k+1$.

Алгебра A конечномерна, поэтому на некотором шаге мы получим набор минимальных инвариантных подпространств $\mathrm{I}_1, \ldots, \mathrm{I}_m$ представления $\tilde{\mathrm{T}}$, дающих в прямой сумме всю алгебру A. Следовательно, левое регулярное представление алгебры A разлагается в прямую сумму ее неприводимых представлений, что и требовалось доказать.

11.54. Заметим, что попутно нами установлено наличие в простой алгебре правой единицы. В действительности имеет место более сильная теорема:

Теорема. *Простая алгебра обладает единицей.*

Пусть A — простая алгебра и e — ее правая единица. Рассмотрим оператор T_e в стандартном представлении $T\colon A \to B(X)$. Для любых $x \in X$ и $a \in A$

$$T_a(T_e x - x) = T_{ae}x - T_a x = 0.$$

Следовательно, поскольку T неприводимо и каждый ненулевой вектор обязан быть циклическим, мы получаем, что $T_e x - x = 0$ для любого $x \in X$; иначе говоря, T_e — тождественный оператор в пространстве X. Но тогда $T_a T_e = T_e T_a = T_a$ для любого $a \in A$, а значит, ввиду точности представления T и $ae = ea = a$. Отсюда e — единица в A, и теорема доказана.

§ 11.6. Структура простых алгебр

В этом параграфе до конца будет решен вопрос о строении простых алгебр. Для этого окажется весьма полезным следующее понятие.

11.61. Пусть X — линейное пространство, A_0 — некоторая подалгебра в B(X). Подмножество в B(X), состоящее из операторов, перестановочных со всеми операторами из A_0, мы будем называть *коммутатором* алгебры A_0 и обозначать через $\overline{A}_0$.

Как легко видеть, множество $\overline{A}_0$ само образует подалгебру в B(X). Коммутатор этой новой подалгебры мы будем обозначать через $\overline{\overline{A}}_0$ и называть *вторым коммутатором* алгебры A_0. Очевидно, $A_0 \subseteq \overline{\overline{A}}_0$.

11.62. В произвольной алгебре A каждый элемент $a \in A$ определяет два оператора из B(A): оператор левого умножения $\tilde{T}_a$, действующий по формуле $T_a b = ab$, и оператор правого умножения R_a, задаваемый равенством $R_a b = ba$. Операторы левого умножения, равно как и операторы правого умножения, образуют, как легко видеть, две подалгебры в B(A); обозначим их соответственно через A_0^l и A_0^r.

Лемма. *Пусть алгебра A обладает единицей. Тогда* $\overline{A_0^l} = A_0^r$ *и* $\overline{A_0^r} = A_0$.

Возьмем $S \in \overline{A_0^l}$, тогда для любых a, $b \in A$ имеет место равенство

$$S(ab) = ST_a b = T_a S b = a(Sb).$$

Положив $b = e$, где e — единица в A, получим $Sa = aS(e)$. Следовательно, оператор S является оператором умножения справа на элемент $Se \in A$, т. е. $S \in A_0^r$.

Таким образом, $\overline{A_0^l} \subseteq A_0^r$ и, поскольку обратное включение очевидно, $\overline{A_0^l} = A_0^r$. Равенство $\overline{A_0^r} = A_0^l$ доказывается совершенно аналогично.

11.63. Т е о р е м а. *Пусть* A — *простая алгебра*, A_0 — *алгебра операторов ее стандартного представления* T: $A \rightarrow B(X)$. *Тогда* $\overline{\overline{A_0}} = A_0$.

Определенную выше алгебру A_0^l можно, очевидно, рассматривать как алгебру операторов левого регулярного представления $\tilde{T}$: $A \rightarrow B(A)$ алгебры A. Согласно теореме *11.53* это представление разлагается в прямую сумму некоторых неприводимых представлений $\tilde{T}^{I_i}$: $A \rightarrow B(I_i)$ $(1 \leqslant i \leqslant m)$, причем, по теореме *11.51*, каждое из этих представлений эквивалентно стандартному.

Сказанное означает следующее: можно найти такой базис $x_1, \ldots, x_n$ в пространстве X и такой базис $f_1^{(i)}, \ldots, f_n^{(i)}$ в каждом из пространств I_i $(1 \leqslant i \leqslant m)$, что для любого $a \in A$ матрица оператора $\tilde{T}_a$ в базисе $f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(m)}$ всего пространства A имеет вид

$$\tilde{T}_a = \begin{Vmatrix} T_a & & & & \\ & T_a & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & T_a \end{Vmatrix}, \tag{1}$$

где в блоках по диагонали стоит матрица оператора T_a в базисе $x_1, \ldots, x_n$, а в остальных местах — нули.

Из правила перемножения матриц вытекает, что матрицы, перестановочные со всеми матрицами вида (1), суть матрицы вида

$$\tilde{S} = \begin{Vmatrix} S_{11} & \ldots & S_{1m} \\ \cdot & \cdot \cdot \cdot & \cdot \\ S_{m1} & \ldots & S_{mm} \end{Vmatrix}, \qquad (2)$$

где в блоках (размера $n \times n$) расположены матрицы, перестановочные со всеми матрицами T_a, $a \in \mathrm{A}$.

Рассмотрим теперь оператор $\mathrm{P} \in \overline{\overline{\mathrm{A}}}_0$ и его матрицу P в базисе $x_1, \ldots, x_n$. Тогда, очевидно, матрица

$$\tilde{P} = \begin{Vmatrix} P & & & & \\ & P & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & P \end{Vmatrix}$$

перестановочна со всеми матрицами вида (2) и, следовательно, задает в базисе $f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(m)}$ пространства A оператор, принадлежащий второму коммутатору алгебры A_0^l.

Простая алгебра по теореме *11.54* обладает единицей. Следовательно, согласно лемме *11.62* $\overline{\overline{\mathrm{A}}}{}_0^l = \overline{\mathrm{A}}{}_0^r = \mathrm{A}_0^l$, а значит, матрица $\tilde{P}$ задает в базисе $f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(m)}$ оператор $\tilde{\mathrm{P}}$, равный $\tilde{\mathrm{T}}_b$ для некоторого $b \in \mathrm{A}$. Но тогда для того же b $\mathrm{P} = \mathrm{T}_b$, и поэтому P принадлежит алгебре A_0. Поскольку P — произвольный элемент из $\overline{\overline{\mathrm{A}}}_0$, теорема доказана.

11.64. Теперь мы в состоянии получить основную теорему о простых алгебрах.

Т е о р е м а. (П е р в а я с т р у к т у р н а я т е о р е м а.) *Всякая простая алгебра изоморфна алгебре всех линейных операторов, действующих в некотором конечномерном пространстве.*

Пусть A — простая алгебра, $\mathrm{T}\colon \mathrm{A} \to \mathrm{B}(\mathrm{X})$ — ее стандартное представление. Достаточно показать, что алгебра A_0 операторов представления T совпадает с $\mathrm{B}(\mathrm{X})$.

Поскольку представление T неприводимо, из леммы Шура немедленно следует, что коммутатор $\overline{\mathrm{A}}_0$ алгебры A_0 состоит из операторов, кратных тождественному. Но тогда второй коммутатор $\overline{\overline{\mathrm{A}}}_0$ совпадает со всей алгеброй $\mathrm{B}(\mathrm{X})$. В то же время по теореме *11.63* $\overline{\overline{\mathrm{A}}}_0 = \mathrm{A}_0$. Следовательно, $\mathrm{A}_0 = \mathrm{B}(\mathrm{X})$, что и требовалось доказать.

Обратим внимание на следующее обстоятельство: в основе всех рассуждений, которые привели к первой структурной теореме, лежит лишь тот факт, что всякая простая алгебра обладает точным неприводимым представлением. Следовательно, мы одновременно доказали и то, что любая алгебра, обладающая точным неприводимым представлением, изоморфна алгебре $\mathrm{B}(\mathrm{X})$. Отсюда уже немедленно вытекает, что имеет место теорема, обратная к теореме *11.41*: *всякая алгебра, обладающая точным представлением, является простой.*

§ 11.7. Структура полупростых алгебр

11.71. В этом параграфе будет показано, что вопрос о строении полупростых алгебр полностью сводится к уже изученному вопросу о строении простых алгебр. Для этого окажется полезным ввести несколько новых понятий.

Определение. *Нормальным рядом* алгебры А называется цепочка алгебр $\mathrm{A} = \mathrm{I}_0 \supseteq \mathrm{I}_1 \supseteq \ldots \supseteq \mathrm{I}_n \supseteq \mathrm{I}_{n+1} = (0)$, в которой каждая из алгебр является двусторонним идеалом предыдущей. *Композиционным рядом* алгебры называется такой ее нормальный ряд, в котором каждый из этих идеалов является максимальным (т. е. не содержащимся ни в каком более широком двустороннем идеале), а алгебра I_n не содержит собственных двусторонних идеалов.

Легко показать, что каждая конечномерная алгебра обладает композиционным рядом. Действительно, среди двусторонних идеалов конечномерной алгебры А существует максимальный; обозначим его I_1. По тем же причинам алгебра I_1 содержит максимальный двусторонний идеал I_2, I_2 — идеал I_3 с теми же свойствами и т. д.

Поскольку исходная алгебра А конечномерна, через конечное число шагов мы придем к некоторой алгебре I_n, уже не имеющей собственных идеалов. Таким образом, полученная цепочка

$$\mathrm{A} = \mathrm{I}_0 \supset \mathrm{I}_1 \supset \ldots \supset \mathrm{I}_n \supset \mathrm{I}_{n+1} = (0)$$

является, очевидно, композиционным рядом алгебры А.

11.72. Прежде чем перейти к специальным свойствам нормальных и композиционных рядов полупростых алгебр, докажем вспомогательное утверждение.

Лемма. *Для любого элемента a полупростой алгебры A существует элемент $b \in \mathrm{A}$ такой, что любая степень элемента ba отлична от нуля.*

Доказательство. Согласно определению полупростой алгебры существует неприводимое представление $\mathrm{T}: \mathrm{A} \longrightarrow \mathrm{B}(\mathrm{X})$ такое, что $\mathrm{T}_a \neq 0$. Тогда для некоторого $x \in \mathrm{X}$, $x \neq 0$, вектор $y = \mathrm{T}_a x$ отличен от нуля и, следовательно, является циклическим вектором неприводимого представления T (*11.32*). Поэтому найдется элемент $b \in \mathrm{A}$ такой, что $\mathrm{T}_b y = x$, а значит, $\mathrm{T}_{ba} x = \mathrm{T}_b(\mathrm{T}_a x) = \mathrm{T}_b y = x$. Это означает, что никакая степень оператора T_{ba}, а следовательно, и элемента $ba \in \mathrm{A}$ не может равняться нулю, что и доказывает лемму.

11.73. Теорема. *Нормальный ряд полупростой алгебры не содержит отличных от нуля тривиальных алгебр.*

Доказательство. Пусть A — полупростая алгебра; $\mathrm{A} = \mathrm{I}_0 \supseteq \mathrm{I}_1 \supseteq \ldots \supseteq \mathrm{I}_n \supseteq \mathrm{I}_{n+1} = (0)$ — ее нормальный ряд. Не теряя общности, можно предположить, что алгебра I_n содержит элемент a, отличный от нуля. Для доказательства теоремы, очевидно, достаточно найти $c \in \mathrm{I}_n$ такой, что $ca \neq 0$.

По лемме *11.72* существует $b \in \mathrm{A}$ такой, что любая степень элемента ba отлична от нуля. Положим $c_k = (ba)^{2^k - 1} b$ ($k = 0, 1, \ldots, n - 1$). Индукцией по k доказать, что $c_k \in \mathrm{I}_{k-1}$. Действительно, для $k = 0$ $c_0 = bab \in \mathrm{I}_1$ ввиду того, что $a \in \mathrm{I}_1$, а возможность индуктивного перехода немедленно вытекает из очевидного соотношения $c_{k+1} = c_k a c_k$ и того факта, что $a \in \mathrm{I}_{k+2}$.

Таким образом, доказано, что элемент $c = c_{n-1}$ принадлежит алгебре I_n. Но $ca = (ba)^{2^{n-1}-1} ba = (ba)^{2^{n-1}} \neq 0$. Тем самым теорема доказана.

11.74. Докажем еще несколько простых утверждений.

Лемма. *Пусть $\mathrm{A} \supseteq \mathrm{I}_1 \supseteq \mathrm{I}_2 \supset 0$ — нормальны ряд алгебры A, причем алгебра I_2 проста. Тогда I_2 является двусторонним идеалом в A.*

Доказательство. По теореме *11.54.* алгебра I_2 обладает единицей e. Поскольку $e \in \mathrm{I}_1$, для любого $a \in \mathrm{A}$

элементы ae и ea принадлежат I_1. Но тогда для любого $b \in \mathrm{I}_2$

$$ab = a(eb) = (ae)b \in \mathrm{I}_2 \text{ и } ba = (be)a = b(ea) \in \mathrm{I}_2,$$

что и доказывает лемму.

11.75. Лемма. *Пусть* A — *произвольная алгебра*, I — *ее двусторонний идеал, обладающий единицей. Тогда в* A *есть двусторонний идеал* J *такой, что* A *разлагается в прямую сумму* I *и* J.

Положим $\mathrm{J} = \{a \in \mathrm{A}: ae = 0\}$, где e — единица алгебры I. Очевидно, J — левый идеал в A, причем, поскольку $b = be + (b - be)$ и $b - be \in \mathrm{J}$ для любого $b \in \mathrm{A}$, имеет место разложение $\mathrm{A} = \mathrm{I} + \mathrm{J}$.

Осталось доказать, что J — правый идеал в A. Для любых $a \in \mathrm{J}$ и $b \in \mathrm{A}$

$$ab = abe + a(b - be).$$

Поскольку $be \in \mathrm{I}$, $be = ebe$, а значит, поскольку $ae = 0$,

$$abe = (ae)be = 0.$$

Таким образом, $ab = a(b - be)$, а следовательно, ab является произведением двух элементов из J. Тем самым $ab \in \mathrm{J}$, и лемма доказана.

11.76. Лемма. *Пусть алгебра* A *разлагается в прямую сумму своих двусторонних идеалов* I *и* J, *причем* I — *максимальный двусторонний идеал в* A. *Тогда алгебра* J *не содержит собственных двусторонних идеалов.*

Пусть J′ — двусторонний идеал в J, не совпадающий с J. Тогда алгебра $\mathrm{J}'' = \mathrm{I} + \mathrm{J}'$ является двусторонним идеалом в A. Но I максимален; следовательно, $\mathrm{J}'' = \mathrm{I}$. Отсюда $\mathrm{J}' = (0)$, и лемма доказана.

11.77. Сформулируем, наконец, основную теорему о строении полупростых алгебр.

Теорема. (Вторая структурная теорема.) *Всякая полупростая алгебра разлагается в прямую сумму своих двусторонних идеалов, каждый из которых является простой алгеброй.*

Доказательство. Пусть A — полупростая алгебра. Как показано в *11.71*, для нее можно построить композиционный ряд

$$\mathrm{A} = \mathrm{I}_0 \supset \mathrm{I}_1 \supset \ldots \supset \mathrm{I}_n \supset \mathrm{I}_{n+1} = (0).$$

Наша теорема, очевидно, содержится как частный случай в следующем утверждении:

(*) *Для любого k, $0 \leqslant k \leqslant n$, алгебра I_{n-k} разлагается в прямую сумму своих двусторонних идеалов, являющихся простыми алгебрами, и, кроме того, обладает единицей.*

Утверждение (*) мы будем доказывать индукцией по k.

Алгебра I_n не содержит собственных двусторонних идеалов и, согласно теореме *11.73*, отлична от тривиальной. Следовательно, алгебра I_n проста и, в частности, обладает единицей. Это доказывает утверждение (*) для $k = 0$.

Пусть теперь наше утверждение справедливо для некоторого k, $0 \leqslant k \leqslant n-1$. Это означает, в частности, что алгебра I_{n-k} содержит единицу, а следовательно, по лемме *11.75*, $\mathrm{I}_{n-k-1} = \mathrm{I}_{n-k} + \mathrm{J}$ для некоторого двустороннего идеала J в I_{n-k-1}.

Поскольку I_{n-k} — максимальный двусторонний идеал в I_{n-k-1} алгебра J согласно лемме *11.76* не содержит собственных двусторонних идеалов. В то же время, применив теорему *11.73* к нормальному ряду

$$\mathrm{A} = \mathrm{I}_0 \supset \mathrm{I}_1 \supset \ldots \supset \mathrm{I}_{n-k-1} \supset \mathrm{J} \supset (0),$$

мы получим, что эта алгебра не является тривиальной. Следовательно, она проста.

По индуктивному предположению, алгебра I_{n-k} разлагается в прямую сумму своих простых подалгебр, являющихся в ней двусторонними идеалами. Обладая единицей, эти же подалгебры согласно лемме *11.74* являются двусторонними идеалами также и в I_{n-k-1}. Из этого факта и равенства $\mathrm{I}_{n-k-1} = \mathrm{I}_{n-k} + \mathrm{J}$ немедленно следует, что требуемое разложение в прямую сумму простых подалгебр имеет место и для алгебры I_{n-k-1}.

Осталось доказать существование единицы у алгебры I_{n-k-1}. Пусть e_1 — единица алгебры I_{n-k} (она существует согласно индуктивному предположению), а e_2 — единица простой алгебры J. Поскольку $ab = ba = 0$ для любых $a \in \mathrm{I}_{n-k}$,

$b=\mathrm{J}$, элемент $e=e_1+e_2$ является, как легко видеть, единицей во всей алгебре I_{n-k-1}.

Таким образом, доказана законность индуктивного перехода, а с ней и все утверждение (*). Как уже отмечалось, наша теорема является частным случаем этого утверждения, именно, при $k=n$. Тем самым она доказана.

Заметим, что мы одновременно установили тот факт, что *всякая полупростая алгебра обладает единицей.*

Найденные в теореме двусторонние идеалы, дающие в прямой сумме заданную полупростую алгебру A, мы будем в дальнейшем называть *простыми составляющими алгебры* A.

11.78. В 11.64 было выяснено, что всякая простая алгебра изоморфна алгебре $\mathrm{B}(\mathrm{X})$ для некоторого конечномерного пространства X, или, что то же самое, алгебре всех квадратных матриц некоторого размера.

Пусть теперь $\mathrm{X}_1, \ldots, \mathrm{X}_n$ — некоторый набор конечномерных пространств. Обозначим через $\mathrm{B}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ множество всех строк вида

$$a=(a_1, \ldots, a_n),$$

где a_k — оператор из алгебры $\mathrm{B}(\mathrm{X}_k)$ (или, если угодно, матрица соответствующего размера). Очевидно, $\mathrm{B}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ является алгеброй относительно «покоординатных» операций, определяемых равенствами

$$\begin{aligned} a+b&=(a_1+b_1, \ldots, a_n+b_n),\\ \lambda a&=(\lambda a_1, \ldots, \lambda a_n),\\ ab&=(a_1b_1, \ldots, a_nb_n), \end{aligned}$$

где $a, b \in \mathrm{B}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$, $a=(a_1, \ldots, a_n)$, $b=(b_1, \ldots, b_n)$, λ — комплексное число.

С учетом сказанного теорема *11.77* допускает следующую эквивалентную формулировку:

Всякая полупростая алгебра изоморфна алгебре $\mathrm{B}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ *для некоторого набора пространств* $\mathrm{X}_1, \ldots, \mathrm{X}_n$.

Заметим еще, что простые составляющие алгебры $\mathrm{B}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ состоят, очевидно, из строк вида $(0, \ldots, 0, a_k, 0, \ldots, 0)$, где k-я координата пробегает всю алгебру $\mathrm{B}(\mathrm{X}_k)$, а на остальных местах — нули; мы будем каждую такую составляющую отождествлять с соответствующей алгеброй $\mathrm{B}(\mathrm{X}_k)$.

11.79. В заключение этого параграфа мы найдем все двусторонние идеалы полупростой алгебры.

Теорема. *Всякий двусторонний идеал полупростой алгебры представляет собой прямую сумму некоторого числа ее простых составляющих.*

Пусть A — полупростая алгебра; тогда согласно *11.78* она изоморфна некоторой алгебре вида $B(X_1, \ldots, X_n)$ с простыми составляющими $B(X_k)$, $1 \leqslant k \leqslant n$.

Пусть теперь I — двусторонний идеал в $B(X_1, \ldots, X_n)$. Обозначим через I_k пересечение I с $B(X_k)$. Поскольку вместе с каждым элементом $a = (a_1, \ldots, a_{k-1}, a_k, a_{k+1}, \ldots, a_n)$, $a \in I$, идеал I содержит и строку $ae_k = (0, \ldots, 0, a_k, 0, \ldots, 0)$, где e_k — единица в $B(X_k)$, имеет место разложение в прямую сумму: $I = I_1 + \ldots + I_n$.

Но для любого k, $1 \leqslant k \leqslant n$, I_k как легко видеть, есть двусторонний идеал в простой алгебре $B(X_k)$; следовательно, либо $I_k = (0)$, либо I_k совпадает со всей $B(X_k)$. Отсюда непосредственно вытекает наше утверждение.

§ 11.8. Строение представлений простых и полупростых алгебр

Знание структуры изученных типов алгебр позволяет без особого труда найти все их представления с точностью до эквивалентности.

