И. М. Гельфанд

Лекции по линейной алгебре

пятое издание, исправленное

Добросвет
МЦНМО
Москва 1998

ББК 22.143 я 7
Г 27

Гельфанд И. М.

Г 27 Лекции по линейной алгебре.— 5-е изд., исправленное.— М.: Добросвет, Московский центр непрерывного математического образования, 1998.— 320 с.

ISBN 5–7913–0016–6

Читателю предлагается пятое, исправленное издание курса лекций И. М. Гельфанда, читавшихся автором в Московском государственном университете на протяжении ряда лет.

Для студентов-математиков и широкого круга специалистов, использующих методы линейной алгебры.

ISBN 5–7913–0016–6

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ К ПЯТОМУ ИЗДАНИЮ

Настоящее пятое издание отличается от предыдущего четвертого исправлением ряда погрешностей и опечаток.

Автор благодарит В. Ю. Радионова и редактора книги В. В. Ященко за полезные замечания.

Сентябрь 1998 г. *И. Гельфанд*

ПРЕДИСЛОВИЕ К ЧЕТВЕРТОМУ ИЗДАНИЮ

В настоящее четвертое издание добавлен новый параграф «Тензорное произведение» (§ 25), написанный совместно с М. И. Граевым. Добавлены также п. 6 в § 9 и текст, напечатанный мелким шрифтом, в конце п. 2 § 23.

Автор благодарит читателей А. Г. Карновского (г. Каунас) и Ю. Г. Шмелакова (г. Москва) за замечания, позволившие исправить ряд опечаток и погрешностей.

Декабрь 1970 г. *И. Гельфанд*

ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ

Настоящее третье издание отличается от второго рядом переделок и добавлений в различных местах книги.

Наиболее существенное добавление — новое доказательство теоремы о приведении матрицы к жордановой нормальной форме (§ 19).

За помощь в переработке книги я благодарю В. Пономарева и З. Я. Шапиро.

Благодарю также редактора книги Н. Я. Виленкина за ряд ценных советов.

Декабрь 1965 г. *И. Гельфанд*

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Второе издание отличается от первого рядом существенных изменений и дополнений. Наиболее крупными из них являются следующие: включены два добавления, помещенные в конце книги: о вычислительных методах линейной алгебры и о теории возмущений, добавлен параграф, посвященный экстремальным свойствам собственных значений, и параграф о λ-матрицах (§§ 17 и 22), заново написана глава о жордановой нормальной форме линейного преобразования, переработана четвертая глава. Кроме того, сделано много более мелких добавлений и изменений. Новый текст написан мною совместно с З. Я. Шапиро.

Выражаю благодарность А. Г. Курошу, предоставившему в мое распоряжение записи своих лекций по тензорной алгебре. За ряд ценных замечаний благодарю С. В. Фомина.

Благодарю также М. Л. Цетлина за помощь при оформлении рукописи и ряд советов.

Сентябрь 1950 г. *И. Гельфанд*

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

В основу настоящей книги положен курс линейной алгебры, читанный автором на механико-математическом факультете Московского государственного университета и в Белорусском государственном университете.

В написании этой книги принял значительное участие Сергей Васильевич Фомин. Его помощь была настолько существенна, что без нее эта книга вряд ли могла быть написана.

Автор выражает благодарность доценту БГУ А. Е. Турецкому, предоставившему в его распоряжение обработанные записки лекций, читанных автором в 1945 г., а также Д. А. Райкову, внимательно прочитавшему рукопись и сделавшему ряд ценных замечаний.

Некоторые места в тексте напечатаны мелким шрифтом. Эти разделы не используются в основном тексте и при первом поверхностном чтении могут быть пропущены.

Январь 1948 г. *И. Гельфанд*

ГЛАВА I

n-МЕРНОЕ ПРОСТРАНСТВО. ЛИНЕЙНЫЕ И БИЛИНЕЙНЫЕ ФОРМЫ

§ 1. Линейное (аффинное) n-мерное пространство

1. Определение линейного пространства. Часто приходится встречаться с объектами, над которыми производятся операции сложения и умножения на числа. Приведем несколько примеров.

1. В геометрии объектами такого рода являются векторы в трехмерном пространстве, т. е. направленные отрезки. При этом, если два направленных отрезка можно совместить параллельным переносом, то считается, что они определяют один и тот же вектор. Поэтому удобно все эти отрезки откладывать от одной какой-либо точки, которую мы будем называть началом координат. Операция сложения векторов, как известно, определяется следующим образом: суммой векторов x и y мы считаем диагональ параллелограмма со сторонами x и y. Известным образом вводится также умножение на числа.

2. В алгебре мы встречаемся с системами n чисел $x = (\xi_1, \xi_2, \ldots, \xi_n)$ (например, строки матрицы, совокупность коэффициентов линейной формы и т. д.). Для таких систем операции сложения и умножения на числа обычно определяются так: суммой систем $x = (\xi_1, \xi_2, \ldots, \xi_n)$ и $y = (\eta_1, \eta_2, \ldots, \eta_n)$ называется

система $x + y = (\xi_1 + \eta_1, \xi_2 + \eta_2, \ldots, \xi_n + \eta_n)$. Произведением системы $x = (\xi_1, \xi_2, \ldots, \xi_n)$ на число λ мы считаем систему $\lambda x = (\lambda\xi_1, \lambda\xi_2, \ldots, \lambda\xi_n)$.

3. В анализе определяются операции сложения функций и умножения их на числа. В дальнейшем мы для определенности будем рассматривать совокупность всех непрерывных функций, заданных на сегменте $[a, b]$.

В приведенных примерах одни и те же операции сложения и умножения на числа производятся над совершенно разными объектами. Для того чтобы изучить все такие примеры с единой точки зрения, мы введем понятие линейного, или аффинного, пространства.

Определение 1. *Множество R элементов $x, y, z, \ldots$ называется линейным (аффинным) пространством, если:*

a) *каждым двум элементам x и y поставлен в соответствие элемент z, называемый суммой элементов x и y; сумма элементов x и y обозначается через $x + y$,*

b) *каждому элементу x и каждому числу λ из некоторого поля поставлен в соответствие элемент λx, называемый произведением элемента x на число λ.*

Эти операции должны удовлетворять следующим требованиям (аксиомам):

I. 1° $x + y = y + x$ (коммутативность).
2° $(x + y) + z = x + (y + z)$ (ассоциативность).
3° *Существует элемент* 0 *такой, что* $x + 0 = x$ *для любого* x. Элемент 0 называется *нулевым элементом.*

4° *Для каждого x существует элемент, обозначаемый через $-x$, такой, что $x + (-x) = 0$.*

II. 1° $1 \cdot x = x$,
2° $\alpha(\beta x) = \alpha\beta(x)$.

III. 1° $(\alpha + \beta)x = \alpha x + \beta x$,
2° $\alpha(x + y) = \alpha x + \alpha y$.

Мы не случайно не сказали, как именно определяются операции сложения и умножения на числа. От этих операций требуется только, чтобы были выполнены сформулированные выше аксиомы. Поэтому всякий раз, когда мы встречаемся с операциями, удовлетворяющими перечисленным выше условиям, мы вправе считать их операциями сложения и умножения на числа, а совокупность элементов, для которых эти операции установлены, — линейным пространством.

Предоставляем читателю проверить, что в приведенных примерах 1–3 эти аксиомы выполнены. Поэтому 1–3 являются примерами линейных пространств.

Рассмотрим еще несколько примеров.

4. Совокупность всех многочленов степени, не превышающей натурального числа n, с обычными операциями сложения многочленов и умножения их на числа образует линейное пространство.

Заметим, что множество многочленов степени n не образует линейного пространства, так как сумма двух многочленов степени n может оказаться многочленом более низкой степени: например

$$(t^n + t) + (-t^n + t) = 2t.$$

5. Элементами пространства R являются матрицы порядка n. Суммой матриц $\|a_{ik}\|$ и $\|b_{ik}\|$ называется матрица $\|a_{ik} + b_{ik}\|$, произведением матрицы $\|a_{ik}\|$ на число λ — матрица $\|\lambda a_{ik}\|$. Нулевым элементом при этом будет матрица, состоящая из одних нулей. Можно проверить, что все аксиомы линейного пространства здесь выполнены.

6. Совокупность всех многочленов степени, не превышающей натурального числа n, и имеющих положительные коэффициенты, не образует линейного пространства: если многочлен $P(x)$ входит в эту совокупность, то $-P(x)$ в нее не входит.

7. Не образует линейного пространства и совокупность непрерывных функций на сегменте $[a, b]$ таких, что $|f(x)| \leqslant 1$: из того, что $|f_1(x)| \leqslant 1$ и $|f_2(x)| \leqslant 1$, не следует $|f_1(x) + f_2(x)| \leqslant 1$.

Элементы линейного пространства мы будем называть *векторами*. То обстоятельство, что это слово часто употребляется в более узком смысле (так, как в примере 1), не должно нас смущать. Геометрические представления, связанные с этим словом, помогут нам уяснить, а иногда и предвидеть, ряд результатов.

Если числа $\lambda, \mu, \ldots$, участвующие в определении линейного пространства, вещественны, то пространство называется *вещественным линейным пространством*. Если же эти числа $\lambda, \mu, \ldots$ берутся из поля комплексных чисел, то R называется *комплексным линейным пространством*.

Более общо, мы можем предполагать, что $\lambda, \mu, \ldots$ — элементы произвольного поля K. Тогда R называется *линейным пространством над полем* K. Многие понятия и теоремы, излагаемые ниже, в частности, все содержание этого параграфа, автоматически переносятся на линейные пространства над любым полем. Однако в главе I мы будем обычно предполагать, что R — вещественное линейное пространство.

2. Число измерений (размерность) пространства. Важную роль в дальнейшем будет играть понятие линейной зависимости и линейной независимости векторов.

Определение 2. *Пусть R — линейное пространство. Векторы $x, y, z, \ldots, v$ называются линейно зависимыми, если существуют такие числа $\alpha, \beta, \gamma, \ldots, \theta$, из которых хотя бы одно отлично от нуля, что*

$$\alpha x + \beta y + \gamma z + \ldots + \theta v = 0. \tag{1}$$

Векторы, не являющиеся линейно зависимыми, называются линейно независимыми. Другими словами, *векторы $x, y, z, \ldots, v$ называются линейно независимыми, если равенство*

$$\alpha x + \beta y + \gamma z + \ldots + \theta v = 0$$

возможно только при $\alpha = \beta = \gamma = \ldots = \theta = 0$.

Пусть векторы $x, y, z, \ldots, v$ линейно зависимы, т. е. пусть они связаны соотношением вида (1), в котором хотя бы один из коэффициентов, например α, отличен от нуля. Тогда

$$\alpha x = -\beta y - \gamma z - \ldots - \theta v$$

и, разделив на α и положив

$$-\frac{\beta}{\alpha} = \lambda, \quad -\frac{\gamma}{\alpha} = \mu, \quad \ldots, \quad -\frac{\theta}{\alpha} = \zeta,$$

получим:

$$x = \lambda y + \mu z + \ldots + \zeta v. \tag{2}$$

Если вектор x выражается через векторы $y, z, \ldots, v$ в виде (2), то мы будем говорить, что x есть *линейная комбинация* векторов $y, z, \ldots, v$.

Таким образом, *если векторы $x, y, z, \ldots, v$ линейно зависимы, то хотя бы один из них является линейной комбинацией остальных.* Мы предоставляем читателю проверить, что верно и обратное, т. е. что *векторы, один из которых есть линейная комбинация остальных, линейно зависимы.*

Упражнения. 1. Проверить, что если среди векторов $x, y, z, \ldots, v$ имеется нулевой вектор, то эти векторы обязательно линейно зависимы.

2. Показать, что если к линейно зависимым векторам $x, y, z, \ldots$ добавить еще произвольные векторы $u, v, \ldots$, то все эти векторы вместе также будут линейно зависимы.

3. Доказать, что если векторы $y, z, \ldots, v$ линейно независимы и вектор x есть их линейная комбинация

$$x = \alpha y + \beta z + \ldots + \delta v, \tag{3}$$

то представление (3) единственно.

Указание. Предположить, что есть другое представление:

$$x = \alpha_1 y + \beta_1 z + \ldots + \delta_1 v, \tag{4}$$

и вычесть равенство (4) из равенства (3).

Перейдем теперь к определению понятия *числа измерений* (*размерности*) пространства.

В совокупности векторов на прямой всякие два вектора пропорциональны, т. е. линейно зависимы. На плоскости можно найти два линейно независимых вектора, но уже всякие три вектора линейно зависимы. Если R — совокупность векторов трехмерного пространства, то три линейно независимых вектора в R найти можно, но всякие четыре вектора линейно зависимы.

Мы видим, что максимальное число линейно независимых векторов на прямой, плоскости, в трехмерном пространстве совпадает с тем, что в геометрии принято называть числом измерений прямой, плоскости, пространства. Естественно поэтому следующее общее

Определение 3. *Линейное пространство R называется n-мерным, если в нем существует n линейно независимых векторов и нет большего числа линейно независимых векторов.*

Если в пространстве R можно найти любое число линейно независимых векторов, то R называется *бесконечномерным.*

Бесконечномерные пространства составляют предмет специального изучения. Мы будем в этой книге заниматься в основном пространствами конечного числа измерений.