11.81. Пусть A — полупростая алгебра; ввиду доказанного в предыдущем параграфе ее можно отождествить с алгеброй $B(X_1, \ldots, X_n)$ для некоторого набора пространств X_k, $1 \leqslant k \leqslant n$. Поэтому вместе с заданной алгеброй A естественным образом возникают n ее представлений T^k: $A \to B(X_k)$, $1 \leqslant k \leqslant n$, действующих по формуле $T^k_a = a_k \in B(X_k)$ для любого $a \in A$, $a = (a_1, \ldots, a_k, \ldots, a_n)$. Поскольку образом представления T^k является вся алгебра $B(X_k)$, все эти представления неприводимы.

Теорема. *Всякое неприводимое представление полупростой алгебры* A *эквивалентно одному из представлений* T^k $(1 \leqslant k \leqslant n)$.

Доказательство. Пусть $A = B(X_1, \ldots, X_n)$ — полупростая алгебра; T: $A \to B(X)$ — ее неприводимое представление. Рассмотрим ядро $Z(T)$ представления T. Будучи двусторонним идеалом в A, это ядро по теореме *11.79* является прямой суммой нескольких простых составляющих алгебры A.

Обозначим через A_1 прямую сумму остальных, не вошедших в $Z(T)$ простых составляющих алгебры A, через $T^{(1)}: A_1 \to B(X)$ — сужение на A_1 исходного представления T. Новое представление $T^{(1)}$ уже является точным и, поскольку образы представлений $T^{(1)}$ и T, очевидно, совпадают, является неприводимым. Алгебра A^1, имея представление такого вида, должна быть простой *(11.64)*. Следовательно, она сводится к одному лишь прямому слагаемому, т. е. совпадает с $B(X_k)$ для некоторого k, $1 \leqslant k \leqslant n$. Отсюда, как легко видеть, для любого $a \in A$, $a = (a_1, \ldots, a_k, \ldots, a_n)$,

$$T_a = T^{(1)}_{a_k}, \quad a_k \in B(X_k).$$

Согласно теореме *11.51* все неприводимые представления простой алгебры эквивалентны; в частности, эквивалентны $T^{(1)}: B(X_k) \to B(X)$ и тождественное представление $T^{(2)}: B(X_k) \to B(X_k)$. Это означает существование изоморфизма $U: X \to X_k$ такого, что $UT^{(1)}_{a_k} = T^{(2)}_{a_k}U$ для любого $a_k \in B(X_k)$. Но по доказанному $T_a = T^{(1)}_{a_k}$ для любого $a \in A$; с другой стороны, из определения представления T^k следует, что $T^k_a = T^{(2)}_{a_k}$. Отсюда $UT_a = T^k_a U$ для любого $a \in A$, что и доказывает эквивалентность представлений T и T^k.

11.82. Перейдем к произвольным представлениям простых и полупростых алгебр. Окажется полезным следующее утверждение общего характера.

Лемма. *Пусть* A — *произвольная алгебра*, $T: A \to B(X)$ — *ее представление*, X_k, $1 \leqslant k \leqslant n$, — *минимальные инвариантные подпространства представления* T *такие, что их линейная оболочка совпадает с* X. *Тогда* X *является прямой суммой нескольких подпространств из этого набора.*

Пересечение инвариантных подпространств представления само является инвариантным подпространством. Поэтому из минимальности заданных подпространств следует, что для любого k пересечение подпространства X_{k+1} с линейной оболочкой предшествующих подпространств X_1, X_k есть либо само X_{k+1}, либо нуль. Таким образом, последовательно выбирая из подпространств X_k те, которые не содержатся в линейной оболочке предыдущих, мы получим подпространства, дающие в прямой сумме всю линейную оболочку подпространств X_k, $1 \leqslant k \leqslant n$, т. е. все X. Тем самым лемма доказана.

11.83. Согласно второй структурной теореме всякая полупростая алгебра A изоморфна некоторой алгебре вида $B(X_1, \ldots, X_n)$. В дальнейшем нам будет удобно рассматривать реализацию алгебры $B(X_1, \ldots, X_n)$ в виде алгебры строк из n матриц соответствующей размерности. Для элемента $a \in A$ число, стоящее на «i, j-м» месте в k-й матрице соответствующей строки, мы будем обозначать через $\lambda_{ij}^{(k)}(a)$. Через $e_{ij}^{(k)}$ будем обозначать тот элемент алгебры A, для которого $\lambda_{ij}^{(k)}(e_{ij}^{(k)}) = 1$, а все остальные места в матрицах соответствующей строки заполнены нулями. Заметим, что $\sum\limits_{i,k} e_{ii}^{(k)} = e$, где e — единица алгебры A.

Лемма. *Пусть* $T: A \to B(X)$ — *представление полупростой алгебры* A. *Пусть, далее, для некоторого* $x \in X$ *и некоторых индексов* i *и* k *вектор* $y = T_{e_{ii}^{(k)}} x$ *отличен от нуля. Тогда* y *принадлежит некоторому минимальному инвариантному подпространству представления* T.

Доказательство. Положим $Y = \{T_a y: a \in A\}$. Поскольку $y = T_{e_{ii}^{(k)}} x$, из правила перемножения матриц следует, что любой элемент $z_1 \in Y$ имеет вид $z_1 = T_b x$, где b — некоторая линейная комбинация элементов $e_{ij}^{(k)}$ (i и k фиксированы). Достаточно показать, что в случае $z \neq 0$ вектор z_1 — циклический вектор в сужении представления T на Y.

Пусть $z_2 \in Y$: $z_2 = T_c x$, где c — другая линейная комбинация тех же элементов. Используя реализацию алгебры A в виде алгебры строк матриц, найдем элемент $a \in A$ такой, что $c = ab$. Но тогда $z_2 = T_c x = T_a(T_b x) = T_a z_1$. Таким образом, вектор z_1 циклический, и лемма доказана.

11.84. Теорема. *Всякое представление полупростой алгебры разлагается в прямую сумму неприводимых представлений и тривиального представления.*

Пусть A — полупростая алгебра, T^0: $A \to B(X^0)$ — ее представление. Рассмотрим оператор T_e^0, где e — единица в A. Для любого $x \in X^0$ равенство

$$x = T_e^0 x + \left(x - T_e^0 x\right)$$

определяет, очевидно, разложение X^0 в прямую сумму подпространств X и X_0, инвариантных относительно T^0, причем сужение представления T^0 на X_0 является тривиальным представлением. Нам осталось доказать, что представление $T: A \longrightarrow B(X)$ — сужение представления T^0 на X — разлагается в прямую сумму неприводимых.

Выберем в X базис $x_1, \ldots, x_m$. Оператор T_e — тождественный оператор в X, поэтому ввиду равенства $e = \sum_{i, k} e_{ii}^{(k)}$ линейная оболочка векторов вида $T_{e_{ii}^{(k)}} x_j$ (по всем возможным индексам i, j, k) совпадает со всем X. По лемме *11.83* каждый отличный от нуля вектор такого вида лежит в некотором минимальном инвариантном подпространстве представления T.

Итак, мы находимся в условиях леммы *11.82*. Но тогда пространство X разлагается в прямую сумму минимальных инвариантных подпространств представления T, и, следовательно, само T разлагается в прямую сумму неприводимых представлений.

Тем самым доказательство теоремы закончено.

11.85. Теоремы *11.81* и *11.84* вместе описывают с точностью до эквивалентности все представления полупростых (в том числе и простых) алгебр.

Полученный результат показывает, в частности, что операторы заданного представления простой алгебры (этот случай мы выделяем для большей наглядности), записываются в некотором базисе матрицами вида

$$\left\| \begin{array}{ccc:c} M & 0 & & \\ & \ddots & & 0 \\ 0 & & M & \\ \hdashline & 0 & & 0 \end{array} \right\|, \tag{3}$$

где M пробегает всю совокупность матриц соответствующего размера. В более общем случае полупростой алгебры соответствующие матрицы суть матрицы вида

$$\left\|\begin{array}{ccccccccc} M_1 & & & & & & & & \\ & \ddots & & & & 0 & & & \\ & & M_1 & & & & & & \\ & & & \ddots & & & & & \\ & & & & M_k & & & & \\ & 0 & & & & \ddots & & & \\ & & & & & & M_k & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{array}\right\|, \tag{4}$$

где стоящие в блоках матрицы $M_1, \ldots, M_k, \ldots$ пробегают независимо друг от друга всю совокупность матриц соответствующего размера (вообще говоря, разного для разных матриц).

11.86. Попутно *нами получено описание всех простых и полупростых матричных алгебр*, т. е. алгебр, которые сами состоят из матриц. Действительно, поставив в соответствие каждой матрице из такой алгебры задаваемый ею оператор (в любом базисе), мы получим тем самым точное представление этой алгебры. Отсюда с учетом предыдущих рассуждений немедленно вытекает следующее утверждение:

Всякая простая (*соответственно полупростая*) *матричная алгебра состоит из матриц вида $P^{-1}\,LP$, где P — некоторая фиксированная невырожденная матрица, а L пробегает совокупность матриц вида* (3) (*соответственно вида* (4)).

Для алгебр с единицей результат получается еще более точным:

Всякая простая матричная алгебра с единицей состоит из матриц вида $P^{-1}LP$, где P — фиксированная невырожденная *матрица, а L — пробегает совокупность всех матриц вида*

$$\left\| \begin{matrix} M & 0 & \dots & 0 \\ 0 & M & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & M \end{matrix} \right\|,$$

где M пробегает всю совокупность матриц соответствующего размера. Всякая полупростая матричная алгебра с единицей состоит из матриц вида $P^{-1}LP$, где P — фиксированная невырожденная матрица, а L пробегает совокупность всех матриц вида

$$\left\| \begin{array}{c|c|c} \begin{matrix} M_1 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & M_1 \end{matrix} & & \\ \hline & \ddots & \\ \hline & & \begin{matrix} M_k & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & M_k \end{matrix} \end{array} \right\|,$$

где матрицы $M_1, \dots, M_k$ пробегают независимо друг от друга всю совокупность матриц соответствующих размеров.

§ 11.9. Некоторые дальнейшие результаты

Итак, мы закончили описание простых и полупростых конечномерных алгебр, а также их представлений. Дальнейшее исследование конечномерных алгебр уже выходит за рамки настоящей главы. Тем не менее полезно для перспективы привести некоторые известные факты в этом направлении.

11.91. Теорема Веддерберна. *Всякая конечномерная алгебра разлагается, как линейное пространство, в прямую сумму своего радикала и некоторой полупростой подалгебры* (см., например, Н. Джекобсон, Теория колец, ИЛ, 1947, стр. 220).

11.92. Радикал конечномерной алгебры состоит только из нильпотентных элементов. Более того, для каждой такой алгебры существует натуральное n такое, что произведение любых n элементов ее радикала равно нулю (см., например, Н. Г. Чеботарев, Введение в теорию алгебр, Гостехиздат, 1949, § 8).

11.93. Всякое представление радикальной алгебры записывается в некотором базисе матрицами с нулями на главной диагонали и ниже ее. (При этом, конечно, не утверждается, что матрицы операторов представления пробегают всю совокупность матриц такого вида; см., например, А. Я. Хелемский, Об алгебрах нильпотентных операторов и связанных с ними категориях, Вестник МГУ, 1963, № 4, стр. 49—55.)

ЗАДАЧИ

1. Доказать, что всякий левый идеал алгебры $B(K_n)$ есть совокупность всех операторов, нуль-многообразие которых содержит некоторое подпространство $K' \subset K_n$.

2. Доказать, что всякий правый идеал I алгебры $B(K_n)$ есть совокупность всех операторов, область значений которых лежит в некотором подпространстве $K' \subset K_n$.

3. Указать все максимальные и минимальные левые и правые идеалы алгебры $B(K_n)$.

4. Для всякой полупростой алгебры B линейных операторов над пространством C_n ввести в C_n скалярное произведение так, чтобы из $A \in B$ следовало $A^* \in B$.

5. (Обращение задачи 4.) Если для некоторой алгебры B линейных операторов над пространством C_n существует такое скалярное произведение (x, y) в C_n, что из $A \in B$ следует $A^* \in B$, то алгебра B полупростая.

6. Если, при выполнении условий задачи 5, коммутатор $\overline{B}$ *(11.61)* пересекается с самой алгеброй B лишь по операторам, кратным единичному, то B — простая алгебра.

7. Доказать, что коммутатор простой алгебры B, состоящей из матриц вида *11.85*, распадающихся на m^2 блоков

$$\begin{Vmatrix} A & 0 & \ldots & 0 \\ 0 & A & \ldots & 0 \\ . & . & . & . \\ 0 & 0 & \ldots & A \end{Vmatrix},$$

представляется (в том же базисе) всеми матрицами вида

$$\left\|\begin{matrix} \lambda_{11}E & \lambda_{12}E & \ldots & \lambda_{1m}E \\ \lambda_{21}E & \lambda_{22}E & \ldots & \lambda_{2m}E \\ \cdot & \cdot & \cdot & \cdot \\ \lambda_{m1}E & \lambda_{m2}E & \ldots & \lambda_{mm}E \end{matrix}\right\|,$$

где λ_{jk} $(j,\ k=1,\ \ldots,\ m)$ — произвольные комплексные числа. В частности, пересечение B с $\overline{\mathrm{B}}$ состоит лишь из матриц, кратных единичной.

8. У какой полупростой матричной алгебры B ее коммутатор $\overline{\mathrm{B}}$ совпадает с самой B?

9. Описать все коммутативные полупростые алгебры $(\mathrm{B}\subset\overline{\mathrm{B}})$.

10. Описать все полупростые матричные алгебры B, для которых $\overline{\mathrm{B}}\subset\mathrm{B}$.

11. Доказать, что для полупростой алгебры всегда $\overline{\overline{\mathrm{B}}}=\mathrm{B}$.

12. Алгебра B состоит из всех многочленов от одного оператора A (и, следовательно, коммутативна, так что $\overline{\mathrm{B}}\supset\mathrm{B}$). Дать критерий равенства $\overline{\mathrm{B}}=\mathrm{B}$.

13. Если алгебра $\mathrm{B}\neq 0$ состоит из нильпотентных элементов (т. е. для каждого $\mathrm{A}\in\mathrm{B}$ имеем $\mathrm{A}^k=0$ при некотором $k=k(\mathrm{A})$), то равенство $c\mathrm{B}=\mathrm{B}$ невозможно ни при каком $c\in\mathrm{B}$.

14. Алгебра B называется нильпотентной, если существует число p такое, что произведение любых p ее элементов равно 0. Показать, что алгебра B, являющаяся суммой своих правых идеалов $\mathrm{B}_1+\ldots+\mathrm{B}_m$, нильпотентна, если нильпотентен каждый идеал B_j $(j=1,\ \ldots,\ m)$.

15. Если конечномерная алгебра B состоит из нильпотентных элементов, то она сама нильпотентна.

16. Пусть B — нильпотентная алгебра операторов в пространстве K_n. Пусть $M_1\subset\mathrm{K}_n$ есть пересечение всех нуль-многообразий всех операторов $\mathrm{A}\in\mathrm{B}$, далее, $M_2\subset\mathrm{K}_n$ — пересечение всех подпространств, которые операторами $\mathrm{A}\in\mathrm{B}$ переводятся в M_1, $M_3\subset\mathrm{K}_n$ — пересечение всех подпространств, которые операторами $\mathrm{A}\in\mathrm{B}$ переводятся в M_2, и т. д. Показать, что имеют место строгие включения

$$0\subset M_1\subset M_2\subset\ldots\subset M_p=\mathrm{K}_n,$$

где p — индекс нильпотентности алгебры B, т. е. наименьшее число p такое, что произведение любых p операторов из B равно 0.

17. Для всякой нильпотентной алгебры B операторов в пространстве K_n существует базис, в котором все операторы $\mathrm{A}\in\mathrm{B}$ записываются матрицами вида

$$A=\left\|\begin{matrix} 0 & A_{12} & A_{13} & \ldots & A_{1,\,p-1} \\ 0 & 0 & A_{23} & \ldots & A_{2,\,p-1} \\ 0 & 0 & 0 & \ldots & A_{3,\,p-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 0 \end{matrix}\right\|,$$

где p — индекс нильпотентности алгебры B (А. Я. Хелемский).

КАТЕГОРИИ КОНЕЧНОМЕРНЫХ ПРОСТРАНСТВ

§ 12.1. Введение

12.11. В последнее время в различных областях математики стали играть большую роль понятие категории и связанные с ним понятия*). Некоторая совокупность множеств с их отображениями друг в друга, некоторая совокупность линейных пространств или алгебр с их морфизмами — все это примеры категорий.

Точное определение категории следующее. Пусть имеется некоторое множество $\mathfrak{A}$ индексов α. Пусть, далее, имеется некоторая совокупность $\mathfrak{K}$ элементов X_α, называемых *объектами категории* $\mathfrak{K}$, и для каждой пары объектов X_β, X_α указано множество $\mathfrak{B}_{\beta\alpha}$ других элементов $A_{\beta\alpha}$, называемых *отображениями* объекта X_α в объект X_β. Предполагается, что при любых α, β, γ определено произведение имеющихся отображений $A_{\alpha\beta}$ и $A_{\beta\alpha}$, которое является отображением X_α в X_γ; при этом требуется, чтобы умножение отображений было ассоциативным, т. е. для любых α, β, γ, δ

$$A_{\delta\gamma}(A_{\gamma\beta}A_{\beta\alpha}) = (A_{\delta\gamma}A_{\gamma\beta})\,A_{\beta\alpha}.$$

В частности, определено и множество $\mathfrak{B}_{\alpha\alpha}$ отображений объекта X_α в себя самого. В этом множестве, следовательно, определено (ассоциативное) умножение отображений. Кроме того, требуется, чтобы множество $\mathfrak{B}_{\alpha\alpha}$ содержало единичный элемент 1_α, который обладает тем свойством, что

$$1_\alpha A_{\alpha\beta} = A_{\alpha\beta}, \quad A_{\gamma\alpha} 1_\alpha = A_{\gamma\alpha}$$

*) См., например, А. Картан и С. Эйленберг. Гомологическая алгебра, ИЛ, М., 1960; А. Гротендик. О некоторых вопросах гомологической алгебры, ИЛ, М., 1961; А. Г. Курош и др. Основы теории категорий, УМН, № 6, 3—52, 1960.

при любых α, β и γ. Будем далее вместо $\mathfrak{B}_{\alpha\alpha}$ писать короче $\mathfrak{B}_{\alpha}$.

Совокупность объектов X_α и отображений $A_{\beta\alpha}$, обладающих перечисленными свойствами, и образует, по определению, *категорию*.

Категория $\mathfrak{K}$ называется *линейной*, если в совокупности $\mathfrak{B}_{\beta\alpha}$ отображений $A_{\beta\alpha}$ (с любыми фиксированными α и β) определены операции сложения отображений и умножения их на числа (из поля K), превращающие совокупность $\mathfrak{B}_{\beta\alpha}$ в линейное пространство над полем K. В линейной категории совокупность $\mathfrak{B}_{\alpha}$ представляет собой алгебру с единицей (над полем K).

12.12. Мы рассмотрим в этой главе линейные категории, элементами которых являются линейные конечномерные пространства над полем C комплексных чисел, а отображениями — линейные отображения (морфизмы) одного такого пространства в другое.

Итак, мы исходим из следующего определения.

Имеется некоторое множество линейных конечномерных комплексных пространств X_α $(\alpha \in \mathfrak{A})$. При каждом α в пространстве X_α задана алгебра $\mathfrak{B}_\alpha$ линейных операторов, переводящих X_α в себя. Для каждой пары индексов β, α задано семейство $\mathfrak{B}_{\beta\alpha}$ линейных операторов $A_{\beta\alpha}$, переводящих X_α в X_β, содержащее вместе с каждыми двумя операторами $A_{\beta\alpha}$ и $B_{\beta\alpha}$ их сумму $A_{\beta\alpha}+B_{\beta\alpha}$ и вместе со всяким оператором $A_{\beta\alpha}$ его произведение $\lambda A_{\beta\alpha}$ на любое комплексное число λ; такое семейство линейных операторов будем в дальнейшем называть *линейным семейством*. В частности, линейное семейство $\mathfrak{B}_{\alpha\alpha}$ совпадает с алгеброй $\mathfrak{B}_\alpha$. Предполагается, что для любых α, β, γ выполнено условие

$$\mathfrak{B}_{\gamma\beta}\mathfrak{B}_{\beta\alpha} \subset \mathfrak{B}_{\gamma\alpha}, \tag{1}$$

т. е. любое произведение $A_{\gamma\beta}A_{\beta\alpha}$ $(A_{\gamma\beta} \in \mathfrak{B}_{\gamma\beta},\ A_{\beta\alpha} \in \mathfrak{B}_{\beta\alpha})$ лежит в $\mathfrak{B}_{\gamma\alpha}$. Совокупность пространств X_α с алгебрами $\mathfrak{B}_\alpha$ и семействами $\mathfrak{B}_{\beta\alpha}$ будем называть *категорией конечномерных пространств*, или просто *категорией*, и обозначать через $\mathfrak{K}$.

Если в каждом пространстве X_α выбран как-нибудь базис, то алгебры $\mathfrak{B}_\alpha$ и линейные семейства $\mathfrak{B}_{\beta\alpha}$ можно отождествить соответственно с алгебрами и линейными семействами соответствующих матриц, что мы в дальнейшем будем систематически использовать.