Найдем в каждом из рассмотренных выше примеров 1–5 размерность соответствующего пространства.

1. Как мы уже указали, в пространстве R примера 1 имеется три линейно независимых вектора, а всякие четыре вектора линейно зависимы. Поэтому R трехмерно.

2. R — пространство, векторами которого являются системы n действительных чисел.

В этом пространстве можно указать n линейно независимых векторов, например

$$\begin{aligned} x_1 &= (1, 0, \ldots, 0),\\ x_2 &= (0, 1, \ldots, 0),\\ &\ldots\ldots\ldots\ldots\\ x_n &= (0, 0, \ldots, 1) \end{aligned}$$

(мы предоставляем читателю доказать, что эти векторы действительно линейно независимы).

Упражнение. Показать, что векторы

$$\begin{aligned} x_1 &= (\eta_{11}, \eta_{12}, \ldots, \eta_{1n}),\\ x_2 &= (0, \quad \eta_{22}, \ldots, \eta_{2n}),\\ x_3 &= (0, \quad 0, \quad \ldots, \eta_{3n}),\\ &\ldots\ldots\ldots\ldots\ldots\\ x_n &= (0, \quad 0, \quad \ldots, \eta_{nn}) \end{aligned}$$

в пространстве R также линейно независимы ($\eta_{11}\eta_{22}\ldots\eta_{nn} \neq 0$).

3. R — пространство непрерывных функций. Пусть N — произвольное целое число. Тогда функции $f_1(t) \equiv 1$, $f_2(t) = t$, $\ldots$, $f_N(t) = t^{N-1}$ образуют совокупность N линейно независимых векторов (доказательство предоставляем читателю). Мы видим, что в этом пространстве имеется произвольное число линейно независимых функций, т. е. R бесконечномерно.

4. R — пространство многочленов степени $\leqslant n-1$. В нем n многочленов $1, t, \ldots, t^{n-1}$ линейно независимы.

5. В пространстве квадратных матриц $\|a_{ik}\|$ порядка n все матрицы, у которых на одном каком-либо месте стоит единица, а на остальных местах нули, линейно независимы.

В примерах 1, 2, 4 и 5 мы нашли систему таких линейно независимых векторов $f_1, \ldots, f_n$, что каждый вектор g есть их линейная комбинация. Чтобы установить, что размерность каждого из этих пространств равна числу векторов $f_1, \ldots, f_n$, нам остается доказать, что в этих пространствах нельзя найти другой системы линейно независимых векторов $g_1, \ldots, g_l$ в количестве, превосходящем n. Этот факт можно вывести из следующей полезной леммы, которой мы неоднократно будем пользоваться и в дальнейшем.

Лемма. *Пусть в линейном пространстве задана система из векторов*

$$f_1, \ldots, f_k.$$

Пусть, далее, каждый из векторов

$$g_1, \ldots, g_l$$

есть линейная комбинация векторов $f_1, \ldots, f_k$. Тогда, если векторы $g_1, \ldots, g_l$ линейно независимы, то $l \leqslant k$.

Другими словами, среди линейных комбинаций k векторов $f_1, \ldots, f_k$ не может быть больше чем k линейно независимых.

Доказательство леммы проведем по индукции. При $k = 1$ она очевидна. Предположим, что лемма верна для $k-1$ векторов $f_1, \ldots, f_{k-1}$, и докажем при этом, что она верна для k векторов.

Итак, пусть среди линейных комбинаций векторов

$$f_1, \ldots, f_k$$

есть линейно независимые векторы $g_1, \ldots, g_l$:

$$\begin{aligned} g_1 &= \alpha_{11} f_1 + \ldots + \alpha_{1k} f_k, \\ g_2 &= \alpha_{21} f_1 + \ldots + \alpha_{2k} f_k, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ g_l &= \alpha_{l1} f_1 + \ldots + \alpha_{lk} f_k. \end{aligned} \tag{5}$$

Нам надо показать, что $l \leqslant k$. Если все коэффициенты при f_k равны нулю, лемма доказана, так как в этом случае, по предположению индукции, имеет место равенство $l \leqslant k-1$, а значит, и подавно, $l \leqslant k$. Пусть хотя бы один из коэффициентов при f_k, например α_{lk}, не равен нулю. Чтобы провести индукцию, мы построим $l-1$ новых линейно независимых векторов, которые будут линейными комбинациями векторов $f_1, \ldots, f_{k-1}$. Для этого из последнего равенства выразим f_k:

$$f_k = \frac{1}{\alpha_{lk}} g_l - \frac{\alpha_{l1}}{\alpha_{lk}} f_1 - \ldots - \frac{\alpha_{l,k-1}}{\alpha_{l,k}} f_{k-1}.$$

Это выражение для f_k подставим теперь в каждое из первых $l-1$ равенств (5) и соберем подобные члены. Мы получим равенства следующего вида:

$$\begin{aligned} g_1 - \frac{\alpha_{1k}}{\alpha_{lk}} g_l &= \beta_{11} f_1 + \ldots + \beta_{1,k-1} f_{k-1}, \\ g_2 - \frac{\alpha_{2k}}{\alpha_{lk}} g_l &= \beta_{21} f_1 + \ldots + \beta_{2,k-1} f_{k-1}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ g_{l-1} - \frac{\alpha_{l-1,k}}{\alpha_{lk}} g_l &= \beta_{l-1,1} f_1 + \ldots + \beta_{l-1,k-1} f_{k-1}. \end{aligned} \tag{6}$$

Эти равенства означают, что каждый из $l-1$ векторов

$$g_1' = g_1 - \frac{\alpha_{1k}}{\alpha_{lk}} g_l, \quad \ldots, \quad g_{l-1}' = g_{l-1} - \frac{\alpha_{l-1,k}}{\alpha_{lk}} g_l$$

есть линейная комбинация векторов $f_1, \ldots, f_{k-1}$. Если мы докажем, что они линейно независимы, то, по предположению индукции, отсюда будет следовать, что $l-1 \leqslant k-1$, т. е. $l \leqslant k$.

Таким образом, нам осталось показать, что векторы $g_1', \ldots, g_{l-1}'$ линейно независимы. Но это почти очевидно. Действительно, предположим, что $\lambda_1, \ldots, \lambda_{l-1}$ — такие числа, что

$$\lambda_1 g_1' + \ldots + \lambda_{l-1} g_{l-1}' = 0,$$

т. е. что

$$\lambda_1\left(g_1 - \frac{\alpha_{1k}}{\alpha_{lk}}g_l\right) + \ldots + \lambda_{l-1}\left(g_{l-1} - \frac{\alpha_{l-1,k}}{\alpha_{lk}}g_l\right) = 0.$$

Раскрывая скобки, получаем:

$$\lambda_1 g_1 + \lambda_2 g_2 + \ldots + \lambda_{l-1}g_{l-1} - \\ - \left(\lambda_1\frac{\alpha_{1k}}{\alpha_{lk}} + \ldots + \lambda_{l-1}\frac{\alpha_{l-1,k}}{\alpha_{lk}}\right)g_l = 0.$$

Так как векторы $g_1, \ldots, g_l$ линейно независимы, то все коэффициенты в последнем равенстве равны нулю. В частности, $\lambda_1 = \lambda_2 = \ldots = \lambda_{l-1} = 0$, а это означает, что векторы $g'_1, \ldots, g'_{l-1}$ линейно независимы. Лемма полностью доказана.

Из доказанной леммы вытекает следующий, часто оказывающийся полезным результат: *если в пространстве R существуют k линейно независимых векторов $f_1, \ldots, f_k$ таких, что каждый вектор из R есть их линейная комбинация, то пространство R k-мерно.*

Доказательство этого факта ввиду его простоты мы оставляем читателю.

В каждом из примеров 2, 4 и 5 такая система была выбрана. Таким образом показано, что в примерах 2 и 4 размерность пространства равна n, а в примере 5 размерность пространства равна n^2.

Упражнение. Показать, что если векторы $f_1, \ldots, f_k$, входящие в условие леммы, линейно зависимы, то $l < k$ (а не только $l \leqslant k$).

3. Базис и координаты в n-мерном пространстве.

Определение 4. *Совокупность n линейно независимых векторов $e_1, e_2, \ldots, e_n$ n-мерного пространства R называется базисом в R.*

Например, в пространстве R, рассмотренном в примере 1 (трехмерном пространстве), базис образуют любые три вектора, не лежащие в одной плоскости.

По определению n-мерного пространства в нем существует n линейно независимых векторов, т. е. существует базис.

Покажем, что произвольную систему из k линейно независимых векторов $f_1, \ldots, f_k$, где $k < n$, можно дополнить до базиса в n-мерном пространстве R.

Пусть $e_1, \ldots, e_n$ — какой-либо базис в R. Если бы каждый из векторов $e_1, \ldots, e_n$ был линейной комбинацией векторов f_i, то, согласно лемме, мы имели бы, что $n \leqslant k$, в то время как, по предположению, $k < n$. Значит, среди векторов $e_1, \ldots, e_n$ есть хотя бы один, например e_{p_1}, не являющийся линейной комбинацией векторов $f_1, \ldots, f_k$. Добавив вектор e_{p_1} к $f_1, \ldots, f_k$, мы получим систему из $k+1$ векторов, которые по-прежнему линейно независимы (почему?).

Если $k + 1 < n$, то среди векторов $e_1, \ldots, e_n$ снова есть вектор e_{p_2}, не являющийся линейной комбинацией векторов $f_1, \ldots, f_k, e_{p_1}$. Добавим и этот вектор к системе. Этот процесс можно продолжить до тех пор, пока мы дойдем до n линейно независимых векторов, т. е. до базиса. Этот базис содержит $f_1, \ldots, f_k$, и тем самым наше утверждение доказано.

Т е о р е м а 1. *Каждый вектор x из n-мерного пространства R можно представить, и при том единственным образом, как линейную комбинацию векторов базиса.*

Д о к а з а т е л ь с т в о. Пусть векторы $e_1, e_2, \ldots, e_n$ образуют базис в R. Присоединим к ним произвольный вектор x из R. Векторов $x, e_1, e_2, \ldots, e_n$ уже $n+1$. Поэтому по определению n-мерного пространства они должны быть линейно зависимы, т. е.

$$\alpha_0 x + \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0, \tag{7}$$

где не все α_i равны нулю. Число α_0 заведомо отлично от нуля, так как иначе из формулы (7) следовала бы линейная зависимость векторов $e_1, e_2, \ldots, e_n$.

Выразим из (7) вектор x:

$$x = -\frac{\alpha_1}{\alpha_0}e_1 - \frac{\alpha_2}{\alpha_0}e_2 - \ldots - \frac{\alpha_n}{\alpha_0}e_n.$$

Мы доказали, что каждый вектор $x \in R$ *) есть линейная комбинация векторов $e_1, e_2, \ldots, e_n$.

Докажем теперь единственность полученного разложения. Предположим, что существуют два разложения:

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

и

$$x = \xi'_1 e_1 + \xi'_2 e_2 + \ldots + \xi'_n e_n.$$

Вычитая, получим:

$$0 = (\xi_1 - \xi'_1)e_1 + (\xi_2 - \xi'_2)e_2 + \ldots + (\xi_n - \xi'_n)e_n.$$

Так как $e_1, e_2, \ldots, e_n$ линейно независимы, то это возможно лишь, если

$$\xi_1 - \xi'_1 = \xi_2 - \xi'_2 = \ldots = \xi_n - \xi'_n = 0,$$

т. е.

$$\xi_1 = \xi'_1, \quad \xi_2 = \xi'_2, \quad \ldots, \quad \xi_n = \xi'_n.$$

Единственность разложения доказана.

Определение 5. *Если $e_1, e_2, \ldots, e_n$ есть базис в n-мерном пространстве и*

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n, \tag{8}$$

то числа $\xi_1, \xi_2, \ldots, \xi_n$ называются координатами вектора x в базисе $e_1, e_2, \ldots, e_n$.

Теорема 1 означает, что *при заданном базисе e_1, $e_2, \ldots, e_n$ каждый вектор имеет координаты и притом однозначно определенные.*

---

*) Запись $x \in R$ означает, что x принадлежит R.

Если вектор x имеет в базисе $e_1, e_2, \ldots, e_n$ координаты $\xi_1, \xi_2, \ldots, \xi_n$, а вектор y — координаты $\eta_1, \eta_2, \ldots, \eta_n$, т. е. если

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

то

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

т. е. вектор $x + y$ имеет координаты $\xi_1 + \eta_1, \xi_2 + \eta_2, \ldots$ $\ldots, \xi_n + \eta_n$. Вектор λx имеет координаты $\lambda\xi_1, \lambda\xi_2, \ldots$ $\ldots, \lambda\xi_n$.

Таким образом, *при сложении векторов x и y их координаты складываются. При умножении вектора x на число λ его координаты умножаются на это число.*

Ясно также, что нулевой вектор, и только он, имеет все координаты равными нулю.