В этой главе мы выясняем, какими могут быть категории линейных пространств при заданных алгебрах $\mathfrak{B}_\alpha$. Мы ограничиваемся полупростыми алгебрами $\mathfrak{B}_\alpha$. В силу *11.86* для полупростой алгебры пространство X_α может быть разбито в прямую сумму подпространств $X_{\alpha j}$, инвариантных относительно всех операторов $A_{\beta\alpha}$, причем в каждом подпространстве $X_{\alpha j}$ алгебра $\mathfrak{B}_\alpha$ есть простая алгебра с единицей, т. е. в некотором базисе она записывается матрицами вида

$$\begin{Vmatrix} C & & & \\ & C & & \\ & & \ddots & \\ & & & C \end{Vmatrix},$$

где C пробегает всю совокупность матриц соответствующего размера.

Мы начинаем с разбора нескольких частных случаев, которые позволят затем сформулировать и общие результаты. В § 12.2 рассматривается случай, когда каждая алгебра $\mathfrak{B}_\alpha$ есть полная алгебра, т. е. алгебра всех линейных операторов, действующих в X_α. Противоположный случай, когда каждая из этих алгебр есть алгебра операторов $\{\lambda X\}$, кратных единичному, разобран в § 12.3. Результаты § 12.4 относятся к случаю простых алгебр $\mathfrak{B}_\alpha$, который является естественным обобщением случая алгебр $\{\lambda X\}$. В § 12.5 речь идет о случае, когда каждая алгебра $\mathfrak{B}_\alpha$ есть алгебра всех диагональных матриц. В § 12.6 общая категория приводится к разобранным в предыдущих параграфах.

12.13. Напомним обозначения и правила действий с матрицами линейных операторов, отображающих линейное пространство X в линейное пространство Y (*4.41—4.43*). Пусть имеется n-мерное пространство X с базисом $e_1, \ldots, e_n$ и m-мерное пространство Y с базисом $f_1, \ldots, f_m$. Линейному оператору A, действующему из X в Y, мы ставим в соответствие $m \times n$-матрицу (т. е. матрицу из m строк и

n столбцов)

$$A=\left\|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{matrix}\right\|,$$

где в k-м столбце стоят числа $a_{1k}, a_{2k}, \ldots, a_{mk}$, представляющие собой коэффициенты разложения вектора $\mathrm{A}e_k \in \mathrm{Y}$ по базису $f_1, \ldots, f_m$.

Пусть, далее, имеется k-мерное пространство Z с базисом $g_1, \ldots, g_k$. Оператору B, действующему из пространства Y в пространство Z, ставится в соответствие $k\times m$-матрица

$$B=\left\|\begin{matrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ b_{k1} & b_{k2} & \dots & b_{km} \end{matrix}\right\|.$$

Оператор $\mathrm{C}=\mathrm{BA}$ действует из X в Z. Ему отвечает $k\times n$-матрица

$$C=\left\|\begin{matrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{k1} & c_{k2} & \dots & c_{kn} \end{matrix}\right\|.$$

Эта матрица получается из матриц B и A с помощью операции умножения:

$$c_{pq}=\sum_{j=1}^{m} b_{pj}a_{jq}, \qquad p=1, \ldots, k;\ q=1, \ldots, n.$$

12.14. Напомним следующий факт, который будет часто полезным (*4.44*).

Лемма. *Пусть дана прямоугольная* $n\times m$-*матрица* $A=\|a_{pq}\|$ (n строк, m столбцов; $p=1, \ldots, n$; $q=1, \ldots, m$). *Если умножить матрицу* A *справа на прямоугольную* $m\times k$-*матрицу* $B=\|b_{rs}\|$, *где все элементы* b_{rs}, *кроме* $b_{r_0s_0}$, *равны* 0, *а* $b_{r_0s_0}=1$, *то получится прямоугольная* $n\times k$-*матрица* F, *в которой все столбцы, кроме* s_0-*го, будут равны* *нулю, а в s_0-м будут стоять элементы r_0-го столбца матрицы A. Если матрицу A умножить* с л е в а *на прямоугольную $l\times n$-матрицу $C=\|c_{rs}\|$, где все элементы c_{rs}, кроме $c_{r_1s_1}$, равны нулю, а $c_{r_1s_1}=1$, то получится прямоугольная $l\times m$-матрица, все строки которой, кроме r-й, будут равны 0, тогда как в r_1-й строке будут стоять элементы s_1-й строки матрицы A.*

12.15. Как следствие получаем: если прямоугольную $n\times m$-матрицу A умножить справа на $m\times k$-матрицу B, а слева на $l\times n$-матрицу C с указанными свойствами, то в полученной $l\times k$-матрице D будет отличен от нуля (возможно) только один элемент, именно, стоящий на скрещении r_1-й строки и s_0-го столбца, а этим элементом будет элемент $a_{s_1r_0}$ исходной матрицы A (т. е. находившийся в ней на скрещении s_1-й строки и r_0-го столбца).

§ 12.2. Случай, когда все данные алгебры $\mathfrak{B}_\alpha$ — полные

12.21. Пусть категория $\mathfrak{K}$ состоит из конечномерных линейных пространств X_α, причем алгебра $\mathfrak{B}_\alpha$ операторов, действующих в X_α, при любом α есть полная алгебра, т. е. алгебра *всех* линейных операторов в X_α. Фиксируя произвольно базис $e_1, \ldots, e_n$ в пространстве X_1 и $f_1, \ldots, f_n$ в пространстве X_2, мы можем отождествить операторы из совокупностей $\mathfrak{B}_{11}$, $\mathfrak{B}_{12}$, $\mathfrak{B}_{21}$, $\mathfrak{B}_{22}$ с соответствующими им матрицами.

Пусть размерность пространства X_1 равна m и размерность пространства X_2 равна n. Предположим, что в совокупности $\mathfrak{B}_{21}$ имеется ненулевой оператор A и ему соответствует прямоугольная $n\times m$-матрица $A=\|a_{pq}\|$, в которой по крайней мере один элемент, например $a_{p_0q_0}$, отличен от нуля. В этом случае можно считать $a_{p_0q_0}=1$. По условию (1) и нашему предположению ($\mathfrak{B}_1$ и $\mathfrak{B}_2$ суть полные алгебры матриц), произведение матрицы A справа на любую квадратную $m\times m$-матрицу и слева на любую квадратную $n\times n$-матрицу снова принадлежит семейству $\mathfrak{B}_{21}$. Однако в силу леммы *12.14* в результате такой операции мы всегда можем получить $n\times m$-матрицу с единственным отличным от нуля элементом 1 в любом наперед заданном месте. Так как линейная комбинация таких матриц дает любую $n \times m$-матрицу, то $\mathfrak{B}_{21}$ содержит все эти матрицы, т. е. *представляет собой полную совокупность операторов, действующих из* X_1 в X_2.

12.22. Мы свяжем сейчас рассматриваемую категорию с некоторым *частично упорядоченным множеством*. Введем соответствующее определение: множество S называется *частично упорядоченным,* если для некоторых пар A, B его элементов установлено соотношение, обозначаемое знаком $\leqslant$ (меньше или равно), причем выполнены следующие аксиомы:

а) Если $A \leqslant B$ и $B \leqslant A$, то $A = B$.

б) Если $A \leqslant B$, $B \leqslant C$, то $A \leqslant C$.

в) Всякое $A \leqslant A$.

Несколько более обшим понятием является понятие *пред-частично упорядоченного множества*. Мы будем так называть множество с отношением $\leqslant$, удовлетворяющим лишь аксиомам б) и в). Если при этом $A \leqslant B$ и $B \leqslant A$, то A и B называются *эквивалентными,* что обозначается знаком $A \sim B$. Проверим, что из $A \sim B$ и $B \sim C$ следует $A \sim C$. Действительно, из $A \leqslant B$, $B \leqslant C$ следует, по б), также $A \leqslant C$ и из $C \leqslant B$, $B \leqslant A$ следует $C \leqslant A$, так что $A \sim C$. Поэтому отношение $\sim$ позволяет разбить все множество S на классы $\mathfrak{A}$, $\mathfrak{B}$, ... ; каждый класс содержит вместе с элементом A все ему эквивалентные, и элементы A, B, принадлежащие к разным классам, не эквивалентны.

Теперь введем отношение $\leqslant$ уже для классов $\mathfrak{A}$, $\mathfrak{B}$. Именно, будем считать, что $\mathfrak{A} \leqslant \mathfrak{B}$, если существуют $A \in \mathfrak{A}$, $B \in \mathfrak{B}$ такие, что $A \leqslant B$. Это определение не зависит от выбора элементов $A \in \mathfrak{A}$ и $B \in \mathfrak{B}$; действительно, пусть, кроме того, $A_1 \in \mathfrak{A}$, $B_1 \in \mathfrak{B}$, так что $A \sim A_1$, $B \sim B_1$. Тогда мы имеем $A_1 \leqslant A \leqslant B \leqslant B_1$, откуда $A_1 \leqslant B_1$, что и требуется. Выполнение свойств б) и в) частично упорядоченного множества для классов $\mathfrak{A}$, $\mathfrak{B}$, ... теперь следует из выполнения этих свойств для элементов A, B, ... Покажем, что для классов $\mathfrak{A}$, $\mathfrak{B}$, ... выполнено и свойство а). Пусть $\mathfrak{A} \leqslant \mathfrak{B}$ и $\mathfrak{B} \leqslant \mathfrak{A}$; выберем произвольно $A \in \mathfrak{A}$ и $B \in \mathfrak{B}$. Мы имеем $A \leqslant B$ и $B \leqslant A$, так что A и B эквивалентны. Но тогда классы $\mathfrak{A}$ и $\mathfrak{B}$ совпадают, $\mathfrak{A} = \mathfrak{B}$, что и требовалось.

Итак, введя в пред-частично упорядоченное множество S, как указано, отношение эквивалентности, мы приходим к *частично упорядоченному множеству классов эквивалентных элементов.*

12.23. Возвращаемся к рассмотрению нашей категории. Из *12.21* следует, что для пары пространств X_1 и X_2 мы имеем четыре возможности: а) $\mathfrak{B}_{12}$ и $\mathfrak{B}_{21}$—обе суть полные совокупности операторов; б) $\mathfrak{B}_{12}$ есть полная совокупность, $\mathfrak{B}_{21}$ состоит только из нуля; в) $\mathfrak{B}_{21}$ есть полная совокупность, $\mathfrak{B}_{12}$ состоит только из нуля; г) $\mathfrak{B}_{12}$ и $\mathfrak{B}_{21}$ состоят только из нуля.

Если $\mathfrak{B}_{12}$ есть полная совокупность, а относительно $\mathfrak{B}_{21}$ не делается никаких предположений, будем писать $X_1 \leqslant X_2$. Аналогичный смысл имеет запись $X_2 \leqslant X_1$.

Покажем, что введенное таким образом отношение $\leqslant$ превращает категорию $\mathfrak{K}$ в пред-частично упорядоченное множество. Действительно, для данного X_1 мы имеем по условию, что $\mathfrak{B}_{11}$—полная совокупность операторов, поэтому $X_1 \leqslant X_1$. Далее, если $X_1 \leqslant X_2$ и $X_2 \leqslant X_3$, то $\mathfrak{B}_{12}$ и $\mathfrak{B}_{23}$ являются полными совокупностями линейных операторов, действующих из X_1 в X_2 и соответственно из X_2 в X_3. Так как все наши пространства имеют размерность $\geqslant 1$, то в совокупности $\mathfrak{B}_{13}$ имеется оператор, отличный от нулевого. Такой оператор можно получить, фиксируя ненулевые векторы $e_1 \in X_1$, $e_2 \in X_2$, $e_3 \in X_3$, как произведение операторов AB, где $A \in \mathfrak{B}_{12}$ переводит e_2 в e_1, а $B \in \mathfrak{B}_{23}$ переводит e_3 в e_2. По *12.21* совокупность $\mathfrak{B}_{13}$ есть полная совокупность операторов, переводящих X_3 в X_1, так что $X_1 \leqslant X_3$. Итак, аксиомы б) и в) выполнены, категория $\mathfrak{K}$ превращена в пред-частично упорядоченное множество.

12.24. В соответствии с 12.22 введем отношение эквивалентности, считая, что $X_1 \sim X_2$, если $X_2 \leqslant X_1$, $X_1 \leqslant X_2$, т. е. если обе совокупности $\mathfrak{B}_{12}$ и $\mathfrak{B}_{21}$—полные совокупности соответствующих линейных операторов. Тогда совокупность пространств X_α разбивается на классы взаимно эквивалентных пространств, и совокупность этих классов, с указанным в *12.22* отношением $\leqslant$, становится частично упорядоченным множеством.

Обратно, *любое частично упорядоченное множество $\mathfrak{X}_\alpha$ классов конечномерных пространств определяет категорию рассматриваемого типа.* Именно, для пространств X_1 и X_2, входящих в один и тот же класс, мы задаем совокупности $\mathfrak{B}_{12}$ и $\mathfrak{B}_{21}$, как полные совокупности операторов; для пространств X_1 и X_3, входящих в классы $\mathfrak{X}_1$ и $\mathfrak{X}_3$, связанные отношением $\mathfrak{X}_1 < \mathfrak{X}_3$ (т. е. $\mathfrak{X}_1 \leqslant \mathfrak{X}_3$, но $\mathfrak{X}_1 \neq \mathfrak{X}_3$), мы задаем $\mathfrak{B}_{13}$, как полную совокупность, и $\mathfrak{B}_{31}$, как состоящую только из нуля. Если же X_1 и X_4 входят в несравнимые классы $\mathfrak{X}_1$ и $\mathfrak{X}_4$, то мы задаем $\mathfrak{B}_{14}$ и $\mathfrak{B}_{41}$, как состоящие только из нуля.

Этим описание категорий рассматриваемого типа закончено.

§ 12.3. Все данные алгебры $\mathfrak{B}_\alpha$—одномерные

12.31. Рассмотрим вначале два простых примера.

а. *Категория* $\mathfrak{K}_1$ состоит из двух пространств X_1 и X_2 одинаковой размерности. Система $\mathfrak{B}_{21}$ представляет собой один оператор A, отображающий взаимно однозначно X_1 на X_2, со всеми его кратными λA, $\lambda \in C$. Система $\mathfrak{B}_{12}$ представляет собой один оператор B, обратный к A, со всеми его кратными μB, $\mu \in C$. Очевидно, что

$$\mathfrak{B}_{12}\mathfrak{B}_{21} = (\lambda X), \quad \mathfrak{B}_{21}\mathfrak{B}_{12} = (\lambda X).$$

б. *Категория* $\mathfrak{K}_2$ состоит из двух произвольных пространств X_1 и X_2, в которых фиксированы подпространства $X_1' \subset X_1$ и $X_2' \subset X_2$. Система $\mathfrak{B}_{21}$ состоит из всех операторов, переводящих X_1 в X_2' и при этом X_1' в (0). Система $\mathfrak{B}_{12}$ состоит из всех операторов, переводящих X_2 в X_1' и одновременно X_2' в (0). Очевидно, что $\mathfrak{B}_{12}\mathfrak{B}_{21} = 0$, $\mathfrak{B}_{21}\mathfrak{B}_{12} = 0$.

Мы покажем, что категориями $\mathfrak{K}_1$ и $\mathfrak{K}_2$ в основном исчерпываются все категории из двух пространств с $\mathfrak{B}_j = (\lambda E)$ $(j = 1, 2)$, а именно, для любой категории $\mathfrak{K}$ имеет место альтернатива: *или* $\mathfrak{B}_{12}\mathfrak{B}_{21} = 0$ — *в этом случае и* $\mathfrak{B}_{21}\mathfrak{B}_{12} = 0$, *и категория* $\mathfrak{K}$ *содержится в некоторой категории типа* $\mathfrak{K}_2$, *или размерности пространств* X_1 *и* X_2 *равны и категория* $\mathfrak{K}$ *есть категория типа* $\mathfrak{K}_1$.

12.32. Пусть дана некоторая категория $\mathfrak{K}$ из двух пространств X_1 и X_2 с условиями $\mathfrak{B}_1 = (\lambda X)$, $\mathfrak{B}_2 = (\lambda X)$.

Обозначим через $N_1 \subset X_1$ пересечение нуль-многообразий (*4.62*) всех операторов $A_{12} \in \mathfrak{B}_{12}$ и через $N_2 \subset X_2$ пересечение нуль-многообразий всех операторов $A_{21} \in \mathfrak{B}_{21}$. Если

$\mathfrak{B}_{21}X_1 \subset N_2$ и $\mathfrak{B}_{12}X_2 \subset N_1$, то мы имеем дело с подкатегорией категории типа $\mathfrak{K}_2$, в которой $X_1' = N_1$, $X_2' = N_2$. Поэтому предположим, что, например, $\mathfrak{B}_{21}X_1$ не содержится в N_2 и, следовательно, есть вектор $x_1 \in X_1$ и оператор $A_{21} \in \mathfrak{B}_{21}$, для которого $A_{21}x = x_2$ не принадлежит N_2.

Покажем, что все операторы $B_{21} \in \mathfrak{B}_{21}$ переводят вектор x_1 в вектор, коллинеарный x_2, и все операторы $C_{12} \in \mathfrak{B}_{12}$ переводят x_2 в вектор, коллинеарный x_1.

Действительно, пусть $A_{21}x_1 = x_2$, $B_{21}x_1 = y_2$. Рассмотрим оператор $C_{12}^0 \in \mathfrak{B}_{12}$,для которого $C_{12}^0 x_2 \neq 0$. Тогда в силу основного условия $C_{12}^0 x_2 = C_{12}^0 A_{21} x_1 = \lambda x_1$, причем $\lambda \neq 0$. Заменяя C_{12}^0 его кратным, можем считать, что $\lambda = 1$. Далее, $B_{21}C_{12}^0 x_2 = B_{21}x_1 = y_2$, и в то же время $B_{21}C_{12}^0 x_2 = \mu x_2$; таким образом, $y_2 = \mu x_2$. Так как, обратно, $x_1 = C_{12}^0 x_2$ и, по определению вектора x_1, он не содержится в N_1, мы имеем аналогично $C_{12}x_2 = \mu x_1$ при любом $C_{12} \in \mathfrak{B}_{12}$.

Покажем, что в данном случае N_1 и N_2 сводятся к нуль-вектору. Пусть $z_1 \in N_1$; тогда $A_{21}(x_1 + z_1) = A_{21}x_1 = x_2$, т. е. вектор x_1 в предыдущем построении может быть заменен на $x_1 + z_1$. В таком случае $C_{12}^0 x_2$ есть кратное и x_1 и $x_1 + z_1$, так что x_1 и z_1 коллинеарны. Так как $x_1 \notin N_1$, то $z_1 = 0$. Итак, $N_1 = (0)$. Аналогично, начиная с x_2, получаем, что и $N_2 = (0)$.

Теперь мы видим, что в качестве x_1 можно было взять любой ненулевой вектор пространства X_1, поскольку всегда имеется оператор $A_{21} \in \mathfrak{B}_{21}$, переводящий x_1 не в нуль. Следовательно, операторы семейства $\mathfrak{B}_{21}$ устанавливают взаимно однозначное соответствие между всеми прямыми пространства X_1 и некоторой совокупностью прямых пространства X_2; более того, даже с совокупностью всех прямых пространства X_2 в силу симметрии нашего построения.

Теперь покажем, что вся совокупность $\mathfrak{B}_{21}$ сводится к семейству кратных одного оператора. Пусть $x_1 \neq 0$ есть произвольный вектор в пространстве X_1 и x_2 — ненулевой вектор, определяющий соответствующую x_1 прямую в пространстве X_2. Мы знаем, что имеется оператор A_{21}^0, переводящий x_1 точно в x_2. Всякий другой оператор $A_{21} \in \mathfrak{B}_{21}$ переводит x_1 в λx_2. Пусть сначала $A_{21}x_1 = \lambda x_2$ и $\lambda \neq 0$. Тогда оператор $B_{21} = \frac{1}{\lambda} A_{21}$ переводит x_1 точно в x_2. Покажем, что B_{21} совпадает с A_{21}^0 *всюду*. Пусть $A_{21}^0 y_1 = y_2$, $B_{21}y_1 = z_2 \neq y_2$; это возможно или при $y_2 \neq 0$, $z_2 = \mu y_2$, $\mu \neq 1$, или при $y_2 = 0$, $z_2 \neq 0$. Рассмотрим ненулевой вектор $z_1 = \alpha x_1 + \beta y_1$, $\alpha \neq 0$, $\beta \neq 0$. Векторы $A^0_{21} z_1$ и $B_{21} z_1$, по доказанному, коллинеарны; но в данном случае это не может иметь места, поскольку при $y_2 \neq 0$

$$A^0_{21}(\alpha x_1 + \beta y_1) = \alpha x_2 + \beta y_2, \qquad B_{21}(\alpha x_1 + \beta y_1) = \alpha x_2 + \beta \mu y_2,$$

а при $y_2 = 0$

$$A^0_{21}(\alpha x_1 + \beta y_1) = \alpha x_2, \quad B_{21}(\alpha x_1 + \beta y_1) = \alpha x_2 + \beta z_2.$$

Таким образом, если $A_{21} x_1 = \lambda x_2$, $\lambda \neq 0$, то $A_{21} = \lambda A^0_{21}$. Пусть теперь $A_{21} x_1 = 0$; тогда, по доказанному, $A^0_{21} + A_{21} = A^0_{21}$ и, следовательно, $A_{21} = 0$. Итак, семейство $\mathfrak{B}_{21}$ сводится к семейству кратных оператора A^0_{21}. Аналогично и семейство $\mathfrak{B}_{12}$ сводится к семейству кратных фиксированного оператора B_{12}. Произведения $A^0_{21} B^0_{12}$ и $B^0_{12} A^0_{21}$ отличны от нуля и, по основному предположению, дают операторы, кратные E. Поэтому A^0_{21} и B^0_{12} можно считать взаимно обратными операторами; но это возможно, лишь если пространства X_1 и X_2 имеют одинаковую размерность. Мы пришли к выводу: *всякая категория, не являющаяся подкатегорией категории типа $\mathfrak{K}_2$, есть категория типа $\mathfrak{K}_1$.*

12.33. Категориями типа $\mathfrak{K}_1$ и $\mathfrak{K}_2$ не исчерпываются возможные категории с двумя пространствами и $\mathfrak{B}_1 = (\lambda E)$, $\mathfrak{B}_2 = (\lambda E)$. Действительно, если из совокупности $\mathfrak{B}_{21}$ категории $\mathfrak{K}$ типа $\mathfrak{K}_2$ мы выберем линейную подсовокупность, не увеличивая N_1 и не уменьшая N_2 (например, наложив на элементы матриц операторов A_{21} подходящее дополнительное линейное однородное условие), то получим категорию $\mathfrak{K}'$, удовлетворяющую поставленным условиям, но не совпадающую с $\mathfrak{K}$. В частично упорядоченном по включению множестве всех категорий с $\mathfrak{B}_j = (\lambda E)$ $(j = 1, 2)$ категории типа $\mathfrak{K}_2$ характеризуются тем, что они максимальны; иными словами, каждая категория типа $\mathfrak{K}_2$, за исключением вырожденных случаев, когда $X'_2 = (0)$ или $X'_1 = (0)$, не может быть расширена с сохранением свойств категории и условий $\mathfrak{B}_j = (\lambda E)$. Действительно, допустим, что категорию $\mathfrak{K}$ типа $\mathfrak{K}_2$ можно расширить добавлением оператора A^0_{21}, принимающего на некотором $x_1 \in X_1 - X'_1$ значение $y_2 \notin X'_2$. Допустим, что $X'_1 \neq 0$. Возьмем оператор $B^0_{12} \in \mathfrak{B}_{12}$, переводящий y_2 в ненулевой вектор $x_1' \in X_1'$. Тогда $B_{12}^0 A_{21}^0 \; x_1 = x_1'$, что противоречит условию.

Далее, пусть к категории $\mathfrak{K}$ добавлен оператор A_{21}, переводящий вектор $x_1 \in X_1'$ не в нуль, а, например, в вектор $y_2' \in X_2'$. Ясно, что $X_2' \neq X_2$, иначе было бы $X_1' = \mathfrak{B}_{12} \; X_2 = \mathfrak{B}_{12} \; X_2 = (0)$ и указанного вектора x_1, переводящегося не в нуль, не могло бы существовать. Поэтому в $\mathfrak{B}_{12}$ имеется оператор B_{12}, переводящий вектор $y_2 \in X_2 - X_2'$ в x_1. Тогда $A_{21} B_{12} y_2 = y_2'$, что противоречит условию.

Аналогично, предполагая $X_2' = 0$, получаем, что нельзя добавить ни одного оператора к совокупности $\mathfrak{B}_{12}$. Итак, рассматриваемая категория $\mathfrak{K}$ типа $\mathfrak{K}_2$ максимальна в предположении, что $X_1' \neq 0$ и $X_2 \neq 0$.

12.34. Случаи вырождения нужно рассмотреть отдельно. Пусть, например, $X_1' = 0$, так что $\mathfrak{B}_{12}$ состоит только из нулевого оператора. Если теперь $X_2' \neq X_2$, то категория не максимальна и в совокупность $\mathfrak{B}_{21}$ можно добавить без нарушения условий $\mathfrak{B}_j = (\lambda E)$ все операторы, действующие из X_1 в X_2. Таким образом, получается «тривиальная» максимальная категория, где $\mathfrak{B}_{12} = (0)$, а $\mathfrak{B}_{21}$ — полная система операторов, действующих из $\mathfrak{B}_1$ в $\mathfrak{B}_2$. Аналогичная максимальная категория получается с $\mathfrak{B}_{21} = (0)$ и $\mathfrak{B}_{12}$ — полной системой. В результате мы получаем, что общая категория типа $\mathfrak{K}_2$ максимальна при условиях: 1) $X_1' \neq 0$, $X_2' \neq 0$, 2) $X_1' = 0$, $X_2' = X_2$; 3) $X_1' = X_1$, $X_2' = 0$.

12.35. Мы переходим теперь к общему случаю категории из любого числа $N \leqslant \infty$ пространств $\{X_\alpha\}$, $\alpha \in \mathfrak{A}$. Здесь имеет место утверждение, аналогичное доказанной нами альтернативе (*12.32*).

Альтернатива. *Если* $\mathfrak{B}_1 = \ldots = \mathfrak{B}_k = (\lambda E)$, *то или произведение* $\mathfrak{B}_{1k} \mathfrak{B}_{k,\,k-1} \ldots \mathfrak{B}_{32} \mathfrak{B}_{21} = 0$, *или пространства* $X_1, \ldots, X_k$ *имеют одинаковую размерность и* $\mathfrak{B}_{ij} = \{\lambda A_{ij}\}$, *где* A_{ij} — *фиксированный обратимый оператор, причем*

$$A_{1k} A_{k,\,k-1} \ldots A_{32} A_{21} = E.$$

Доказательство. Допустим, что произведение $\mathfrak{B}_{1k} \, \mathfrak{B}_{k,\,k-1} \ldots \mathfrak{B}_{21}$ содержит ненулевой оператор, равный, следовательно, λE с $\lambda \neq 0$. Пусть размерность пространства X_j есть r_j; тогда матрицы $A_{1k} \in \mathfrak{B}_{1k}, \ldots, A_{21} \in \mathfrak{B}_{21}$, дающие в произведении λE, имеют размеры

$$r_1 \times r_k,\ r_k \times r_{k-1},\ \ldots,\ r_3 \times r_2,\ r_2 \times r_1.$$

Рассмотрим произведение матриц $A_{1k}A_{k,\,k-1}\ldots A_{32}$. Эта матрица, имеющая размеры $r_1 \times r_2$, действующая, следовательно, из X_2 в X_1. Рассмотрим категорию $\mathfrak{K}_0$ из двух пространств X_1, X_2, где совокупности операторов $\mathfrak{B}^0_{12}$ и $\mathfrak{B}^0_{21}$ следующие:

$$\mathfrak{B}^0_{21} = \mathfrak{B}_{21},\quad \mathfrak{B}^0_{12} = \mathfrak{B}_{1k}\mathfrak{B}_{k,\,k-1}\ldots\mathfrak{B}_{32}$$

($\mathfrak{B}^0_{12}$ — линейная оболочка соответствующих произведений). Так как, по условию,

$$A_{1k}A_{k,k-1}\ldots A_{21} = \lambda E,$$

то, по доказанному, X_1 и X_2 имеют одинаковую размерность $r_1 = r_2$, $\mathfrak{B}^0_{21} = \mathfrak{B}_{21} = (\lambda A^0_{21})$, где A^0_{21} — обратимая матрица $\mathfrak{B}^0_{12} = (\mu A^{0-1}_{21})$.

Далее, мы имеем

$$A^{0-1}_{21}A_{1k}A_{k,\,k-1}\ldots A_{32} = \lambda E,$$

а рассуждая, как и раньше, получаем, что $r_2 = r_3$, $\mathfrak{B}_{32} = \lambda A^0_{32}$, где A^0_{32} — обратимая матрица. После k шагов приходим к требуемому утверждению.

12.36. В дальнейшем, рассматривая категорию из N пространств, мы можем предполагать, что все циклические произведения $\mathfrak{B}_{\alpha\beta}\mathfrak{B}_{\beta\gamma}\ldots\mathfrak{B}_{\delta\alpha}$ равны нулю; в противном случае мы просто отождествили бы соответствующие пространства одинаковой размерности.

Рассмотрим вначале следующую конкретную категорию, которую будем обозначать $\mathfrak{K}^N_2$. Выберем в X_1 произвольно $N-1$ подпространств, которые обозначим $X_{12}, \ldots, X_{1N}$. При попарно различных $j, k, l, \ldots$ положим

$$X_{1jk} = X_{1j} \cap X_{1k},\ X_{1jkl} = X_{1j} \cap X_{1k} \cap X_{1l}$$

и т. д., образуя последовательно пересечения пространств X_{1j} по два, по три и т. д. Если N конечно, последним пересечением будет $X_{12\ldots N}$ — пересечение всех $N-1$ выделенных подпространств, если N бесконечно, последнего пересечения не будет. Такое же построение проведем во всех остальных пространствах $X_2, \ldots$, причем индекс всего пространства будет первым индексом в обозначении всех его подпространств.

Каждой системе попарно различных индексов $j, \ldots, k$ теперь отвечает однозначно некоторое из выделенных подпространств. Построим совокупность $\mathfrak{B}_{21}$. По определению она состоит из всех операторов, переводящих X_1 в X_2 так, что каждое из подпространств $X_{1j\ldots k}$ переходит в $X_{21j\ldots k}$, если в последовательности $j, \ldots, k$ нет индекса 2, или в (0), если в ней индекс 2 имеется. Аналогично построим все совокупности $\mathfrak{B}_{\beta\alpha}$.

Докажем, что мы получили категорию.

Пусть имеются некоторые операторы $A_{21} \in \mathfrak{B}_{21}$ и $B_{32} \in \mathfrak{B}_{32}$. Рассмотрим оператор $B_{32}A_{21}$. Он переводит пространство X_1 в X_3. Подпространство $X_{1j\ldots k}$ оператор A_{21} переводит в $X_{21j\ldots k}$, а полученное подпространство оператор B_{32} переводит в $X_{321j\ldots k} \subset X_{31j\ldots k}$. Таким образом, оператор $B_{32}A_{21}$ входит в $\mathfrak{B}_{31}$, что и требуется.

Если имеется цепь операторов $A_{1j}\ldots A_{k1}$, переводящая пространство X_1 в себя, то результирующий оператор переводит X_1 в $X_{1j\ldots k1} = 0$, что и согласуется с требованием $\mathfrak{B}_1 = (\lambda E)$.

12.37. Покажем теперь, что *любая категория из* $N \leqslant \infty$ *пространств* $X_1, \ldots$ *с* $\mathfrak{B}_j = (\lambda E)$ *содержится в некоторой категории типа* $\mathfrak{K}_2^N$. Обозначим через X_{jk} полный образ пространства X_k в пространстве X_j под воздействием всех операторов $\mathfrak{B}_{jk}$ и далее через $X_{jkl\ldots sm}$ полный образ в пространстве X_j пространства X_m под воздействием всех операторов вида $A_{jk}A_{kl}\ldots A_{sm}$ (порядок индексов существен!). Покажем, что $X_{jkl\ldots sm}$ содержится в пересечении $X_{jk}, X_{jl}, \ldots$ $\ldots, X_{jm}$. Действительно, если $z \in X_{jkl\ldots sm}$, то

$$z \in \sum_\alpha A^\alpha_{jk} A^\alpha_{kl} \ldots A^\alpha_{pq} \ldots A^\alpha_{sm} z^\alpha_m,$$

где $z^\alpha_m \in X_m$, или

$$z = \sum_\alpha A^\alpha_{jk} \ldots A^\alpha_{pq} y^\alpha_q,$$

где $y^\alpha_q = \sum\limits_\alpha \ldots A^\alpha_{sm} z^\alpha_m \in X_q$, а так как $A^\alpha_{jk} \ldots A^\alpha_{pq} \in \mathfrak{B}_{jq}$, то $z \in X_{jq}$, что и требуется.

Заметим, далее, что A_{ij} переводит $X_{j\ldots m}$ в $X_{ij\ldots m}$. Теперь ясно, что наша категория содержится в категории типа

$\mathfrak{K}_2^N$ с определяющими подпространствами X_{jk}; в частности, все максимальные категории должны иметь тип $\mathfrak{K}_2^N$.

Неясно, однако, при каких условиях на подпространства X_{jk} категория типа $\mathfrak{K}_2^N$ является максимальной. (Напомним, что в случае двух пространств X_1 и X_2 необходимые и достаточные условия максимальности категории $\mathfrak{K}_2$ состояли в том, что пространства X_{12} и X_{21} либо оба отличны от нуля, либо одно есть нуль, а другое — все пространство.)

§ 12.4. Все данные алгебры $\mathfrak{B}_\alpha$ — простые

В этом случае каждая алгебра $\mathfrak{B}_\alpha$ в некотором базисе записывается матрицами вида

$$\left\|\begin{array}{cccc} C & & & \\ & C & & \\ & & \ddots & \\ & & & C \end{array}\right\|, \tag{2}$$

где C — одна и та же матрица фиксированного размера r_α, причем во всем $\mathfrak{B}_\alpha$ она пробегает всю совокупность квадратных матриц с r_α строками и столбцами (*11.86*).

12.41. Рассмотрим вначале категорию $\mathfrak{K}$ из двух пространств X_1 и X_2 размерностей n_1 и n_2 с размерами квадратов m_1 и m_2 и числом их k_1 и k_2 (так что $n_1 = k_1 m_1$, $n_2 = k_2 m_2$). Матрица $\mathcal{A}_{21}$ категории $\mathfrak{K}$ может быть разбита на блоки следующим образом:

$$\mathcal{A}_{21} = \begin{array}{r} m_2\{ \\ \\ \\ \\ \end{array}\left\|\begin{array}{c|c|c|c} \overbrace{A_{11}}^{m_1} & A_{12} & \cdots & A_{1k_1} \\ \hline A_{12} & \cdots & \cdots & \cdots \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline A_{k_2 1} & \cdots & \cdots & A_{k_2 k_1} \end{array}\right\|\Bigg\} n_2.$$

Аналогично можно представить матрицу $\mathscr{B}_{12}$:

$$\mathscr{B}_{12}=\left.\left\|\begin{array}{cccc} \overbrace{B_{11} \quad B_{12}}^{} & & \cdots & B_{1k_2} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ B_{k_1 1} & & \cdots & B_{k_1 k_2} \end{array}\right\|\right\}n_1.$$

(m_1 — высота первой строки блоков.)

Мы утверждаем: *или произведение* $\mathscr{A}_{21}\mathscr{B}_{12}=0$ (*при любых* $\mathscr{A}_{21}\in\mathfrak{B}_{21}$ *и* $\mathscr{B}_{12}\in\mathfrak{B}_{12}$), *или* $k_1=k_2$ *и все матрицы* A_{jk} *являются кратными одной (произвольной) фиксированной матрицы* A, *равно как и все матрицы* B_{jk} *являются кратными одной (произвольной) фиксированной матрицы* B; *при этом коэффициенты кратностей образуют взаимно обратные матрицы порядка* $k_1=k_2$. Категорию второго типа будем обозначать в дальнейшем через $\mathfrak{K}_3$.

Для доказательства заметим, что в категорию $\mathfrak{K}$ вместе с матрицами $\mathscr{A}_{21}$ и $\mathscr{B}_{12}$ входят их произведения (с соответствующей стороны) на матрицы C_1 и C_2 вида (2), поэтому наряду с равенством

$$\mathscr{A}_{21}\mathscr{B}_{12}=C_2$$

имеет место и равенство

$$\mathscr{A}_{21}C_1\mathscr{B}_{12}=C_2$$

с произвольной матрицей C_1 вида (2). Вспоминая правило умножения блочных матриц (*4.51*), мы можем написать уравнения (в блоках):

$$\left.\begin{array}{l} A_{11}CB_{11}+A_{12}CB_{21}+\ldots+A_{1k_1}CB_{k_1 1}= \\ =A_{21}CB_{12}+A_{22}CB_{22}+\ldots+A_{2k_1}CB_{k_1 2}= \\ =A_{k_2 1}CB_{1k_2}+A_{k_2 2}CB_{2k_2}+\ldots+A_{k_2 k_1}CB_{k_1 k_2}, \\ A_{11}CB_{12}+A_{12}CB_{22}+\ldots+A_{1k_1}CB_{k_1 2}=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{array}\right\} \qquad (3)$$

Возьмем в качестве матрицы C матрицу, имеющую единственный ненулевой элемент 1 на скрещении r-й строки и s-го столбца ($r\leqslant m_1$, $s\leqslant m_1$). Вообще, если A — любая $m_2\times m_1$-матрица, а B — любая $m_1\times m_2$-матрица, то произведение ACB есть $m_2\times m_2$-матрица первого ранга, у которой на скрещении p-й строки и q-го столбца стоит элемент $a_{pr}b_{sq}$. Поэтому равенства (3) можно переписать в форме

$$\left.\begin{array}{l} a_{pr}^{11}b_{sq}^{11}+a_{pr}^{12}b_{sq}^{21}+\ldots+a_{pr}^{1k_1}b_{sq}^{k_1 1}= \\ =a_{pr}^{21}b_{sq}^{12}+a_{pr}^{22}b_{sq}^{22}+\ldots+a_{pr}^{2k_1}b_{sq}^{k_1 2}= \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ =a_{pr}^{k_2 1}b_{sq}^{1k_2}+a_{pr}^{k_2 2}b_{sq}^{2k_2}+\ldots+a_{pr}^{k_2 k_1}b_{sq}^{k_1 k_2}, \\ a_{pr}^{11}b_{sq}^{12}+a_{pr}^{12}b_{sq}^{22}+\ldots+a_{pr}^{1k_1}b_{sq}^{k_1 2}=0, \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \end{array}\right\} \tag{4}$$

где индексы наверху означают номера соответствующих матриц.

Уравнения (4) можно трактовать как матричное равенство

$$\tilde{A}_{pr}\tilde{B}_{sq}=\begin{Vmatrix} a_{pr}^{11} & a_{pr}^{12} & \ldots & a_{pr}^{1k_1} \\ a_{pr}^{21} & a_{pr}^{22} & \ldots & a_{pr}^{2k_1} \\ \cdot & \cdot & \cdot & \cdot \\ a_{pr}^{k_2 1} & a_{pr}^{k_2 2} & \ldots & a_{pr}^{k_2 k_1} \end{Vmatrix}\begin{Vmatrix} b_{sq}^{11} & b_{sq}^{12} & \ldots & b_{sq}^{1k_2} \\ b_{sq}^{21} & b_{sq}^{22} & \ldots & b_{sq}^{2k_2} \\ \cdot & \cdot & \cdot & \cdot \\ b_{sq}^{k_1 1} & b_{sq}^{k_1 2} & \ldots & b_{sq}^{k_1 k_2} \end{Vmatrix}=$$

$$=\underbrace{\begin{Vmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{Vmatrix}}_{k_2}\Bigg\} k_2.$$

Аналогично имеет место и равенство

$$\tilde{B}_{sq}\tilde{A}_{pr}=\begin{Vmatrix} b_{sq}^{11} & b_{sq}^{12} & \ldots & b_{sq}^{1k_2} \\ b_{sq}^{21} & b_{sq}^{22} & \ldots & b_{sq}^{2k_2} \\ \cdot & \cdot & \cdot & \cdot \\ b_{sq}^{k_1 1} & b_{sq}^{k_1 2} & \ldots & b_{sq}^{k_1 k_2} \end{Vmatrix}\begin{Vmatrix} a_{pr}^{11} & a_{pr}^{12} & \ldots & a_{pr}^{1k_1} \\ a_{pr}^{21} & a_{pr}^{22} & \ldots & a_{pr}^{2k_1} \\ \cdot & \cdot & \cdot & \cdot \\ a_{pr}^{k_1 1} & a_{pr}^{k_1 2} & \ldots & a_{pr}^{k_1 k_1} \end{Vmatrix}=$$

$$=\underbrace{\begin{Vmatrix} \mu & & & \\ & \mu & & \\ & & \ddots & \\ & & & \mu \end{Vmatrix}}_{k_1}\Bigg\} k_1.$$

Мы видим, что матрицы $\tilde{A}_{pr}$ и $\tilde{B}_{sq}$ (p, r, s, q — параметры этих матриц) образуют категорию, связывающую пространство E_1 размерности k_1 с пространством E_2 размерности k_2 при условиях $R_1 = (\lambda Y)$, $R_2 = (\mu Y)$. Мы можем теперь применить альтернативу, установленную в *12.32*. Именно, если $k_1 \neq k_2$, то на самом деле всегда $\lambda = 0$, $\mu = 0$; если же мы допускаем, что хотя бы при одной системе индексов p, q, r, s получается $\lambda \neq 0$ (или $\mu \neq 0$), то $k_1 = k_2$ и все матрицы $\tilde{A}_{pr}$ представляют собой кратные одной обратимой матрицы, а все матрицы $\tilde{B}_{sq}$ — кратные обратной матрицы:

$$\tilde{A}_{pr} = \lambda_{pr}\tilde{A}, \quad \tilde{B}_{sq} = \mu_{sq}\tilde{B}.$$

Матрица $\tilde{A}_{pr}$ состоит из элементов матриц A_{ij}, находящихся на скрещении p-й строки и r-го столбца. Мы видим, что выполняется условие

$$a_{pr}^{ij} = \lambda \tilde{a}^{ij},$$

где $\tilde{a}^{ij}$ — элементы матрицы $\tilde{A}$. Поэтому мы приходим к выводу, что *все матрицы A_{ij} являются кратными фиксированной матрицы $A = \| \lambda_{pr} \|$ с коэффициентами a^{ij}*. Аналогичный вид имеют матрицы B_{ij}, чем и завершается доказательство нашего утверждения.