П р и м е р ы. 1. Для случая трехмерного пространства наше определение координат вектора совпадает с имеющимся в аналитической геометрии определением координат вектора в некоторой (вообще говоря, не прямоугольной) системе координат.

2. Пусть R — пространство, векторами которого являются системы $x = (\xi_1, \xi_2, \ldots, \xi_n)$ из n чисел. Возьмем базис (см. упражнение на стр. 13)

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

Найдем координаты $\eta_1, \eta_2, \ldots, \eta_n$ вектора $x = (\xi_1, \xi_2, \ldots, \xi_n)$ в этом базисе. По определению

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

т. е.

$$(\xi_1, \xi_2, \ldots, \xi_n) = \eta_1(1, 1, \ldots, 1) + \eta_2(0, 1, \ldots, 1) + \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ + \eta_n(0, 0, \ldots, 1) = (\eta_1, \eta_1 + \eta_2, \ldots, \eta_1 + \eta_2 + \ldots + \eta_n).$$

Таким образом, числа $\eta_1, \eta_2, \ldots, \eta_n$ находятся из следующей системы уравнений:

$$\begin{aligned} &\eta_1 = \xi_1, \\ &\eta_1 + \eta_2 = \xi_2, \\ &\cdots\cdots\cdots \\ &\eta_1 + \eta_2 + \ldots + \eta_n = \xi_n, \end{aligned}$$

откуда

$$\eta_1 = \xi_1, \quad \eta_2 = \xi_2 - \xi_1, \quad \ldots, \quad \eta_n = \xi_n - \xi_{n-1}.$$

Рассмотрим теперь в R базис, в котором связь между координатами вектора $x = (\xi_1, \xi_2, \ldots, \xi_n)$ и числами $\xi_1, \xi_2, \ldots, \xi_n$, определяющими этот вектор, наиболее проста. Пусть

$$\begin{aligned} &e_1 = (1, 0, \ldots, 0), \\ &e_2 = (0, 1, \ldots, 0), \\ &\cdots\cdots\cdots\cdots \\ &e_n = (0, 0, \ldots, 1) \end{aligned}$$

Тогда

$$x = (\xi_1, \xi_2, \ldots, \xi_n) = \\ = \xi_1(1, 0, \ldots, 0) + \xi_2(0, 1, \ldots, 0) + \ldots + \xi_n(0, 0, \ldots, 1) = \\ = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n.$$

Таким образом, *в пространстве R, где каждый вектор определяется как система n чисел $(\xi_1, \xi_2, \ldots, \xi_n)$, эти числа можно трактовать как координаты век-*

тора $x = (\xi_1, \xi_2, \ldots, \xi_n)$ *в базисе* $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0)$, $\ldots$, $e_n = (0, 0, \ldots, 1)$.

Упражнение. Доказать, что в любом базисе

$$\begin{aligned} e_1 &= (a_{11}, a_{12}, \ldots, a_{1n}), \\ e_2 &= (a_{21}, a_{22}, \ldots, a_{2n}), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n &= (a_{n1}, a_{n2}, \ldots, a_{nn}) \end{aligned}$$

координаты $\eta_1, \eta_2, \ldots, \eta_n$ вектора $x = (\xi_1, \xi_2, \ldots, \xi_n)$ суть линейные комбинации чисел $\xi_1, \xi_2, \ldots, \xi_n$.

3. R — пространство, векторами которого являются многочлены степени $\leqslant n-1$. Простейшим базисом является совокупность векторов $e_1 = 1$, $e_2 = t$, $\ldots$, $e_n = t^{n-1}$. Координатами многочлена $P(t) = a_0 t^{n-1} + a_1 t^{n-2} + \ldots + a_{n-1}$ в этом базисе являются, как легко видеть, его коэффициенты $a_0, a_1, \ldots, a_{n-1}$.

Выберем теперь другой базис:

$$e'_1 = 1,\ e'_2 = t - a,\ e'_3 = (t-a)^2,\ \ldots,\ e'_n = (t-a)^{n-1}.$$

Каждый многочлен $P(t)$ может быть по формуле Тейлора представлен в виде:

$$P(t) = P(a) + P'(a)(t-a) + \ldots + \frac{P^{(n-1)}(a)}{(n-1)!}(t-a)^{n-1}.$$

Таким образом, в этом базисе $P(t)$ имеет координаты

$$P(a), P'(a), \ldots, \frac{P^{(n-1)}(a)}{(n-1)!}.$$

4. Изоморфизм n-мерных пространств. В разобранных выше примерах некоторые пространства с точки зрения рассматриваемых здесь свойств не отличаются друг от друга. Таковы, например, обычное трехмерное пространство R примера 1 и пространство R', в котором векторы определяются как тройки действительных чисел. В самом деле, выбрав в R определенную систему координат, мы можем каждому вектору из R поставить в соответствие совокупность трех его координат, т. е. вектор пространства R'. При сложении векторов координаты их складываются, а при умножении на число все координаты вектора умножаются на это число. Поэтому геометрические факты, вытекающие из определения линейного пространства, которые имеют место в R, мы можем параллельно изложить как в R, так и в пространстве R' троек чисел.

Поскольку единственными операциями, которые введены в линейных пространствах, являются операции сложения векторов и умножения вектора на число, то естественно ввести следующее

Определение 6. *Линейные пространства R и R' называются изоморфными, если между векторами $x \in R$ и векторами $x' \in R'$ можно установить взаимно однозначное соответствие* *) $x \leftrightarrow x'$ *так, что если вектору x соответствует вектор x', а вектору y соответствует вектор y', то*

1° *вектору $x + y$ соответствует вектор $x' + y'$,*

2° *вектору λx соответствует вектор $\lambda x'$.*

Из определения изоморфизма следует, что если $x, y, \ldots$ — векторы из R, а $x', y', \ldots$ — соответствующие им векторы из R', то равенство $\lambda x + \mu y + \ldots = 0$ равносильно равенству $\lambda x' + \mu y' + \ldots = 0$. Следовательно, линейно независимым векторам из R соответствуют линейно независимые векторы из R', и обратно.

Возникает вопрос, какие пространства изоморфны между собой и какие нет.

---

*) Соответствие, установленное между элементами двух множеств R и R', называется взаимно однозначным, если:

1° каждому элементу из R соответствует один и только один элемент из R',

2° каждый элемент из R' соответствует при этом одному и только одному элементу из R.

Два пространства различной размерности заведомо не изоморфны друг другу.

В самом деле, пусть R и R' изоморфны. Из сделанного выше замечания следует, что максимальное число линейно независимых векторов в R и R' одно и то же, т. е. размерности пространств R и R' равны. Следовательно, пространства различной размерности не могут быть между собой изоморфны.