12.42. Итак, если $\mathscr{A}_{12}\mathscr{B}_{21} \neq 0$, то $k_1 = k_2$ и категория $\mathfrak{K}$ имеет вид

$$\mathscr{A}_{21} = \left\| \begin{array}{c|c|c|c} \tilde{a}^{11}\Lambda & \tilde{a}^{12}\Lambda & \ldots & \tilde{a}^{1k_1}\Lambda \\ \hline \tilde{a}^{21}\Lambda & \ldots & \ldots & \ldots \\ \hline \ldots & \ldots & \ldots & \ldots \\ \hline \tilde{a}^{k_1 1}\Lambda & \ldots & \ldots & \tilde{a}^{k_1 k_1}\Lambda \end{array} \right\|,$$

$$\mathscr{B}_{12} = \left\| \begin{array}{c|c|c|c} \tilde{l}^{11}M & \tilde{b}^{12}M & \ldots & \tilde{b}^{1k_1}M \\ \hline \tilde{b}^{21}M & \ldots & \ldots & \ldots \\ \hline \ldots & \ldots & \ldots & \ldots \\ \hline \tilde{b}^{k_1 1}M & \ldots & \ldots & \tilde{b}^{k_1 k_1}M \end{array} \right\|.$$

Матрицы $\Lambda = \| \lambda_{pq} \|$ и M прямоугольные: $1 \leqslant p \leqslant m_2$, $1 \leqslant q \leqslant m_1$. Среди матриц Λ, участвующих в данной категории, имеется ненулевая Λ_0 (поскольку $\mathfrak{B}_{21}\mathfrak{B}_{12} \neq 0$), и поэтому среди них должна быть и любая $m_2 \times m_1$-матрица, поскольку из ненулевой матрицы умножением справа на C_2 и слева на C_1 можно получить любую такую матрицу. Итак, *если* $\mathfrak{B}_{21}\mathfrak{B}_{12} \neq 0$, *семейство* R_{21} *состоит из матриц вида*

$$\left\| \begin{array}{c|c|c} \tilde{a}^{11}\Lambda & \ldots & \tilde{a}^{1k_1}\Lambda \\ \hline \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ \hline \tilde{a}^{k_1 1}\Lambda & & \tilde{a}^{k_1 k_1}\Lambda \end{array} \right\|, \tag{5}$$

где $\tilde{A} = \| \tilde{a}^{ij} \|$ — *фиксированная обратимая матрица, а* Λ *пробегает всю совокупность* $m_2 \times m_1$-*матриц.*

Аналогичная картина будет иметь место, если $\mathfrak{B}_{12}\mathfrak{B}_{21} \neq 0$. Теперь ясно, что неравенства $\mathfrak{B}_{12}\mathfrak{B}_{21} \neq 0$ и $\mathfrak{B}_{21}\mathfrak{B}_{12} \neq 0$ имеют место или не имеют места одновременно.

12.43. Полученный результат может быть сформулирован в терминах тензорных произведений. Эта точка зрения позволяет, кроме того, выяснить и некоторые дополнительные обстоятельства. Приведем необходимые нам определения.

Пусть даны k-мерное пространство X с базисом $e_1, \ldots, e_k$ и m-мерное пространство Y с базисом $f_1, \ldots, f_m$. *Тензорным произведением* $X \times Y = Z$ пространств X и Y называется совокупность всех конечных формальных сумм вида

$$\sum_{\nu=1}^{p} x_\nu \times y_\nu,$$

где $x_\nu \in X$, $y_\nu \in Y$. При этом предполагается, что

$$[x_1 \times y] + [x_2 \times y] = [(x_1 + x_2) \times y],$$
$$[x \times y_1] + [x \times y_2] = [x \times (y_1 + y_2)].$$

$$\sum_{\nu=1}^{p} \lambda_\nu x_\nu \times y_\nu = \sum_{\nu=1}^{p} x_\nu \times \lambda_\nu y_\nu = \sum_{\nu=1}^{p} \lambda_\nu [x_\nu \times y_\nu].$$

Отсюда следует, что $X \times Y$ — линейное пространство, оно имеет размерность km, и его базисом могут служить векторы вида $e_i \times f_j$ $(i = 1, \ldots, k;\ j = 1, \ldots, m)$. Таким образом, векторы пространства Z имеют вид

$$g = \sum_{i=1}^{k} \sum_{j=1}^{m} c_{ij} [e_i \times f_j].$$

Мы можем преобразовать эту запись, выполнив суммирование по индексу i; получим $g = \sum_{j=1}^{m} \left[\sum_{i=1}^{k} c_{ij} e_i \times f_j \right] = \sum_{j=1}^{m} x_j \times f_j$, где $x_j = \sum_{i=1}^{k} c_{ij} e_i$ — произвольные (уже не обязательно базисные) векторы пространства X.

12.44. Пусть даны оператор A, действующий из пространства X_1 в пространство X_2, и оператор B, действующий из пространства Y_1 в пространство Y_2. Определим тензорное произведение $C = A \times B$ как оператор, действующий из пространства $Z_1 = X_1 \times Y_1$ в пространство $Z_2 = X_2 \times Y_2$ по формуле (дополнительный индекс — номер пространства)

$$C\,[e_i^1 \times f_j^1] = Ae_i^1 \times Bf_j^1. \tag{6}$$

Если $Ae_i^1 = \sum a_{i\lambda} e_\lambda^2$ и $Bf_j^1 = \sum b_{j\mu} f_\mu^2$, то формула (6) преобразуется к виду

$$C\,[e_i^1 \times f_j^1] = \sum_{\lambda=1}^{k_2} \sum_{\mu=1}^{m_2} a_{i\lambda} b_{j\mu}\,[e_\lambda^2 \times f_\mu^2].$$

Выясним структуру матрицы оператора C относительно базисов $e_i^1 \times f_j^1$ и $e_k^2 \times f_l^2$, упорядоченных в $X_1 \times Y_1$ по правилу

$$e_1^1 \times f_1^1;\ e_2^1 \times f_1^1,\ \ldots,\ e_{k_1}^1 \times f_1^1;\ e_1^1 \times f_2^1,\ e_2^1 \times f_2^1,\ \ldots,\ e_{k_1}^1 \times f_2^1,\ \ldots$$
$$\ldots,\ e_1^1 \times f_{m_1}^1,\ e_2^1 \times f_{m_1}^1,\ \ldots,\ e_{k_1}^1 \times f_{m_1}^1,$$

и аналогично в пространстве $X_2 \times Y_2$. Матрица C согласно определениям *12.13* имеет вид

$$\left|\begin{matrix}
a_{11}b_{11} & a_{12}b_{11} & \ldots & a_{1k_2}b_{11} & \ldots & a_{11}b_{1m_2} & a_{12}b_{1m_2} & \ldots & a_{1k_2}b_{1m_2} \\
a_{21}b_{11} & a_{22}b_{11} & \ldots & a_{2k_2}b_{11} & \ldots & a_{21}b_{1m_2} & a_{22}b_{1m_2} & \ldots & a_{2k_2}b_{1m_2} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
a_{k_11}b_{11} & a_{k_12}b_{11} & \ldots & a_{k_1k_2}b_{11} & \ldots & a_{k_11}b_{1m_2} & a_{k_12}b_{1m_2} & \ldots & a_{k_1k_2}b_{1m_2} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
a_{11}b_{m_11} & a_{12}b_{m_11} & \ldots & a_{1k_2}b_{m_11} & \ldots & a_{11}b_{m_1m_2} & a_{12}b_{m_1m_2} & \ldots & a_{k_2k_2}b_{m_1m_2} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
a_{k_11}b_{m_11} & a_{k_12}b_{m_11} & \ldots & a_{k_1k_2}b_{m_11} & \ldots & a_{k_11}b_{m_1m_2} & a_{k_12}b_{m_1m_2} & \ldots & a_{k_1k_2}b_{m_1m_2}
\end{matrix}\right|$$

или (используем блок-запись)

$$C = \left\|\begin{matrix} Ab_{11} & \ldots & Ab_{1m_2} \\ \cdot & \cdot & \cdot \\ Ab_{m_11} & \ldots & Ab_{m_1m_2} \end{matrix}\right\|.$$

12.45. Применяя *12.44* и *12.41*, мы видим: операторы алгебры $\mathfrak{B}_1$, рассмотренной выше, суть операторы в тензорном произведении m_1-мерного пространства X_1 и k_1-мерного пространства Y_1, которые являются тензорными произведениями любого оператора $C \in \mathfrak{B}(X_1)$ на единичный оператор $E \in \mathfrak{B}(Y_1)$.

Операторы системы $\mathfrak{B}_{21}$, рассмотренной выше, являются тензорными произведениями любых операторов $A_{21} \in \mathfrak{B}(X_2, X_1)$ на фиксированный обратимый оператор $\tilde{A}_{21} \in \mathfrak{B}(Y_2, Y_1)$ и аналогично операторы системы $\mathfrak{B}_{12}$ являются тензорными произведениями любых операторов $M_{12} \in \mathfrak{B}(X_1, X_2)$ на $\tilde{A}^{-1}$.

12.46. Для произведения тензорных произведений операторов справедлива очевидная формула

$$(A \times B)(C \times D) = (AC) \times (BD).$$

Поэтому, перемножая операторы категории A_{21} и B_{12}, мы находим, в частности,

$$(\Lambda \times \tilde{A})(M \times \tilde{A}^{-1}) = (\Lambda M) \times (\tilde{A}\tilde{A}^{-1}) = (\Lambda M) \times E \in \mathfrak{B}_1,$$

откуда следует выполнение свойства категории.

12.47. Найдем инвариантные пространства алгебры операторов $\mathfrak{B} = \{C \times E\}$, действующих в пространстве $Z = X \times Y$. Таким инвариантным подпространством является тензорное произведение $X \times Y_0$, где $Y_0 \in Y$ есть любое подпространство, поскольку

$$(C \times E)(X \times Y_0) = CX \times EY_0 \, (X \times Y_0).$$

Покажем, что никаких иных инвариантных подпространств в пространстве Z нет. Действительно, пусть $z = \sum x_j \times y_i$ — любой вектор пространства Z. Можно считать, что x_j линейно независимы. Рассмотрим оператор C, переводящий векторы x_j в заданные векторы $\bar{x}_j \in X$. Тогда

$$(C \times E) \sum x_j \times y_i = \sum \bar{x}_j \times y_j.$$

Таким образом, во всякое подпространство, инвариантное относительно всех операторов $C \times E$, вместе с каждым вектором $\sum x_j \times y_j$ попадает любой вектор $\sum \bar{x}_j \times y_j$, откуда и вытекает требуемое.

Если мы применим к инвариантному подпространству $X_1 \times Y_{10}$ в $Z_1 = X_1 \times Y_1$ оператор категории $\Lambda \times \bar{A}$, то в силу произвольности матрицы Λ получим в пространстве

$Z_2 = X_2 \times Y_2$ в качестве образа подпространство $\Lambda X_1 \times AY_{10} = X_2 \times Y_{20}$. Таким образом, *операторы категории устанавливают взаимно однозначное соответствие между инвариантными подпространствами в пространствах* Z_1 *и* Z_2 *и одновременно между обычными подпространствами пространств* Y_1 *и* Y_2.

12.48. Все изложенное справедливо при условии $\mathfrak{B}_{21}\mathfrak{B}_{12} \neq 0$ (или, что то же самое, при $\mathfrak{B}_{12}\mathfrak{B}_{21} \neq 0$). Если $\mathfrak{B}_{12}\mathfrak{B}_{21} = \mathfrak{B}_{21}\mathfrak{B}_{12} = 0$, то изложенная схема уже не работает и матрица категории вообще не состоит из блоков, кратных фиксированной матрице Λ. Мы находимся в условиях *12.32* и можем применить соответствующий результат: *категория есть часть некоторой категории типа* K_2; какая именно, уточним дальше.

12.49. Переходим к случаю категории из любого числа пространств X_α с алгебрами $\mathfrak{B}_\alpha$ рассматриваемого типа. Назовем пространства Z_1 и Z_2 *родственными*, если $\mathfrak{B}_{12}\mathfrak{B}_{21} \neq 0$ и, следовательно, матрицы $\mathfrak{B}_{12}$ имеют вид (5). Ясно, что соотношение родства транзитивно: если Z_1 родственно с Z_2, а Z_2 родственно с Z_3, то Z_1 родственно с Z_3, так как в силу произвольности матриц Λ в произведении $\mathfrak{B}_{32}\mathfrak{B}_{21}$ имеются ненулевые матрицы.

Поэтому всю совокупность пространств X_α можно разбить на непересекающиеся классы взаимно родственных пространств. Если X_1 и X_2 входят в разные классы, то $\mathfrak{B}_{12}\mathfrak{B}_{21} = \mathfrak{B}_{21}\mathfrak{B}_{12} = 0$.

Теперь мы можем с некоторыми изменениями повторить схему *12.36*. Допустим, что наши пространства разбиты на некоторое число классов $G_1, \ldots, G_r, \ldots$ взаимно родственных пространств. Пространства, входящие в класс G_r, имеют вид $X_{rj} \times Y_r$, где Y_r означает фактически одно пространство, в котором действуют обратимые операторы. Рассмотрим лишь пространства Y_r и построим из них категорию $\mathfrak{K}_2^N$, как было указано в *12.36*, с тем, чтобы выполнялось условие $\mathfrak{B}_{is}\mathfrak{B}_{si} = 0$ (произвольно выбираем подпространства $Y_{r\nu}$ и строим пересечения $Y_{r\nu\mu}, Y_{r\nu\mu\tau}, \ldots$). Эта категория состоит из операторов A_{is}, отображающих Y_s в Y_i и при этом переводящих подпространства $Y_{s\nu\mu}, \ldots \subset Y_s$ в подпространства

$Y_{is\nu\mu}, \ldots \subset Y_i$. Категория, которую мы обозначим через $\mathfrak{K}_4^N$, для пространства Z_j строится так: *если Z_j и Z_k родственны, то операторы $A_{jk} \in \mathfrak{B}_{jk}$ описаны выше; если же $Z = X_j \times Y_j$ и $Z_k = X_k \times Y_k$ принадлежат к разным классам, то операторы A_{jk} — любые операторы, переводящие Z_k в Z_j и при этом переводящие инвариантное подпространство $X_k \times Y_{k\nu\mu} \ldots$ в инвариантное подпространство $X_j \times Y_{jk\nu\mu} \ldots$*

Проверим, что любая категория $\mathfrak{K}$ с простыми кольцами $\mathfrak{B}_j$ содержится в категории типа $\mathfrak{K}_4^N$. Пусть $Z_k = X_k \times Y_k$ и $Z_j = X_j \times Y_j$ принадлежат к различным классам родственных пространств. Пусть Z_{jk} — полный образ пространства Z_k в пространстве Z_j под воздействием всех операторов семейства $\mathfrak{B}_{jk}$. Очевидно, что Z_{jk} есть инвариантное подпространство в Z_j и, следовательно, имеет вид $X_j \times Y_{jk}$, где Y_{jk} — некоторое подпространство в Y_j. Обозначим через $Z_{jkl\ldots sm}$ полный образ в Z_j пространства Z_m под воздействием всех операторов вида $A_{jk}A_{kl} \ldots A_{sm}$. Это — также инвариантное подпространство, и легко установить, действуя, как и в *12.36*, что оно содержится в пересечении $Z_{jk}, \ldots, Z_{jm}$. Отсюда видно, что наша категория является частью категории вида $\mathfrak{K}_4^N$, что мы и утверждали.

§ 12.5. Все данные алгебры $\mathfrak{B}_\alpha$ — полные алгебры диагональных матриц

В этом случае в каждом из рассматриваемых пространств X_α выделяется фиксированный базис, относительно которого все матрицы операторов $A_\alpha \in \mathfrak{B}_\alpha$ диагональны. С помощью таких базисов операторы $A_{\beta\alpha} \in \mathfrak{B}_{\beta\alpha}$ также записываются определенными матрицами (прямоугольными), так что наша задача может быть поставлена как задача из теории матриц.

12.51. Рассмотрим вначале категорию $\mathfrak{K}$ из двух пространств X_1 и X_2. Пусть $A_{12} \in \mathfrak{B}_{12}$ — любой оператор; согласно определению категории произведение

$$B_{12} = A_1 A_{12} B_2 \tag{7}$$

также принадлежит $\mathfrak{B}_{12}$, если A_1 и B_2 суть соответствующие диагональные матрицы.

Возьмем в качестве A_1 матрицу с единственным ненулевым элементом 1 на скрещении j-й строки и j-го столбца, а в качестве B_2 матрицу с единственным ненулевым элементом 1 на скрещении k-й строки и k-го столбца. Тогда согласно *12.14* в матрице B_{12} будет (возможно) отличен от нуля только один элемент, стоящий на скрещении j-й строки и k-го столбца, и этим элементом будет элемент a_{jk} матрицы A_{12}. Таким образом, операция (7) заменяет в матрице A_{12} все элементы нулями, кроме элемента a_{jk}, который она оставляет неизменным.

Отсюда можно сделать вывод о структуре системы $\mathfrak{B}_{12}$: *система $\mathfrak{B}_{12}$ представляет собой совокупность всех матриц, элементы которых на фиксированном множестве мест произвольны, в то время как остальные равны нулю.*

12.52. Будем обозначать в матрицах системы $\mathfrak{B}_{12}$ фиксированное множество мест, на котором разрешаются любые элементы, через S_{12}.

Мы должны выяснить теперь, как связаны множества S_{12} и S_{21}. Возьмем матрицу $A_{12} \in \mathfrak{B}_{12}$ с единственным отличным от нуля элементом, равным 1, на скрещении j_1-й строки и k_1-го столбца, $(j_1, k_1) \in S_{12}$, и любую матрицу $B_{21} \in \mathfrak{B}_{21}$ со всеми возможными отличными от нуля элементами на местах S_{21}. Произведения $C_1 = A_{12}B_{21}$ и $D_2 = B_{21}A_{12}$ согласно условию представляют собой диагональные матрицы. С другой стороны, согласно *12.13* мы имеем: в C_1 все строки равны 0, кроме j_1-й, а в j_1-й строке стоят элементы k_1-й строки матрицы B_{21}. Так как должна получиться диагональная матрица, то мы делаем вывод, что в k_1-й строке матрицы B_{21} все элементы, кроме стоящего в j_1-м столбце, равны нулю. В D_2 все столбцы равны нулю, кроме k_1-го, а в k_1-м столбце стоят элементы j_1-го столбца матрицы A_{12}; опять-таки, поскольку должна получиться диагональная матрица, то в j_1-м столбце матрицы A_{12} все элементы равны 0, кроме стоящего в k-й строке. Итак, *если* $(j_1, k_1) \in S_{12}$, *то в матрицах класса $\mathfrak{B}_{12}$ все элементы j_1-го столбца и k_1-й строки, кроме их пересечения, равны* 0. Этого достаточно, чтобы мы, зная множество S_{12}, могли сделать вывод о строении множества S_{21}. Переставляя строки и столбцы матриц $\mathfrak{B}_{12}$ (что равносильно перестановке элементов в базисах пространств X_1 и X_2), мы можем добиться, чтобы вначале шли строки и столбцы, в которых нет представителей множества S_{12}; далее — строки и столбцы, в которых только по одному представителю этого множества, и, наконец,— строки и столбцы, в которых не менее чем по два его представителя:

$$A_{12}\sim\begin{array}{c@{\;}c}
 & \begin{array}{ccc} 1\qquad\quad\gamma & \qquad\qquad\delta & \qquad\qquad n \end{array}\\
\begin{array}{c} 1\\ \\ \\ \alpha\\ \\ \\ \beta\\ \\ m \end{array} &
\left\|\begin{array}{c|c|c}
0\ \dots & 0\ \dots & 0\ \dots\\
0\ \dots & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
\hline
0\ \dots & 1\ \dots & 0\ \dots\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
0\ \dots & \dots\ 1 & 0\ \dots\\
\hline
0\ \dots & \cdot\ \cdot\ \cdot & 1\ 1\dots\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & 1\ \dots\\
0\ \dots & 0\ \dots & 1\ \dots
\end{array}\right\|
\end{array}. \qquad (8)$$

На этой схеме места, занятые множеством S_{12}, отмечены единицами, остальные — нулями.

Построим теперь матрицу $B_{21}\in\mathfrak{B}_{21}$; в ней n строк и m столбцов:

$$B_{21}\sim\begin{array}{c@{\;}c}
 & \begin{array}{ccc} 1\qquad\quad\alpha & \qquad\qquad\beta & \qquad\qquad m \end{array}\\
\begin{array}{c} 1\\ \\ \\ \gamma\\ \\ \\ \delta\\ \\ n \end{array} &
\left\|\begin{array}{c|c|c}
\cdot\ \cdot\ \cdot & 0\ \dots & 0\ \dots\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
\hline
0\ \dots & 1\ \dots & 0\ \dots\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
0\ \dots & \dots\ 1 & 0\ \dots\\
\hline
0\ \dots & \cdot\ \cdot\ \cdot & 0\ \dots\\
\cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\\
0\ \dots & 0\ \dots & 0\ \dots
\end{array}\right\|
\end{array}. \qquad (9)$$

Поскольку в матрице A_{12} на скрещении строки с номером $\alpha+1$ и столбца с номером $\gamma+1$ стоит 1, в матрице B_{21} на скрещении столбца с номером $\alpha+1$ и строки с номером $\gamma+1$ может стоять 1, и во всяком случае остальные элементы этой строки и этого столбца равны 0. То же относится ко всем строкам с номерами от $\gamma+1$ до δ и столбцам с номерами от $\alpha+1$ до β. Если в столбце матрицы A_{12} с номером $\delta+1$ имеются две единицы, все элементы соответствующей строки матрицы B_{21} равны 0. То же относится ко всем тем столбцам, идущим за δ-м, в которых имеется по две единицы. Если же в столбце матрицы A_{12} оказывается только одна единица, то имеются две единицы в соответствующей строке с номером $>\beta$, что приводит к равенству нулю столбца матрицы B_{21} с тем же номером. В результате весь правый нижний угол матрицы B_{21} заполняется нулями. Действительно, если взять место (j, k) в этом углу и рассмотреть соответствующее место (k, j) в матрице A_{12}, то k-я строка или j-й столбец имеет в матрице A_{12} по крайней мере две единицы (иначе мы поместили бы эти строку и столбец раньше). Значит, k-й столбец или j-я строка в B_{21} состоит сплошь из нулей; следовательно, на месте (j, k) должен быть нуль. Левый нижний и правый верхний углы в матрице B_{21} также сплошь состоят из нулей; действительно, если бы в матрице B_{21} оказалась 1 где-нибудь в левом нижнем углу, например на месте (j, k), то, по симметрии построения, в матрице A_{12} все элементы j-го столбца, кроме того, который стоит в k-й строке (т. е. в правом верхнем углу матрицы A_{12}!), должны были бы быть равными нулю. В этом столбце обязательно есть единицы в правом нижнем углу. Что же касается элементов левого верхнего угла матрицы B_{21}, то они могут быть произвольными. Ясно, что нашу категорию можно расширить, присоединив к множеству S_{12} *все* элементы нижнего правого угла матрицы (8) (если S_{12} еще содержит не все эти элементы) и к множеству S_{21} *все* элементы левого верхнего угла матрицы (9). После этого категория K станет максимальной, так как уже нельзя будет расширить S_{12} без уменьшения S_{21}. На языке геометрии, максимальная категория из двух пространств устроена так: *пространство* X_1 *разбито в прямую сумму трех подпространств, положим* X_1^0, X_1^1, X_1^2, *и пространство* X_2 *разбито в прямую сумму трех подпространств* X_2^0, X_2^1, X_2^2, *причем* $\dim \mathrm{X}_1^1 = \dim \mathrm{X}_2^1$; *действие оператора* A_{12} *состоит в том, что* X_1^0 *переходит в* 0, X_1^1 *отображается диагональной матрицей в* X_2^1, X_1^2 *любым образом отображается в* X_2^2, *а действие оператора* B_{21} *состоит в том, что* X_2^0 *любым образом отображается в* X_1^0, X_2^1 — *диагональной матрицей в* X_1^1 *и* X_2^2 — *в нуль*. Произвольная (не максимальная) категория отличается от максимальной тем, что операторы, переводящие X_1^2 в X_2^2, не произвольны, а соответствуют матрицам с некоторыми закрепленными местами для нулей. То же можно сказать об операторах, переводящих X_2^0 в X_1^0; при этом между местами этих нулей может не быть уже никакой связи.

12.53. Теперь займемся описанием категории $\mathfrak{K}$ с любым числом пространств X_α.

Прежде всего ясно, что каждая подкатегория категории $\mathfrak{K}$, образованная парой пространств X_α, X_β и соответствующими системами $\mathfrak{B}_{\beta\alpha}$ и $\mathfrak{B}_{\alpha\beta}$, устроена так, как было описано выше: в матрицах этих систем выделены множества $S_{\beta\alpha}$ и $S_{\alpha\beta}$, и $\mathfrak{B}_{\beta\alpha}$ состоит из всех матриц, у которых на множестве $S_{\beta\alpha}$ стоят любые числа, а вне этого множества — нули, и аналогично устроена система $\mathfrak{B}_{\beta\alpha}$. Мы будем обозначать через S любое подмножество мест в совокупности матриц фиксированного размера $m\times n$. Множество элементов главной диагонали квадратных матриц будем обозначать D.

Совокупность всех $m\times n$-матриц, имеющих на местах S произвольные элементы, а вне S — нули, обозначим через $\mathfrak{B}_{mn}(S)$. Пусть имеется множество S_1, определенное на $m\times n$-матрицах, и множество S_2, определенное на $n\times p$-матрицах.

Определим на $m\times p$-матрицах множество S (будем называть его произведением S_1S_2), которое состоит из всех тех мест в $m\times p$-матрицах, на которых могут получаться ненулевые элементы в произведении $\mathfrak{B}_{mn}(S_1)\mathfrak{B}_{np}(S_2)$. Иначе говоря, место (j, k) принадлежит совокупности S_1S_2 тогда и только тогда, когда существует номер j такой, что (i, j) принадлежит S_1, а (j, k) принадлежит S_2.

Пусть $S_{11}, \ldots, S_{1r}$ — совокупность множеств на $m\times n$-матрицах и $S_{21}, \ldots, S_{2q}$ — совокупность множеств на $n\times p$-матрицах. Из определения произведения легко вытекает общая формула

$$\bigcup_{j=1}^{r} S_{1j} \bigcup_{i=1}^{q} S_{2i} = \bigcup_{j=1}^{r} \bigcup_{i=1}^{q} S_{1j}S_{2i}. \tag{10}$$

С помощью операции произведения S-множеств условия категории этого параграфа можно записать так:

$$S_{\alpha\beta}S_{\beta\alpha}\subset D, \quad S_{\alpha\beta}S_{\beta\gamma}\subset S_{\alpha\gamma}. \tag{11}$$

12.54. Теперь мы построим семейство некоторых конкретных категорий. Задать категорию $\mathfrak{K}$ — значит задать все семейства $R_{\alpha\beta}$ или, что в данном случае то же, задать все множества $S_{\alpha\beta}$. Выберем произвольно S_{21} и затем подберем такое S_{12}, чтобы выполнялись условия $S_{21}S_{12} \subset D$, $S_{12}S_{21} \subset D$; как это сделать, описано выше. Допустим, что уже построены S_{jk} для всех j и k, меньших n, с выполнением свойств категории (11). Покажем, как построить S_{jn} и S_{nj} при $j < n$. Выберем S_{n1} произвольно, а S_{1n} с выполнением условия $S_{1n}S_{n1} \subset D$, $S_{n1}S_{1n} \subset D$. Пусть уже выбраны S_{jn} и S_{nj} для всех $j < k$ с выполнением условий (11) и нам остается выбрать S_{nk} и S_{kn}. Для искомых множеств S_{nk} и S_{kn} должны быть выполнены условия, вытекающие из (11):

а) $S_{nk}S_{kn} \subset D, \quad S_{kn}S_{nk} \subset D;$

б) $S_{jk}S_{kn} \subset S_{jn}, \quad S_{kn}S_{nj} \subset S_{kj}, \quad S_{in}S_{nk} \subset S_{ik}, \quad S_{nk}S_{ki} \subset S_{ni};$

в) $S_{kn} \supset S_{ki}S_{in}, \quad S_{nk} \supset S_{nj}S_{jk}.$

Условия а), б) ограничивают множества S_{nk} и S_{kn} сверху, условия в) — снизу. Покажем, что все эти условия совместны. Положим, например,

$$S_{kn} = \bigcup_{i=1}^{n-1} S_{ki}S_{in}, \qquad S_{nk} = \bigcup_{j=1}^{n-1} S_{nj}S_{jk}.$$

Тогда, по предположению индукции и в силу формулы (10),

$$S_{nk}S_{kn} = \bigcup_{j=1}^{n-1} S_{nj}S_{jk} \bigcup_{i=1}^{n-1} S_{ki}S_{in} = \bigcup_{j=1}^{n-1} \bigcup_{i=1}^{n-1} S_{nj}S_{jk}S_{ki}S_{in} \subset$$
$$\subset \bigcup_{j=1}^{n-1} \bigcup_{i=1}^{n-1} S_{nj}S_{ji}S_{in} \subset \bigcup_{i=1}^{n-1} S_{ni}S_{in} \subset D,$$

$$S_{jk}S_{kn} = S_{jk} \bigcup_{i=1}^{n-1} S_{ki}S_{in} = \bigcup_{i=1}^{n-1} S_{jk}S_{ki}S_{in} \subset \bigcup_{i=1}^{n-1} S_{ji}S_{in} \subset S_{jn}.$$

Мы проверили первое из соотношений а) и первое из соотношений б). Ясно, что остальные справедливы по тем же соображениям.

Итак, индукция оправдана и наше построение корректно.

Можно, разумеется, строить категорию с помощью любых S_{kn} и S_{nk}, удовлетворяющих условиям а) — в), а не только того специального вида, который мы использовали при доказательстве совместности этих условий. Таким образом, мы получаем широкое семейство конкретных категорий, в каждой из которых полностью произвольными являются лишь множества S_{n1}, все же остальные множества $S_{\alpha\beta}$ подчинены дополнительным условиям а)—в).

12.55. Покажем, что *любая категория* $\mathfrak{K}$ *с условиями* $\mathfrak{B}_{\alpha\beta}\mathfrak{B}_{\beta\alpha}\subset\mathfrak{B}(D)$ *принадлежит к этому семейству*. Действительно, в категории определены множества S_{n1} и S_{1n} для всех n; все остальные множества S_{nk} и S_{kn} должны удовлетворять условиям а)—в), а это и значит, что категория принадлежит к числу описанных.

Было бы интересно выяснить, какой вид имеют максимальные категории описанного семейства.

§ 12.6. Категории и прямые суммы

12.61. Пусть имеется категория $\mathfrak{K}$ с основными пространствами X_j, алгебрами $\mathfrak{B}_j$ и системами $\mathfrak{B}_{jk}$. Мы укажем, как построить категорию, в которой основными пространствами являются прямые суммы пространств X_j (в произвольных комбинациях) и основными алгебрами — соответствующие прямые суммы алгебр $\mathfrak{B}_j$. Пусть X_j — прямая сумма пространств $X_j^1, \ldots, X_j^{k_j}$, $\mathfrak{B}_j$ — прямая сумма соответствующих алгебр $\mathfrak{B}_j^1, \ldots, \mathfrak{B}_j^{k_j}$ (т. е. оператор $\bar{A}\in\overline{\mathfrak{B}}_j$ в пространстве X_j^s действует, как любой из операторов алгебры $\mathfrak{B}_j^s$). Нужно определить оператор $\bar{A}_{ji}\subset\mathfrak{B}_{ji}$. Мы задаем $\bar{A}_{ji}$ с помощью блок-матрицы

$$\bar{A}_{ji}=\begin{Vmatrix} A_{ji}^{11} & A_{ji}^{12} & \ldots & A_{ji}^{1k_i} \\ A_{ji}^{21} & A_{ji}^{22} & \ldots & A_{ji}^{2k_i} \\ \cdot & \cdot & \cdot & \cdot \\ A_{ji}^{k_j1} & A_{ji}^{k_j2} & \ldots & A_{ji}^{k_jk_i} \end{Vmatrix}, \tag{12}$$

где блок A_{ji}^{pq} отвечает произвольному оператору категории, который действует из пространства X_i^q в X_j^p, $p=1, \ldots, k_j$; $q=1, \ldots, k_i$. Покажем, что при этом получается категория. Действительно, если

$$\bar{B}_{il}=\begin{Vmatrix} B_{il}^{11} & B_{il}^{12} & \ldots & B_{il}^{1k_l} \\ B_{il}^{21} & B_{il}^{22} & \ldots & B_{il}^{2k_l} \\ \cdot & \cdot & \cdot & \cdot \\ B_{il}^{k_i1} & B_{il}^{k_i2} & \ldots & B_{il}^{k_ik_l} \end{Vmatrix},$$

то

$$\overline{A}_{ji}\overline{B}_{ij}=\begin{Vmatrix} A_{ji}^{11}B_{il}^{11}+A_{ji}^{12}B_{il}^{21}+\ldots+A_{ji}^{1k_i}B_{il}^{k_i1} & \ldots \\ \ldots & \ldots \\ \ldots & \ldots \\ \ldots & \ldots \end{Vmatrix}$$

и все получающиеся суммы произведений принадлежат снова к соответствующим системам согласно определению категории $\mathfrak{K}$. Итак, наше правило приводит к новой категории $\overline{\mathfrak{K}}$, которую будем называть *расширением* категории $\mathfrak{K}$.

12.62. Оказывается, что верно и обратное: если основные пространства X_j, фигурирующие в некоторой категории, суть прямые суммы некоторых пространств X_j^i, $i=1, \ldots, k_j$, и соответствующие алгебры $\overline{\mathfrak{B}}_j$ — прямые суммы алгебр $\mathfrak{B}_j^i$, $i=1, \ldots, k_j$, операторов, действующих в X_j^i, то вся категория $\mathfrak{K}'$ есть расширение $\overline{\mathfrak{K}}$ в указанном смысле некоторой категории $\mathfrak{K}$, построенной по пространству X_j^i и алгебрам $\mathfrak{B}_j^i$.

Действительно, пусть имеется категория $\mathfrak{K}$ указанного вида.

В базисе, выбранном в подпространствах X_j^i, матрица $\overline{A}_j$ алгебры $\mathfrak{B}_j$ имеет квазидиагональный вид

$$\overline{A}_j=\begin{Vmatrix} A_j^1 & & & & \\ & A_j^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & A_j^{k_j} \end{Vmatrix},$$

причем квадратная матрица A_j^i имеет r_j^i строк и столбцов.

Матрицу $\overline{A}_{ji}$ категории $\mathfrak{K}$ можно представить в форме блок-матрицы

$$\overline{A}_{ji}=\begin{Vmatrix} A_{ji}^{11} & A_{ji}^{12} & \ldots & A_{ji}^{1k_i} \\ A_{ji}^{21} & A_{ji}^{22} & \ldots & A_{ji}^{2k_i} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ A_{ji}^{k_j1} & A_{ji}^{k_j2} & \ldots & A_{ji}^{k_jk_i} \end{Vmatrix},$$

при этом блок A_{ji}^{pq} имеет r_j^p строк и r_i^q столбцов. С блоком A_{ji}^{pq} естественно можно связать оператор, переводящий пространство X_i^q в пространство X_j^p.

Используя все такие операторы, мы построим новую категорию $\mathfrak{K}_{ji}^{pq}$, для которой основными пространствами будут пространства X_i^q, основными алгебрами — алгебры $\mathfrak{B}_i^q$ и основными системами — системы $\mathfrak{B}_{ji}^{pq}$ операторов, определенных матрицами A_{ji}^{pq}. Докажем, что указанный набор объектов определяет категорию.

Пусть имеются оператор A_{ji}^{pq}, действующий из X_i^q в X_j^p, и оператор A_{kj}^{rp}, действующий из X_j^p в X_k^r; покажем, что их произведение A_{kj}^{rp} принадлежит классу $\mathfrak{B}_{ki}^{rq}$. Действительно, в категории $\mathfrak{K}$ имеется матрица, в которой все блоки, кроме (p, q)-го, равны нулю, а блок с индексами (p, q) занимает матрица A_{ji}^{pq}, и есть другая матрица с единственным отличным от нуля блоком A_{kj}^{rp}. Произведение этих матриц принадлежит категории $\mathfrak{K}$, и в нем имеется единственный ненулевой блок $A_{kj}^{rp}A_{ji}^{pq}$, что и требуется.

Итак, все условия категории выполнены; правда, не определены еще операторы, действующие из пространства X_j^i в X_j^s с тем же нижним индексом. Однако можно положить все такие операторы равными нулю, что не нарушит требований категории.

12.63. Так как каждая полупростая алгебра операторов, действующих в пространстве X_j, позволяет разбить пространство X_j в прямую сумму пространств X_j^i, в которых действует уже простая алгебра, мы видим, что структура общей категории с полупростыми алгебрами приводится к структуре категории с простыми алгебрами. (Этот вопрос мы рассмотрели в § 12.4.) В соответствующем базисе матрица каждого оператора $\mathfrak{A}_{ji}$ категории $\mathfrak{K}$ имеет вид (11), причем каждый из блоков A_{ji}^{pq} есть оператор семейства $\mathfrak{B}_{ji}^{pq}$ некоторой категории $\mathfrak{K}_{ji}^{pq}$ с основными пространствами X_j^q и X_i^q и простыми алгебрами $\mathfrak{B}_j^q$ и $\mathfrak{B}_i^q$. Некоторые блоки матрицы $\mathfrak{A}_{ji} \in \mathfrak{B}_{ji}$ могут быть тождественно (для всех $\mathfrak{A}_{ji}$) равными нулю. Если обозначить их множество через S_{ji}, то возникает вопрос, как связаны между собой эти множества при разных индексах j и i. Подобный вопрос для случая одномерных блоков мы рассматривали в § 12.5. Метод, использованный там, применим и в данном случае и приводит к следующему результату. *Если категория, определяемая скрещением j-й блок-строки и i-го блок-столбца матрицы $\mathfrak{A}_{21} \in \mathfrak{B}_{21}$ типа $\mathfrak{K}_1$ (12.31) или $\mathfrak{K}_3$ (12.41), т. е. связана с обратимыми матрицами, то в i-й блок-строке и в j-м блок-столбце матрицы $\mathfrak{A}_{12}$ все блоки, кроме находящегося в их скрещении, определяют нулевые категории. Если указанная категория типа $\mathfrak{K}_2$ (12.31), то в указанных блоках находятся матрицы категории типа $\mathfrak{K}_2$, дающие в произведении с данной нуль.*

Это позволяет судить и об устройстве общей категории, как в § 12.5.

Замечание. А. Я. Хелемский (Вестник МГУ, серия Математика и механика, 1963, № 4, стр. 49—55) нашел категории, соответствующие нильпотентным алгебрам $\mathfrak{B}$.

ОТВЕТЫ И УКАЗАНИЯ К ЗАДАЧАМ

К главе 1

1. *Ответ.* а) $+$, б) $+$.

2. *Ответ.* $a_{11}a_{32}a_{23}a_{44}$, $a_{41}a_{12}a_{23}a_{34}$, $a_{31}a_{42}a_{23}a_{44}$.

3. *Ответ.* $(-1)^{n(n-1)/2}$.

4. *Указание.* Рассмотреть определитель, все элементы которого равны единице.

5. *Ответ.* $\Delta=(mq-np)(ad-bc)$.

6. *Указание.* Умножить первый столбец на 10^4, второй — на 10^3, третий — на 10^2, четвертый — на 10^1 и прибавить к последнему столбцу; использовать, далее, следствие *1.45*.

7. *Ответ.* $\Delta_1=-29\,400\,000$, $\Delta_2=394$.

8. *Указание.* Очевидно, $P(x)$ есть многочлен 4-й степени. Можно подсчитать его старший коэффициент, а затем определить его корни из условия совпадения строк определителя.

Ответ. $P(x)=-3(x^2-1)(x^2-4)$.

9. *Указание.* Прибавить все столбцы к первому.

Ответ. $\Delta=[x+(n-1)a](x-a)^{n-1}$.

10. *Указание.* Заменив x_n на x и использовав идею решения задачи 8, получить соотношение

$$\Delta(x_1, x_2, \ldots x_{n-1}, x)=\Delta(x, \ldots, x_{n-1})\cdot(x-x_1)\ldots(x-x_{n-1}).$$

Ответ. $\Delta(x_1, x_2, \ldots, x_n)=(x_2-x_1)(x_3-x_1)(x_3-x_2)\ldots\times$
$\times(x_n-x_{n-1})\ldots(x_n-x_{n-1})$.

11. *Ответ.* $C_1=0$, $C_2=2$, $C_3=-2$, $C_4=0$, $C_5=3$.

12. *Ответ.*

$$\sum M_{j_1,\, j_2,\, \ldots,\, j_k}^{i_1,\, i_2,\, \ldots,\, i_k}\cdot \overline{A}_{j_1,\, j_2,\, \ldots,\, j_k}^{i'_1,\, i'_2,\, \ldots,\, i'_k}=0,$$

где $i_1<i_2<\ldots<i_k$ и $i'_1<i'_2<\ldots<i'_k$ фиксированы, причем хотя бы одно из i_α не совпадает с соответствующим i'_α.

13. *Указание.* Достаточно, чтобы соответствующий определитель 4-го порядка был отличен от нуля.

14. *Указание.* Использовать результаты *1.96—1.97*.

К главе 2

1. *Ответ.* Нет, ибо в пределах этой совокупности нельзя умножить на -1.

2. *Ответ.* Нет, ибо в пределах этой совокупности нельзя сложить два вектора, симметричных относительно заданной прямой.

3. *Ответ.* Да. В частности, «нулем» пространства Р служит число $1 \in \text{P}$.

4. *Указание.* Использовать *1.96*.

5. *Указание.* Предполагая наличие зависимости

$$C_1 t^{\alpha_1} + C_2 t^{\alpha_2} + \ldots + C_k t^{\alpha_k} \equiv 0,$$

разделить на t^{α_k} и продифференцировать, далее использовать индукцию по k.