Теорема 2. *Все пространства, имеющие одну и ту же размерность n, изоморфны друг другу.*

Доказательство. Пусть R и R' — два n-мерных пространства. Выберем в R базис $e_1, e_2, \ldots, e_n$ и в R' какой-либо базис $e_1', e_2', \ldots, e_n'$. Поставим в соответствие вектору

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n \tag{9}$$

вектор

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

т. е. линейную комбинацию векторов e_i' с теми же коэффициентами, что и в (9).

Это соответствие взаимно однозначно. В самом деле, каждый вектор x может быть однозначно представлен в виде (9). Поэтому числа ξ_i, а значит, и вектор x', определяются по вектору x однозначно. Ввиду равноправности, в нашем построении, пространств R и R', каждому x' отвечает элемент из R и притом только один.

Из установленного закона соответствия сразу следует, что если $x \leftrightarrow x'$ и $y \leftrightarrow y'$, то $x+y \leftrightarrow x'+y'$ и $\lambda x \leftrightarrow \lambda x'$. Изоморфизм пространств R и R', таким образом, доказан.

Итак, единственной существенной характеристикой конечномерного линейного пространства является его размерность.

В § 3 мы еще вернемся к понятию изоморфизма по другому поводу.

5. Подпространства линейного пространства.

Определение 7. *Подпространством R' пространства R называется совокупность элементов из R таких, что они сами образуют линейное пространство относительно уже введенных в R операций сложения и умножения на числа.*

Иначе говоря, совокупность R' элементов $x, y, \ldots$ из R образует линейное подпространство пространства R, если из $x \in R'$, $y \in R'$ следует $x + y \in R'$, $\lambda x \in R'$.

Примеры. 1. Нулевое подпространство, т.е. подпространство, состоящее из единственного элемента — нуля.

2. Все пространство R.

Нулевое подпространство и все пространство называются обычно *несобственными* подпространствами. Приведем несколько более содержательных примеров подпространств.

3. R — трехмерное пространство. Рассмотрим какую-либо плоскость в R, проходящую через начало координат. Совокупность R' всех векторов, лежащих в этой плоскости, есть подпространство.

4. В пространстве, векторами которого являются системы n чисел $x = (\xi_1, \xi_2, \ldots, \xi_n)$, совокупность всех тех векторов $x = (\xi_1, \xi_2, \ldots, \xi_n)$, для которых $\xi_1 = 0$, образует подпространство. Более общо: совокупность векторов $x = (\xi_1, \xi_2, \ldots, \xi_n)$, удовлетворяющих условию

$$a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n = 0,$$

где $a_1, a_2, \ldots, a_n$ — какие-то фиксированные числа, образует подпространство.

5. В пространстве всех непрерывных функций совокупность многочленов степени $\leqslant n$ является подпространством.

Очевидно, что во всяком подпространстве R' какого-либо пространства R содержится нулевой элемент пространства R.

Поскольку любое подпространство само по себе является линейным пространством, то все такие понятия, как базис, число измерений пространства и т. д., которые мы ввели выше, применимы и к подпространствам. Так как в подпространстве не может быть больше линейно независимых векторов, чем во всем пространстве, то *размерность любого подпространства не превосходит размерности всего пространства.*

Упражнения. 1. Доказать, что если подпространство R' пространства R имеет ту же размерность, что и все пространство R, то оно совпадает с R.

2. Доказать, что если R_1 и R_2 — подпространства пространства R, и если $R_1 \subset R_2$ и размерности R_1 и R_2 совпадают, то $R_1 = R_2$.

В каждом пространстве R можно строить подпространства следующим общим приемом: возьмем в R произвольное (конечное или бесконечное) множество векторов $e, f, g, \ldots$; тогда *совокупность R' всех линейных комбинаций выбранных векторов $e, f, g, \ldots$ есть подпространство пространства R.* Действительно, складывая между собой и умножая на числа линейные комбинации векторов $e, f, g, \ldots$, мы снова получим линейные комбинации векторов $e, f, g, \ldots$, т. е. элементы из R'. Полученное таким образом подпространство R' называется *подпространством, порожденным векторами* $e, f, g, \ldots$ Оно является наименьшим линейным подпространством, содержащим данные векторы $e, f, g, \ldots$

Подпространство R', порожденное линейно независимыми векторами $e_1, e_2, \ldots, e_k$, является k-мерным и векторы $e_1, e_2, \ldots, e_k$ образуют в нем базис. Действительно, в R' имеется система k линейно независимых векторов, именно, сами векторы $e_1, e_2, \ldots, e_k$. С другой стороны, если $x_1, x_2, \ldots, x_l$ — произвольные линейно независимые векторы из R', то так как они являются линейными комбинациями векторов $e_1, e_2, \ldots, e_k$, то, согласно лемме п. 2, $l \leqslant k$. Следовательно, R' k-мерно и набор векторов $e_1, e_2, \ldots, e_k$ есть один из возможных базисов в R'.

Упражнение. Показать, что в n-мерном пространстве существуют подпространства всех меньших размерностей.

Если исключить из рассмотрения не представляющее интереса нулевое подпространство, то самыми простыми являются одномерные подпространства. Базис всякого такого подпространства состоит из одного вектора e_1. Таким образом, одномерное подпространство состоит из векторов вида αe_1, где α — произвольное число.

Прибавим к каждому из векторов αe_1 один и тот же вектор x_0. Мы получим совокупность векторов вида $x = x_0 + \alpha e_1$, где α пробегает все числа, а e_1 и x_0 — фиксированные векторы. Эту совокупность векторов естественно, по аналогии с трехмерным пространством, назвать *прямой* в линейном пространстве R.

Аналогично, векторы вида $\alpha e_1 + \beta e_2$, где e_1 и e_2 — фиксированные линейно независимые векторы, а α и β — произвольные числа, образуют двумерное подпространство. Совокупность векторов

$$x = x_0 + \alpha e_1 + \beta e_2,$$

где x_0 — фиксированный вектор, мы называем *плоскостью* (двумерной).

Упражнения. 1. Показать, что в пространстве, где векторами являются системы n действительных чисел $(\xi_1, \xi_2, \ldots, \xi_n)$, совокупность векторов, удовлетворяющих соотношению

$$a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n = 0$$

($a_1, a_2, \ldots, a_n$ — фиксированные числа, не все равные нулю), образует подпространство размерности $n-1$.

2. Показать, что если два подпространства R_1 и R_2 пространства R имеют общим лишь нулевой вектор, то сумма их размерностей не превосходит размерности R.

3. Показать, что размерность подпространства, порожденного векторами $e, f, g, \ldots$, равна максимальному числу линейно независимых векторов среди них.