6. *Указание.* Показать, что нуль-вектор также допускает единственное разложение по системе $e_1, e_2, \ldots, e_n$. Отсюда вывести линейную независимость векторов этой системы.

7. *Ответ.* Да, из одного вектора — любого элемента $x \in P$, отличного от 1.

8. *Ответ.* 1.

9. *Ответ.* Пересечение — прямая пересечения двух плоскостей в обычном смысле. Сумма — все пространство.

10. Использовать *2.34*.

11. *Ответ.* Нет. Его можно заменить на любой другой вектор этой гиперплоскости.

12. *Ответ.* В «точечной» интерпретации: каждая гиперплоскость вместе с любыми двумя своими точками содержит проходящую через них прямую.

13. *Ответ.* В общем случае $p+q+1$, если это число не превосходит размерности всего пространства.

14. *Ответ.* $p+q+r+2$, если это число не превосходит размерности всего пространства.

15. *Ответ.* Поставить в соответствие каждому положительному числу его логарифм.

К главе 3

1. *Указание.* В матрице ранга 1 столбцы пропорциональны.

2. *Указание.* Нужно так написать условия принадлежности вектора y подпространству L, чтобы в них участвовали только миноры k-го порядка матрицы A. Но $y \in \text{L}$ тогда и только тогда, когда матрица B, полученная присоединением к матрице A столбца из координат вектора y, имеет ранг k, или, что то же самое, каждый ее минор $(k+1)$-го порядка равен нулю. Разлагая каждый минор $(k+1)$-го порядка матрицы B по последнему столбцу, можно получить некоторую систему уравнений относительно координат вектора y с коэффициентами — минорами k-го порядка матрицы A.

3. *Указание.* Использовать *1.51—1.52*.

4. *Ответ.* $x = (c_1, c_2, c_3, c_4, c_5)$, где $c_1 = -16 + c_3 + c_4 + 5c_5$, $c_2 = 23 - 2c_3 - 2c_4 - 6c_5$.

5. *Ответ.* Если $(\lambda - 1)(\lambda + 2) \neq 0$, то

$$x = -\frac{\lambda+1}{\lambda+2}, \quad y = \frac{1}{\lambda+2}, \quad z = \frac{(\lambda+1)^2}{\lambda+2}.$$

Если $\lambda = 1$, система имеет решения, зависящие от двух параметров. Если $\lambda = -2$, система несовместна.

6. *Ответ.* Если матрицы $\begin{Vmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{Vmatrix}$ и $\begin{Vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{Vmatrix}$ имеют одинаковый ранг.

7. *Ответ.* Матрицы $\begin{Vmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \dots & \dots \\ a_n & b_n \end{Vmatrix}$ и $\begin{Vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ \dots & \dots & \dots \\ a_n & b_n & c_n \end{Vmatrix}$ имеют одинаковый ранг.

8. *Ответ.*

$$x^{(1)} = (1, -2,\ 1,\ 0,\ 0),$$
$$x^{(2)} = (1, -2,\ 0,\ 1,\ 0),$$
$$x^{(3)} = (5, -6,\ 0,\ 0,\ 1).$$

9. *Ответ.* Например,

$$x = \begin{Vmatrix} -16 \\ 23 \\ 0 \\ 0 \\ 0 \end{Vmatrix} + c_1 \begin{Vmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{Vmatrix} + c_2 \begin{Vmatrix} 1 \\ -2 \\ 0 \\ 1 \\ 0 \end{Vmatrix} + c_3 \begin{Vmatrix} 5 \\ -6 \\ 0 \\ 0 \\ 1 \end{Vmatrix}.$$

Здесь в первом столбце выписаны координаты вектора x_0 — частного решения неоднородной системы; в остальных столбцах — координаты векторов $y^{(1)}$, $y^{(2)}$, $y^{(3)}$, образующих нормальную фундаментальную систему решений соответствующей однородной системы.

10. *Ответ.* Ранг A_1 равен 3; базисный минор, например, в левом верхнем углу. Ранг A_2 равен пяти, базисный минор совпадает с определителем матрицы.

11. *Указание.* Перевести минор M в левый верхний угол и затем, применив процедуру *3.62*, показать, что столбцы матрицы A, начиная с $(r+1)$-го, можно сделать нулевыми.

12. *Указание.* Если $P \neq 0$, искать матрицу A в виде

$$\begin{Vmatrix} P & 0 & x \\ 0 & 1 & y \end{Vmatrix}.$$

13. *Указание.* Или ранг матрицы $\| a_{jk} \|$ равен n, или он меньше n.

14. *Указание.* Использовать теорему Кронекера — Капелли.

15. *Указание.* Использовать результат задачи 14.

16. *Указание.* См. *6.63* (18).

К главе 4

1. *Ответ.* Также n.

2. *Ответ.* в) и ж).

3. *Ответ.* Да.

4. *Ответ.*

$$\text{а) } A_{(e)} = \begin{Vmatrix} -1 & -1 & 2 \\ 1 & -3 & 3 \\ -1 & -5 & 5 \end{Vmatrix}; \quad \text{б) } A_{(x)} = \begin{Vmatrix} 2 & 0 & -2 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{Vmatrix}.$$

5. *Ответ.* $\mathrm{AB\,AB} \neq \mathrm{A}^2\mathrm{B}^2$.

6. *Ответ.* $\mathrm{AB} - \mathrm{BA} = \mathrm{E}$.

7. *Ответ.* $(\mathrm{A}+\mathrm{B})^2 = \mathrm{A}^2 + \mathrm{AB} + \mathrm{BA} + \mathrm{B}^2$, $(\mathrm{A}+\mathrm{B})^3 = \mathrm{A}^3 + \mathrm{A}^2\mathrm{B} + \mathrm{ABA} + \mathrm{AB}^2 + \mathrm{BA}^2 + \mathrm{BAB} + \mathrm{B}^2\mathrm{A} + \mathrm{B}^3$.

8. *Указание.* Применить метод индукции.

9. *Ответ.* Размерность пространства равна mn. В качестве базисных операторов можно взять такие, которым соответствуют матрицы A_{ij} с единственным ненулевым элементом на пересечении i-й строки и j-го столбца.

10. *Ответ.* $AB = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$.

11. *Ответ.* $A^n = \begin{Vmatrix} 1 & n \\ 0 & 1 \end{Vmatrix}$, $B^n = \begin{Vmatrix} \cos n\varphi & -\sin n\varphi \\ \sin n\varphi & \cos n\varphi \end{Vmatrix}$.

12. *Ответ.* $A = \begin{Vmatrix} a & b \\ c & -a \end{Vmatrix}$, где $bc = -a^2$.

13. *Ответ.*

$$\text{а)} \begin{Vmatrix} -9 & -2 & -10 \\ 6 & 14 & 8 \\ -7 & 5 & -5 \end{Vmatrix}, \quad \text{б)} \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}.$$

15. *Указание.* Найти след обеих частей равенства.

16. *Указание.* Три уравнения для неизвестных элементов матриц A и B приводятся к уравнениям для трех миноров неизвестной матрицы с двумя строками и тремя столбцами. Использовать задачу 12 к гл. 3.

17. *Указание.* Использовать *4.54*.

18. *Указание.* Записать элементы минора M через элементы, находящиеся в первых r строках, и применить теорему *4.54*.

19. *Указание.* Использовать решение задачи 18.

20. *Указание.* См. *4.54*.

21. *Ответ.*

$$A^{-1} = \begin{Vmatrix} 5 & -2 \\ -2 & 1 \end{Vmatrix}, \quad B^{-1} = \begin{Vmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{Vmatrix}, \quad C^{-1} = C.$$

23. *Ответ.* Если A — нулевая матрица, то X — любая. Если $\det A \neq 0$, то X — нулевая матрица. Если $\det A = 0$, но A — ненулевая матрица, то ее строки пропорциональны; пусть $\alpha:\beta$ есть отношение соответствующих элементов первой и второй строк матрицы A; тогда

$$X = \begin{Vmatrix} -\beta p & \alpha p \\ -\beta q & \alpha q \end{Vmatrix}$$

при любых p и q.

24. *Указание.* Использовать *1.51—1.52*.

25. *Ответ.* Нет.

26. *Указание.* Например, $\mathrm{A}_1[a_0 + a_1 t + \cdots] = \mathrm{A}[a_0 + a_1 t + \cdots] + \lambda a_0$.

27. *Указание.* Оператор A переводит линейно независимые векторы снова в линейно независимые.

28. *Указание.* Применить равенство $AB = BA$ к собственному вектору оператора B.

31. *Указание.* Использовать результат задачи 30.

32. *Указание.* Подбирая должным образом оператор B и используя задачу 28, свести решение к задаче 31.

34. *Указание.* Использовать разложение оператора $A^2 - \mu^2 E$ на множители.

35. *Ответ.* а) $\lambda_1 = 2$, $f_1 = (1, 0, 0)$; $\lambda_2 = 1$, $f_2 = (1, 0, 1)$; $\lambda_3 = -1$, $f_3 = (0, 1, -1)$; б) $\lambda_1 = -1$; $f_1 = (1, 0, 0)$; $\lambda_2 = \lambda_3 = 1$, $f_2 = (1, 0, 1)$; $f_3 = (0, 1, 1)$; в) $\lambda_1 = 2$, $f_1 = (1, 0, 0)$; г) $\lambda_1 = 1$, $f_1 = (1, 0, 0, -1)$; $\lambda_2 = (0, 1, 0, 0)$.

36. *Указание.* Включение $T(A^k) \subset N(A^m)$ необходимо и достаточно для равенства $A^{k+m} = 0$

37. *Указание.* Пусть $f_1, \ldots, f_r$ — базис области значений оператора A, так что для любого $x \in K_n$ имеем

$$Ax = \sum_{j=1}^{r} a_j(x) f_j.$$

Положить $A_j x = a_j(x) f_j$ $(j = 1, \ldots, n)$.

К главе 5

1. *Указание.* Первый вектор нового базиса есть x.

2. *Указание.* Выбрать новый базис $f_1, f_2, \ldots, f_n$ так, чтобы первые k векторов составляли базис подпространства K'. Записать условие $x \in K'$ системой координатных уравнений в новом базисе. Используя формулы перехода, построить соответствующую систему координатных уравнений в исходном базисе.

3. *Указание.* Использовать определение гиперплоскости и задачу 2.

4. *Ответ.* Матрица искомого преобразования $C = BA^{-1}$.

5. *Указание.* Пусть $e_1, \ldots, e_n$ — произвольный базис в K_n и $f(x) = \sum_{k=1}^{n} c_k \xi_k$, где $\xi_1, \ldots, \xi_n$ — координаты вектора x. Формулы преобразования координат начать с уравнения

$$\eta_1 = \sum_{k=1}^{n} c_k \xi_k.$$

6. *Указание.* Использовать *4.83* и инвариантность характеристического многочлена (*5.53*).

7. *Указание.* Выбрать базис так, чтобы первые m его векторов лежали в подпространстве $R^{(\lambda_0)}$. Показать, что для этого базиса многочлен $\det \| A_{(f)} - \lambda E \|$ имеет множителем $(\lambda - \lambda_0)^m$. Использовать инвариантность характеристического многочлена (*5.53*).

К главе 6

1. *Ответ.* В базисе $x_n, x_{n-1}, \ldots, x_1$.

2. *Указание.* См. *6.44*.

3. *Ответ:*

$$\left\|\begin{matrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{matrix}\right\|.$$

4. *Ответ.* Нет: $E_2(A) = (\lambda-2)(\lambda-1)^2$, $E_2(B) = (\lambda-1)(\lambda^2-5\lambda-2)$.

5. *Ответ.* $E_{n-1}(A_1) = E_{n-1}(A_2) = (1-\lambda)^n$, $E_{n-2}(A_1) = E_{n-2}(A_2) = 1$;

$$E_{n-1}(A_3) = (n-\lambda)^n, \quad E_{n-2}(A_3) = 1;$$

$$E_{n-1}(A_4) = \prod_{k=1}^{n} (\lambda-k), \quad E_{n-2}(A_4) = 1.$$

6. *Указание.* $E_{n-1}(A) = (\alpha-\lambda)^n$, $E_{n-2}(A) = 1$.

7. *Ответ.* Диагональная матрица, на диагонали которой стоят некоторые из корней многочлена $P(\lambda)$.

8. *Ответ.* На диагонали жордановых клеток стоят некоторые из корней многочлена $P(\lambda)$, и размеры клеток не больше кратности соответствующего корня.

9. *Указание.* Векторы x, $\mathrm{A}x$, A^2x линейно зависимы.

10. *Ответ.* Многочлены от $\mathrm{A}_m(a)$.

11. *Ответ.* Матрицы вида

$$B_{mn} = \left\|\begin{matrix} b_1 & b_2 & b_3 & & \ldots b_n \\ 0 & b_1 & b_2 & & \ldots b_{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots\ b_1 & \ldots b_{n-m+1} \end{matrix}\right\| \quad (n \geqslant m)$$

или

$$B_{mn} = \left\|\begin{matrix} b_1 & b_2 & \ldots & b_n \\ 0 & b_1 & \ldots & b_{n-1} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & b_1 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \end{matrix}\right\| \quad (n \leqslant m).$$

12. *Ответ.* Матрицы вида

$$\left\|\begin{matrix} B_{m_1m_1} & B_{m_1m_2} & \ldots & B_{m_1m_k} \\ \cdot & \cdot & \cdot & \cdot \\ B_{m_km_1} & B_{m_km_2} & \ldots & B_{m_km_k} \end{matrix}\right\|,$$

где блоки $B_{m_im_j}$ указаны в задачах 10 и 11.

13. *Ответ.* Матрицы вида

$$\left\|\begin{matrix} B_{m_1m_1} & 0 & \ldots & 0 \\ 0 & B_{m_2m_2} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & B_{m_km_k} \end{matrix}\right\|.$$

14. *Ответ.* Для каждой группы жордановых блоков с одним и тем же характеристическим корнем — блок, указанный в задаче 12. Остальные элементы — нули.

15. *Ответ.* Если кратность каждого характеристического корня равна размеру соответствующей жордановой клетки (или: характеристический многочлен совпадает с минимальным; или: все элементарные делители, кроме самого старшего, равны 1).

16. *Указание.* Использовать *6.63* (18).

К главе 7

1. *Ответ.* Тензор второго ранга, два раза ковариантный.

2. *Ответ.* Например,

$$\eta_1^2-\eta_2^2-\eta_3^2,$$

где

$$\eta_1=\frac{1}{2}\xi_1+\frac{1}{2}\xi_2+\xi_3,\quad \eta_2=\frac{1}{2}\xi_1-\frac{1}{2}\xi_2,\quad \eta_3=\xi_3.$$

3. *Указание.* Использовать *7.93.*

4. *Указание.* Использовать *4.54.* и *7.15.*

5. *Ответ.* Например, $A(x, y)=\sigma_1\tau_1+\sigma_2\tau_2+\sigma_3\tau_3$, где σ_i и τ_i $(i=1, 2, 3)$ — новые координаты векторов x и y. При этом формулы перехода к новому базису следующие:

$$\sigma_1=\xi_1+\xi_2,\quad \sigma_2=\xi_2+2\xi_3,\quad \sigma_3=\xi_3.$$

6. *Указание.* Сначала изменить нумерацию координат таким образом, чтобы матрица билинейной формы $A(x, y)$ преобразовалась к виду, допускающему применение метода Якоби.

7. *Указание.* $\|-a_{ik}\|$ есть матрица положительно определенной формы.

$$\textit{Ответ.}\ a_{11}<0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{vmatrix}>0,\ \ldots,\ (-1)^n\det\|a_{ik}\|>0.$$

8. *Указание.* Использовать замечание к *7.96.*

9. *Указание.* Рассмотреть эту форму на базисных векторах.

10. *Указание.* Последняя строка определителя состоит из элементов

$$a_k^{(n)}=(-1)^{k-1}A(e_1, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n),\ k=1, \ldots, n.$$

11. *Указание.* Первую пару базисных векторов e_1, e_2 найти из уравнения $A(e_1, e_2)=1$; построить подпространство L, определяемое уравнениями $A(e_1, x)=0$, $A(e_2, x)=0$, и, если форма $A(x, y)$ в нем не равна тождественно нулю, определить векторы $e_3, e_4\in$ L так, чтобы $A(e_3, e_4)=1$; и т. д.

12. *Указание.* Рассмотреть форму $A(x, x)+\varepsilon\sum_{j=1}^{n}x_j^2$ $(\varepsilon>0)$ и применить критерий *7.96.*

13. *Указание.* Пусть $x^{(j)}=\{\xi_1^{(j)}, \ldots, \xi_n^{(j)}\}$ $(j=1, \ldots, r)$ — базис подпространства K′. Тогда K″ состоит из векторов $y=\{\eta_1, \ldots, \eta_n\}$, удовлетворяющих системе

$$\mathrm{A}(x^{(j)}, y)=\sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_{ik}\xi_i^{(j)}\right)\eta_k=0, \qquad j=1, \ldots, r.$$

Матрица коэффициентов системы есть произведение невырожденной матрицы $\|a_{ik}\|$ формы $\mathrm{A}(x, y)$ и матрицы $\|\xi_i^{(j)}\|$ ранга r. Использовать *4.67*.

14. *Ответ.* $\mathrm{A}'=\mathrm{A}$.

15. *Указание.* Если $y=\{\eta_1, \ldots, \eta_n\}$ есть решение системы (44), то $(b, y)=(\mathrm{A}x, y)=(x, \mathrm{A}'y)=0$. Обратно: система (44) есть условие сопряженности вектора y и векторов $a_j=\{a_{j1}, \ldots, a_{jn}\}$; если $(b, y)=0$ для всех таких y, то x лежит в линейной оболочке векторов $a_1, \ldots, a_n$.

16. *Указание.* См. задачу 37 к гл. 4.

17. *Указание.* См. задачу 1 к гл. 3.

18. *Указание.* Сначала рассмотреть случай неотрицательных форм ранга 1, использовать задачу 17. Затем применить задачу 16.

К главе 8

1. *Ответ.* Нет, ибо не выполняется аксиома *8.21* б) и аксиома *8.21* а) для $\lambda=-1$.

2. *Ответ.* Нет, ибо не выполняется аксиома *8.21* б).

3. *Ответ.* Можно. Это равносильно изменению масштаба на осях.

4. *Указание.* Обозначить через e_1, e_2, e_3 векторы, идущие из фиксированной вершины тетраэдра по трем его ребрам, и найти векторные выражения для остальных его ребер.

Ответ. 90°.

5. *Ответ.* 90°, 60°, 30°.

6. *Ответ.*

$$\sqrt{\int_a^b (x(t)+y(t))^2\,dt}\ \begin{cases} \leqslant \sqrt{\displaystyle\int_a^b x^2(t)\,dt}+\sqrt{\displaystyle\int_a^b y^2(t)\,dt}, \\ \geqslant \left|\sqrt{\displaystyle\int_a^b x^2(t)\,dt}-\sqrt{\displaystyle\int_a^b y^2(t)\,dt}\right|. \end{cases}$$

7. *Ответ.* $\cos\varphi=\dfrac{1}{\sqrt{n}}$.

8. *Ответ.* а) $g=(3, 1, -1, -2)$, $h=(2, 1, -1, 4)$,

б) $g=(1, 7, 3, 3)$, $h=(-4, -2, 6, 0)$.

9. *Указание.* Использовать определение угла *8.33* и ортогональность вектора h ко всем векторам подпространства R′.

10. *Указание.* Умножить скалярно равенство *8.51* (18) на вектор g_0.

11. *Указание.* Использовать *8.54* и *8.52*.

12. *Ответ.* $y_1 = \boldsymbol{i}$, $y_2 = y_3 = 0$, $y_4 = -2\boldsymbol{j}$, $y_5 = 0$, $y_6 = 5\boldsymbol{k}$.

13. *Ответ.* (1, 2, 1, 3); (10, −1, 1, −3); (19, −87, −61, 71).

14. *Указание.* Предполагая, что размерность R'' больше, чем размерность R', рассмотреть вектор $e'' \in \mathrm{R}''$, ортогональный к проекции подпространства R' на R''; использовать задачу 10.

15. *Ответ.* $A_n = \dfrac{(2n)!}{2^n (n!)^2}$.

16. *Ответ.* $P_n(-1) = (-1)^n$.

17. *Указание.* Записать искомые коэффициенты через скалярные произведения.

18. *Указание.* Использовать решения задач 15 и 16.

19. *Указание.* Разложить $Q(t)$ по многочленам Лежандра.

Ответ. $Q(t) = \dfrac{1}{A_n} P_n(t)$.

20. *Ответ.* $\| P_n(t) \|^2 = \dfrac{2}{2n+1}$.

21. *Ответ.* $k(A) = |\det A|$.

22. *Указание.* Использовать *4.75*.

23. *Указание.* Речь идет о том, чтобы сравнить высоты двух гиперпараллелепипедов.

24. *Указание.* Неравенства

$$\frac{V[x_1, x_2, \ldots, x_m]}{V[x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_m]} \leqslant \frac{V[x_1, \ldots, x_k]}{V[x_1, \ldots, x_{k-1}]}$$
$$(k = 1, 2, \ldots, m)$$

получаются из неравенства задачи 23. Перемножить их для всех $k = 1, 2, \ldots, m$, произвести сокращения и извлечь корень $(m-1)$-й степени. Геометрический смысл: объем m-мерного гиперпараллелепипеда не превосходит произведения корней $(m-1)$-й степени из объемов его $(m-1)$-мерных граней.

25. *Указание.* Написать неравенство задачи 24. для $x_{s_1}, x_{s_2} \ldots$ $\ldots, x_{s_r}$ и перемножить такие неравенства для всех допустимых значений $s_1, s_2, \ldots, s_r$.

26. *Указание.* Требуется построить в 2^m-мерном пространстве гиперпараллелепипед, у которого проекции ребер на каждую ось не превосходят по абсолютной величине числа M и объем которого в точности равен $M^n n^{n/2}$.

Для $M = 1$ матрица A_m координат искомых векторов в 2^m-мерном пространстве может быть задана рекуррентной формулой

$$A_m = \left\| \begin{matrix} A_{m-1} & A_{m-1} \\ A_{m-1} & -A_{m-1} \end{matrix} \right\|, \quad A_1 = \left\| \begin{matrix} 1 & 1 \\ 1 & -1 \end{matrix} \right\|.$$

Примечание. Для $n \neq 2^m$ оценка $M^n n^{n/2}$ может быть улучшена.

27. *Указание.* Пусть $G^{\perp}$ означает ортогональное дополнение подпространства G в R. Для каждого $x \in N(A)$ и для каждого $z \in R$

$$(A'z,\ x) = (z,\ Ax) = 0,$$

откуда $A'z \in N^{\perp}(A)$, т. е. $T(A') \subset N^{\perp}(A)$, $T^{\perp}(A') \supset N(A)$. Для каждого $x \in T^{\perp}(A)$ и для каждого $y \in R$

$$(A'x,\ y) = (x,\ Ay) = 0,$$

откуда $A'x = 0$, т. е. $x \in N(A')$; таким образом, $T^{\perp}(A) \subset N(A')$, $T^{\perp}(A') \subset N(A)$. Отсюда $N(A) = T^{\perp}(A')$, $N^{\perp}(A) = T(A')$. Аналогично доказывается второе утверждение.

28. *Указание.* Сравнить *8.93* с *1.51—1.52*.

29. *Указание.* Использовать *4.54*.

30. *Указание.* Углы треугольника однозначно определяются по его сторонам. Другой метод: симметричная билинейная форма $(Qx,\ Qy)$ однозначно восстанавливается по квадратичной $(Qx,\ Qx)$.

31. *Указание.* Данный равноугольный оператор C преобразует ортогональный и нормированный базис $e_1, e_2, \ldots, e_n$ в ортогональный базис $f_1' = \alpha_1 f_1$, $f_2' = \alpha_2 f_2$, $\ldots$, $f_n' = \alpha_n f_n$, где $f_1, f_2, \ldots, f_n$ нормированы. Пусть Q — изометрический оператор, переводящий векторы $f_1, f_2, \ldots, f_n$ в $e_1, e_2, \ldots, e_n$; тогда матрица равноугольного оператора QC диагональна. Показать, что условие $\alpha_i = \alpha_j$ позволяет построить пару ортогональных векторов, которые переходят в неортогональные в результате применения оператора QC.

32. *Указание.* Достаточно показать, что Q — равноугольный оператор (задача 31). В предположении, что имеется прямой угол, который переходит не в прямой, построить параллелограмм, площадь которого в результате применения оператора Q изменится.

33. *Указание.* Обобщить конструкцию задачи 32.

34. *Указание.* Применив к данным системам процесс ортогонализации, получить ортогональные и нормированные системы $e_1, e_2, \ldots$ и $f_1, f_2, \ldots$ Используя *8.53*, показать, что векторы $x_1, x_2, \ldots$ выражаются через $e_1, e_2, \ldots$ по таким же формулам, по каким векторы $y_1, y_2, \ldots$ выражаются через $f_1, f_2, \ldots$ Оператор Q задать путем отображения системы $e_1, e_2, \ldots$ на систему $f_1, f_2, \ldots$

35. *Указание.* Рассмотреть конечные системы $e_1', e_1'', e_2', e_2'', \ldots$ $\ldots, e_k', e_k''$ и $f_1', f_1'', f_2', f_2'', \ldots, f_k', f_k''$, полученные при определении углов между подпространствами R' и R'' и подпространствами S' и S''. Согласно построению

$$(e_i',\ e_i'') = (f_i',\ f_i'') = \cos\varphi_i,\ (i = 1,\ 2,\ \ldots,\ k),$$

$$(e_i',\ e_j') = (f_i',\ f_j') = 0,\ (e_i'',\ e_j'') = (f_i'',\ f_j'') = 0 \qquad (i \neq j);$$

далее показать, что $(e_i',\ e_j'') = (f_i',\ f_j'') = 0$ (используя задачу 9). Затем применить результат задачи 34.

36. *Указание.* Использовать задачу 11.

37. *Указание.* Выбрать в подпространствах L_1 и L_2 базисы $e_1, \ldots, e_m$ и $f_1, f_2, \ldots, f_m$, полученные при построении углов $\alpha_1, \ldots, \alpha_m$.

В пространстве R построить базис $e_1, e_2, \ldots, e_m, e_{m+1}, \ldots, e_n$, первые векторы которого получены путем ортогонализации векторов $e_1, e_2, \ldots, e_m, f_1, f_2, \ldots, f_m$. Разложить векторы $x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_m$ по построенному базису. Показать, что матрицы этих разложений имеют только по одному минору m-го порядка, если не считать миноров, заведомо равных нулю. Использовать, далее, выражение объема параллелепипеда через миноры соответствующей матрицы.

38. *Указание.* Использовать задачу 2 к гл. 3 и задачу 17 к гл. 4.

39. *Указание.* Проверить утверждение задачи в специальном базисе, первые k векторов которого принадлежат к подпространству $L\{x_1, x_2, \ldots, x_k\}$. Для перехода к общему случаю использовать задачу 17 к гл. 4, причем показать, что $\det \| a_i^{(j)} \| = 1$.

40. *Указание.* Сначала рассмотреть случай $k=2$.

41. *Указание.* Выбрать в пространстве R базис, как указано в задаче 37, и проверить утверждение задачи в этом базисе. Для перехода к общему случаю действовать так же, как в задаче 39.

42. *Указание.* Использовать *4.54.*

43. *Указание.* Рассмотреть ортогональное дополнение Z к инвариантному (относительно A) подпространству H всех векторов x, для которых $P(A)x=0$. Подпространство Z также инвариантно относительно оператора A, следовательно, и относительно $[P(A)]^{k-1}$. Но если $z \in Z$, то $[P(A)]^{k-1}z \in H$, откуда $[P(A)]^{k-1}z=0$. Отсюда получить, что $[P(t)]^{k-1}$ есть аннулирующий многочлен оператора A.

К главе 9

1. *Указание.* Использовать *9.34.*

2. *Указание.* Оператор B имеет базис из собственных векторов $e_1, \ldots, e_n$ с положительными собственными значениями $\mu_1, \ldots, \mu_n$. Отсюда $B^2e_i=\mu_i^2e_i$, и если мы желаем удовлетворить равенству $B^2=A$, необходимо, чтобы e_i были собственными векторами оператора A, а числа μ_i^2 совпадали с λ_i. Но этого и достаточно для $B^2=A$.

3. *Указание.* Предварительно преобразовать базис так, чтобы матрица данного оператора приняла диагональный вид.

$$\textit{Ответ.}\quad \sqrt{A}=\left\|\begin{matrix}3&2&0\\2&4&2\\0&2&5\end{matrix}\right\|.$$

4. *Указание.* Оператор A′A самосопряжен, и выражение $(A'Ax, x)=(Ax, Ax)$ неотрицательно при любом $x \in R_n$. Если A невырожден, то это выражение положительно при любом $x \in R_n$.

5. *Указание.* Использовать равенство $(AB)^*=B^*A^*$.

6. *Указание.* Оператор AA^* самосопряжен и положителен (задача 4), поэтому можно найти самосопряженный и положительный оператор S так, чтобы иметь $S^2=AA^{*\prime}$. Далее построить оператор Q по формуле $Q=S^{-1}A$ и показать, что Q—унитарный оператор.

7. *Указание.* Использовать задачи 5 и 2.

8. *Указание.* Пусть $R' \subset R_n$—подпространство, порожденное собственными векторами оператора A′A с ненулевыми собственными значениями, и R″—ортогональное дополнение к R′. На R′ положить V равным унитарной составляющей A (так что $\sqrt{A'A}\,Vx = Ax$), на R'' положить $Vx = 0$.

9. *Указание.* Использовать задачи 28 и 29 к гл. 4.

10. *Указание.* Применить к векторам жорданова базиса оператора A (*6.37*) процесс ортогонализации.

К главе 10

1. *Ответ.*

а) $4\eta_1^2+\eta_2^2-2\eta_3^2$;
$$\eta_1=\frac{2}{3}\xi_1-\frac{2}{3}\xi_2+\frac{1}{3}\xi_3,$$
$$\eta_2=\frac{2}{3}\xi_1+\frac{1}{3}\xi_2-\frac{2}{3}\xi_3,$$
$$\eta_3=\frac{1}{3}\xi_1+\frac{2}{3}\xi_2+\frac{2}{3}\xi_3;$$

б) $10\eta_1^2+\eta_2^2+\eta_3^2$;
$$\eta_1=\frac{1}{3}\xi_1+\frac{2}{3}\xi_2-\frac{2}{3}\xi_3,$$
$$\eta_2=\frac{2}{\sqrt{5}}\xi_1-\frac{1}{\sqrt{5}}\xi_2,$$
$$\eta_3=\frac{2}{3\sqrt{5}}\xi_1+\frac{4}{3\sqrt{5}}\xi_2+\frac{\sqrt{5}}{3}\xi_3;$$

в) $\eta_1^2-\eta_2^2+3\eta_3^2+5\eta_4^2$;
$$\eta_1=\frac{1}{2}\xi_1+\frac{1}{2}\xi_2+\frac{1}{2}\xi_3+\frac{1}{2}\xi_4,$$
$$\eta_2=\frac{1}{2}\xi_1+\frac{1}{2}\xi_2-\frac{1}{2}\xi_3-\frac{1}{2}\xi_4,$$
$$\eta_3=\frac{1}{2}\xi_1-\frac{1}{2}\xi_2+\frac{1}{2}\xi_3-\frac{1}{2}\xi_4,$$
$$\eta_4=\frac{1}{2}\xi_1-\frac{1}{2}\xi_2-\frac{1}{2}\xi_3+\frac{1}{2}\xi_4;$$

г) $\eta_1^2+\eta_2^2+\eta_3^2-3\eta_4^2$;
$$\eta_1=\frac{\sqrt{2}}{2}\xi_1+\frac{\sqrt{2}}{2}\xi_2,$$
$$\eta_2=\frac{\sqrt{2}}{2}\xi_2+\frac{\sqrt{2}}{2}\xi_4,$$
$$\eta_3=\frac{1}{2}\xi_1-\frac{1}{2}\xi_2+\frac{1}{2}\xi_3-\frac{1}{2}\xi_4,$$
$$\eta_4=\frac{1}{2}\xi_1-\frac{1}{2}\xi_2-\frac{1}{2}\xi_3+\frac{1}{2}\xi_4.$$

2. *Ответ.* Максимум при $x=(\pm 1, 0, 0)$, где $A(x, x)=1$.

Минимум при $x=(0, 0, \pm 1)$, где $A(x, x)=\frac{1}{3}$.

При $x=(0, \pm 1, 0)$, где $A(x, x)=\frac{1}{2}$,— минимакс (т. е. при движении по единичной сфере в одну сторону от точки x функция $A(x, x)$ возрастает, а в другую — убывает).

3. *Указание*. А именно, на подпространствах, натянутых на соответствующие канонические базисные векторы.

4. *Указание*. Коэффициент λ_k равен наименьшему из максимумов формы $A(x, x)$ на некоторой системе подпространств, а коэффициент μ_k равен наименьшему из максимумов формы $B(x, x)$ на той же системе подпространств.

5. *Ответ*. $\dfrac{y}{x} = \pm\dfrac{1}{2}$.

6. *Ответ*. $A(x, x) = \eta_1^2 + \eta_2^2 + \eta_3^2$, $B(x, x) = \eta_1^2 + 2\eta_2^2 + 3\eta_3^2$,

$\xi_1 = \eta_1 - \eta_2 + 2\eta_3$, $\xi_2 = \eta_2 - \eta_3$, $\xi_3 = \eta_3$.

7. *Указание*. Вопрос сводится к единственности канонического базиса у симметричного оператора с попарно различными собственными значениями.

8. *Указание*. Обобщить *7.44*.

9. *Ответ*. а) Однополостный гиперболоид с осевой линией вдоль оси y; б) однополостный гиперболоид с осевой линией вдоль оси x; в) круговой параболоид с осевой линией вдоль оси x; г) круговой параболоид с осевой линией вдоль оси y, сдвинутый на 1 вдоль этой оси; д) гиперболический параболоид.

10. *Ответ*.

а) $x_1^2 + 2y_1^2 + 3z_1^2 = 6$;
$$3(x-1) = -x_1 + 2y_1 + 2z_1,\quad 3y = 2x_1 - y_1 + 2z_1,\quad 3(z+1) = 2x_1 + 2y_1 - z_1;$$

б) $x_1^2 + 2y_1^2 - 3z_1^2 = 6$;
$$3(x+1) = -x_1 + 2y_1 + 2z_1,\quad 3(y+1) = 2x_1 - y_1 + 2z_1,\quad 3z = 2x_1 + 2y_1 - z_1;$$

в) $y_1^2 = 2x_1$;
$$3(x-m) = 2x_1 + 2y_1 + z_1,\quad 3(y+2m) = 2x_1 - y_1 - 2z_1,\quad 3(z+2m) = -x_1 + 2y_1 - 2z_1$$
(m произвольно).

11. *Указание*. Полуоси эллипсоида определяются по каноническим коэффициентам соответствующей квадратичной формы. Использовать результаты *10.25*.

К главе 11

1. *Указание*. Пусть K' — пересечение нуль-многообразий всех операторов, входящих в левый идеал $J \subset L(K_n)$, и пусть r — размерность K'. Выберем базис в K_n так, чтобы первые r базисных векторов лежали в K'. Матрицы всех операторов $A \in J$ имеют r первых столбцов из одних нулей. Пусть размерность J есть m и $A_1, \ldots, A_m$ — линейно независимые операторы в J. Рассмотрим матрицу из $n-r$ столбцов и mn строк, полученную при записывании всех матриц

$A_1 \ldots A_m$ одна над другой и отбрасывании r первых (нулевых) столбцов. Ее ранг равен $n-r$; следовательно, имеются $n-r$ базисных строк. Их линейные комбинации дают всевозможные строки из $n-r$ элементов. Использовать *4.44*.

2. *Указание*. Введя невырожденную билинейную форму (x, y), рассмотреть совокупность J^* всех операторов A^*, сопряженных к операторам $\mathrm{A} \in \mathrm{J}$. Эта совокупность есть левый идеал. Использовать задачу 1.

3. *Ответ*. Максимальный левый идеал алгебры $\mathrm{B}(\mathrm{K}_n)$ — совокупность всех операторов, переводящих в нуль некоторый фиксированный вектор пространства K_n. Минимальный левый идеал — совокупность всех операторов, переводящих в 0 некоторое фиксированное $(n-1)$-мерное подпространство пространства K_n. Максимальный правый идеал — совокупность всех операторов, переводящих все пространство K_n в фиксированное $(n-1)$-мерное подпространство. Минимальный правый идеал — совокупность всех операторов, переводящих все пространство K_n в фиксированную прямую.

4. *Указание*. В базисе $e_1, \ldots, e_n$, в котором матрицы операторов $\mathrm{A} \in \mathrm{B}$ записываются в форме *11.85*, положить

$$(x, y) = \sum \xi_j \overline{\eta}_j \quad \left(x = \sum \xi_j e_j, \; y = \sum \eta_j e_j\right).$$

5. *Указание*. Вместе со всяким инвариантным (относительно алгебры B) подпространством $\mathrm{C}' \in \mathrm{C}_n$ является инвариантным и его ортогональное дополнение. Разложить C_n в ортогональную прямую сумму неприводимых инвариантных подпространств. Каждый оператор $\mathrm{A} \neq 0$ (из алгебры B) хотя бы в одном из них действует как ненулевой оператор.

6. *Указание*. Получить из представления *11.85*, что коммутатор полупростой, но не простой матричной алгебры B пересекается с самой B *не только* по матрицам, кратным единичной.

7. *Указание*. Записать искомую матрицу в виде блочной матрицы из m^2 блоков, выписать условие коммутируемости и использовать лемму Шура.

8. *Ответ*. У алгебры B всех диагональных матриц

$$A = \begin{Vmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{Vmatrix}$$

($\lambda_1, \ldots, \lambda_n$ — произвольные комплексные числа). Всякая матричная алгебра $\mathrm{B} = \overline{\mathrm{B}}$ приводится к этому виду в некотором базисе.

9. *Ответ*. Алгебра B всех операторов, оставляющих собственными данную систему подпространств, дающих в прямой сумме все C_n, удовлетворяет условию $\mathrm{B} \subset \overline{\mathrm{B}}$. Всякая алгебра B с $\mathrm{B} \subset \overline{\mathrm{B}}$ приводится к этому виду.

10. *Ответ*. Пространство C_n разложено в прямую сумму подпространств $\mathrm{C}^{(1)}, \ldots, \mathrm{C}^{(k)}$, и алгебра B состоит из всех операторов, инвариантных в каждом $C^{(j)}$, $j=1, \ldots, k$. Коммутатор $\overline{B}$ состоит из операторов, кратных единичному в каждом $C^{(j)}$ ($j=1, \ldots, k$).

11. *Указание.* Если B есть прямая сумма $B^{(1)}\dotplus\ldots\dotplus B^{(k)}$, то $\overline{\overline{B}}=\overline{\overline{B}}^{(1)}\dotplus\ldots\dotplus\overline{\overline{B}}^{(k)}$.

12. *Ответ.* Кратность каждого характеристического корня оператора A равна размеру соответствующей жордановой клетки (см. задачу 15 к гл. 6).

13. *Указание.* Если $CB=B$, то для некоторого $A\in B$ имеем $CA=C$. Отсюда $C=CA=C(CA)=C^2A=C^3A=\ldots$

15. *Указание.* Пусть $A_1, \ldots, A_m$ — базис алгебры B. Тогда, если алгебра B не нильпотентна, один из правых идеалов $A_1B, \ldots, A_mB$, например A_1B, не нильпотентен (задача 14). При этом $A_1B\neq B$ (задача 13) и проблема сведена к аналогичной для алгебры меньшей размерности.

16. *Указание.* Если $M_i=M_{i+1}$, то для любого вектора $x\notin M_i$ найдется оператор $A_1\in B$ такой, что $A_1x\notin M_i=M_{i+1}$; далее, найдется $A_2\in B$ такой, что $A_2A_1x\in M_i$ и т. д. Если $M_p\neq K_n$, то для $x\in M_{p+1}-M_p$ найдется оператор $A_p\in B$ такой, что $A_px\in M_p-M_{p-1}$, затем $A_{p-1}\in B$ такой, что $A_{p-1}A_px\in M_{p-1}-M_{p-2}$, и т. д., так что $A_1A_2\ldots A_px\neq 0$.

17. *Указание.* Использовать подпространства $M_1, \ldots, M_p$ задачи 16.

с линейной оболочкой $L(1, t, t^2, \ldots, t^n)$ и, следовательно, совпадает с совокупностью L_n, что и требовалось.

б. *Вектор $p_n(t)$ ортогонален к подпространству L_{n-1}.* Достаточно проверить, что многочлен $p_n(t)$ ортогонален в $R_2(-1, 1)$ к функциям $1, t, \ldots, t^{n-1}$.

При доказательстве мы будем использовать формулу интегрирования по частям в определенном интеграле, предполагаемую известной из анализа. Производные, входящие в эту формулу, для многочленов суть те же самые производные, которые мы рассматривали в *6.73а* с чисто алгебраической точки зрения. В частности, многочлен

$$[(t^2-1)^n] = (t-1)^n (t+1)^n$$

в силу *6.73в* при $t = \pm 1$ имеет равные 0 производные порядков $0, 1, \ldots, n-1$.

Итак, будем вычислять скалярное произведение t^k и $p_n(t)$. Интегрируя по частям, мы получаем

$$(t^k, p_n(t)) = \int_{-1}^{+1} t^k [(t^2-1)^n]^{(n)}\, dt =$$

$$= t^k [(t^2-1)^n]^{(n-1)} \Big|_{-1}^{+1} - k \int_{-1}^{+1} t^{k-1} [(t^2-1)^n]^{(n-1)}\, dt.$$

Внеинтегральный член полученного выражения в силу сказанного выше обращается в нуль. Оставшийся интеграл снова берем по частям, и продолжаем этот процесс, пока показатель при t не снизится до нуля:

$$(t^k, p_n(t)) = -kt^{k-1} [(t^2-1)^n]^{(n-2)} \Big|_{-1}^{+1} +$$

$$+ k(k-1) \int_{-1}^{+1} t^{k-2} [(t^2-1)^n]^{(n-2)}\, dt =$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$= \pm k! \int_{-1}^{+1} [(t^2-1)^n]^{(n-k)}\, dt = \pm k! [(t^2-1)^n]^{(n-k-1)} \Big|_{-1}^{+1} = 0,$$

что и требуется.