6. Разложение пространства R в прямую сумму подпространств. Сумма и пересечение подпространств. Пусть заданы два подпространства n-мерного пространства R. Обозначим их R_1 и R_2.

О п р е д е л е н и е 8. *Если каждый вектор x пространства R можно, и притом единственным образом, представить как сумму двух векторов*

$$x = x_1 + x_2,$$

где $x_1 \in R_1$, а $x_2 \in R_2$, то говорят, что пространство R разложено в прямую сумму подпространств R_1 и R_2.

Это обычно записывают так:

$$R = R_1 \oplus R_2.$$

Т е о р е м а 3. *Для того чтобы пространство R разлагалось в прямую сумму подпространств R_1 и R_2, достаточно, чтобы*:

1. *Подпространства R_1 и R_2 имели только один общий вектор $x = 0$ (нулевой вектор).*

2. *Сумма размерностей этих подпространств была равна размерности пространства R.*

Д о к а з а т е л ь с т в о. Выберем некоторый базис $e_1, \ldots, e_k$ в подпространстве R_1 и базис $f_1, \ldots, f_l$ в подпространстве R_2. Поскольку сумма размерностей R_1 и R_2 есть n, то общее число этих векторов $k + l = n$.

Покажем, что векторы

$$e_1, \ldots, e_k, \quad f_1, \ldots, f_l$$

линейно независимы, т. е. образуют базис пространства R. Действительно, пусть

$$\lambda_1 e_1 + \ldots + \lambda_k e_k + \mu_1 f_1 + \ldots + \mu_l f_l = 0;$$

отсюда

$$\lambda_1 e_1 + \ldots + \lambda_k e_k = -\mu_1 f_1 - \ldots - \mu_l f_l.$$

Левая часть этого равенства есть вектор из R_1, а правая из R_2. Так как, по условию, единственный общий вектор R_1 и R_2 есть нулевой вектор, то

$$\begin{aligned} \lambda_1 e_1 + \ldots + \lambda_k e_k = 0, \\ \mu_1 f_1 + \ldots + \mu_l f_l = 0. \end{aligned} \tag{10}$$

Но каждый из наборов $e_1, \ldots, e_k$ и $f_1, \ldots, f_l$ состоит из линейно независимых векторов, так как это базисы в R_1 и R_2. Поэтому из первого равенства (10) следует, что

$$\lambda_1 = \ldots = \lambda_k = 0,$$

а из второго следует, что

$$\mu_1 = \ldots = \mu_l = 0.$$

Следовательно, система $e_1, \ldots, e_k, f_1, \ldots, f_l$ состоит из n линейно независимых векторов, т. е. это есть базис в пространстве R.

Мы доказали, что при выполнении условий теоремы существует базис, первые k векторов которого образуют базис в R_1, а последние l — базис в R_2.

Произвольный вектор x из R можно разложить по векторам этого базиса

$$x = \alpha_1 e_1 + \ldots + \alpha_k e_k + \beta_1 f_1 + \ldots + \beta_l f_l.$$

При этом

$$x_1 = \alpha_1 e_1 + \ldots + \alpha_k e_k \in R_1$$

и

$$x_2 = \beta_1 f_1 + \ldots + \beta_l f_l \in R_2.$$

Таким образом,

$$x = x_1 + x_2,$$

где $x_1 \in R_1$ и $x_2 \in R_2$. Покажем, что это разложение единственно. Предположим, что существуют два разложения:

$$x = x_1 + x_2, \quad \text{где} \quad x_1 \in R_1, \quad x_2 \in R_2,$$

и

$$x = x_1' + x_2', \quad \text{где} \quad x_1' \in R_1, \quad x_2' \in R_2.$$

Вычитая второе равенство из первого, получаем:

$$0 = x_1 - x_1' + x_2 - x_2',$$

откуда

$$x_1 - x_1' = x_2' - x_2.$$

Так как вектор, стоящий в левой части равенства, принадлежит R_1, а вектор, стоящий в правой части, принадлежит R_2, то каждый из этих векторов равен нулю, т. е.

$$x_1' = x_1,$$
$$x_2' = x_2.$$

Единственность разложения доказана.

Допустим, что нам задано два произвольных подпространства R_1 и R_2 линейного пространства R.

Легко проверить, что совокупность векторов, принадлежащих обоим этим подпространствам, также есть подпространство R_0 пространства R.

Это подпространство называется *пересечением* R_1 и R_2 и обозначается

$$R_0 = R_1 \cap R_2.$$

Например, если R_1 и R_2 — два двумерных подпространства трехмерного пространства (две плоскости, проходящие через начало координат), то $R_1 \cap R_2$ есть одномерное подпространство (прямая, по которой пересекаются эти плоскости).

По двум подпространствам R_1 и R_2 можно построить еще одно подпространство, которое называется их *суммой.* Оно определяется следующим образом.

Векторами этого подпространства являются всевозможные суммы вида

$$x = x_1 + x_2, \tag{11}$$

где $x_1 \in R_1$, $x_2 \in R_2$.

Легко проверить, что элементы вида (11) образуют подпространство. Это подпространство $\tilde{R}$ называется суммой подпространств R_1 и R_2 и обозначается

$$\tilde{R} = R_1 + R_2.$$

Заметим, что, в отличие от прямой суммы двух подпространств, запись элемента из R в виде (11) может быть неоднозначной.

У п р а ж н е н и е. Показать, что сумма двух различных двумерных подпространств трехмерного пространства R есть все это пространство.

Имеет место следующая теорема:

Т е о р е м а 4. *Пусть заданы два подпространства R_1 и R_2 пространства R. Тогда сумма размерностей R_1 и R_2 равна размерности их суммы плюс размерность пересечения.*

Д о к а з а т е л ь с т в о. Выберем в пересечении $R_0 = R_1 \cap R_2$ базис

$$e_1, \ldots, e_k. \tag{12}$$

Дополним этот базис с одной стороны до базиса в R_1:

$$e_1, \ldots, e_k, f_1, \ldots, f_l \tag{13}$$

и с другой стороны до базиса в R_2:

$$e_1, \ldots, e_k, g_1, \ldots, g_m. \tag{14}$$

Покажем, что векторы

$$f_1, \ldots, f_l, e_1, \ldots, e_k, g_1, \ldots, g_m \tag{15}$$

образуют базис в сумме $\tilde{R} = R_1 + R_2$.

Сначала покажем, что эти векторы линейно независимы. Действительно, пусть

$$\lambda_1 f_1 + \ldots + \lambda_l f_l + \mu_1 e_1 + \ldots + \mu_k e_k + \nu_1 g_1 + \ldots + \nu_m g_m = 0.$$

Тогда

$$\lambda_1 f_1 + \ldots + \lambda_l f_l + \mu_1 e_1 + \ldots + \mu_k e_k = -\nu_1 g_1 - \ldots - \nu_m g_m.$$

Левая часть этого равенства есть вектор из R_1, правая — из R_2. Таким образом, эта правая часть есть одновременно вектор из R_1 и из R_2, т. е. принадлежит R_0 и, значит, выражается как линейная комбинация базиса $e_1, \ldots, e_k$ подпространства R_0:

$$-\nu_1 g_1 - \ldots - \nu_m g_m = c_1 e_1 + \ldots + c_k e_k.$$

В силу линейной независимости векторов (14) это возможно только, когда все коэффициенты — нули. В частности, $\nu_1 = \ldots = \nu_m = 0$, т. е.

$$\lambda_1 f_1 + \ldots + \lambda_l f_l + \mu_1 e_1 + \ldots + \mu_k e_k = 0.$$

Из линейной независимости векторов (13) получаем, что и все коэффициенты $\lambda_1, \ldots, \lambda_l, \mu_1, \ldots, \mu_k$ равны нулю. Таким образом, линейная независимость системы (15) доказана.

Покажем теперь, что всякий вектор $x \in \tilde{R}$ выражается как линейная комбинация векторов этой системы. По определению $\tilde{R}$ вектор x можно представить в виде:

$$x = x_1 + x_2,$$

где $x_1 \in R_1$, $x_2 \in R_2$. Так как $x_1 \in R_1$, то его можно представить как линейную комбинацию векторов (13). Аналогично $x_2 \in R_2$ и x_2 можно представить как линейную комбинацию векторов (14). Складывая, получим, что вектор x представим как линейная комбинация системы (15).

Итак мы получили, что векторы

$$f_1, \ldots, f_l,\ e_1, \ldots, e_k,\ g_1, \ldots, g_m,$$

с одной стороны, линейно независимы и, с другой стороны, всякий вектор из $\tilde{R}$ есть их линейная комбинация. В силу замечания на стр. 16 отсюда следует, что эти векторы образуют базис в $\tilde{R}$. Итак, мы имеем k векторов (12), образующих базис в R_0, $k+l$ векторов (13), образующих базис в R_1, $k+m$ векторов (14), образующих базис в R_2, и $k+l+m$ векторов (15), образующих базис в $\tilde{R} = R_1 + R_2$. Утверждение теоремы обращается, таким образом, в тождество

$$(k+l) + (k+m) = (k+l+m) + k.$$

Теорема доказана.

Упражнение. Проверить теорему для случая, когда R_1 и R_2 — двумерные подпространства трехмерного пространства.

Из доказанной теоремы следует, например, что двум 15-мерным подпространствам «тесно» в 28-мерном пространстве — они пересекаются по крайней мере по двумерному подпространству (плоскости). Действительно, сумма их размерностей равна 30, а размерность суммы не может, конечно, превосходить размерности всего пространства, т. е. 28.

Упражнения. 1. Каково наименьшее число измерений пространства, в котором две плоскости могут пересечься в точке?

2. Доказать, что если $R_1 \cap R_2$ есть нулевое подпространство, то $\tilde{R} = R_1 + R_2$ есть прямая сумма R_1 и R_2, т. е. $\tilde{R} = R_1 \oplus R_2$.

Из результата этого упражнения видно, что теорема 3 этого пункта есть частный случай теоремы 4.

Упражнение. Показать, что если имеется разложение R в прямую сумму
$$R = R_1 \oplus R_2,$$
то пересечение R_1 и R_2 равно нулю и, следовательно, сумма размерностей этих подпространств равна n.

7. Преобразование координат при изменении базиса. Пусть $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$ — два базиса n-мерного пространства. Пусть, далее, каждый вектор e'_i выражается через векторы первого базиса формулами

$$\left.\begin{aligned} e'_1 &= a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n, \\ e'_2 &= a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e'_n &= a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n. \end{aligned}\right\} \qquad (16)$$

Тогда переход от первого базиса ко второму задается матрицей $\boldsymbol{A} = \|a_{ik}\|$, определитель которой отличен от нуля *).

Обозначим через ξ_i координаты вектора x в первом базисе, а через ξ'_i — его координаты во втором базисе. Найдем, как выражаются координаты ξ'_i через ξ_i.

Мы имеем:
$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n = \xi'_1 e'_1 + \xi'_2 e'_2 + \ldots + \xi'_n e'_n.$$
Подставив в это равенство вместо e'_i их выражения через e_i, получим:
$$\begin{aligned} x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n = \xi'_1(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n) &+ \\ + \xi'_2(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n) &+ \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ + \xi'_n(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n)&. \end{aligned}$$
Так как e_i линейно независимы, то коэффициенты при них в правой и левой частях равенства одинаковы.

*) Если бы определитель матрицы $\boldsymbol{A}$ был равен нулю, то векторы $e'_1, e'_2, \ldots, e'_n$ были бы линейно зависимы.

Получаем:

$$\left.\begin{aligned} \xi_1 &= a_{11}\xi_1' + a_{12}\xi_2' + \ldots + a_{1n}\xi_n', \\ \xi_2 &= a_{21}\xi_1' + a_{22}\xi_2' + \ldots + a_{2n}\xi_n', \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \xi_n &= a_{n1}\xi_1' + a_{n2}\xi_2' + \ldots + a_{nn}\xi_n'. \end{aligned}\right\} \tag{17}$$

Сравним формулы (16) и (17). Между ними есть два существенных отличия: во-первых, поменялись местами штрихованные и нештрихованные буквы и, во-вторых, в формулах (16) при суммировании меняется первый индекс, а в формулах (17) второй.

Таким образом, *координаты ξ_i вектора x в первом базисе выражаются через координаты того же вектора x во втором базисе с помощью матрицы $\boldsymbol{A}'$, транспонированной к $\boldsymbol{A}$.*

Этот результат можно представить и в другой форме. Решим уравнения (17) относительно $\xi_1', \xi_2', \ldots, \xi_n'$. Получим:

$$\begin{aligned} \xi_1' &= b_{11}\xi_1 + b_{12}\xi_2 + \ldots + b_{1n}\xi_n, \\ \xi_2' &= b_{21}\xi_1 + b_{22}\xi_2 + \ldots + b_{2n}\xi_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \xi_n' &= b_{n1}\xi_1 + b_{n2}\xi_2 + \ldots + b_{nn}\xi_n, \end{aligned}$$

где b_{ik} являются элементами матрицы, обратной к матрице $\boldsymbol{A}'$. Таким образом, мы видим, что *координаты вектора преобразуются с помощью матрицы $\boldsymbol{B} = \|b_{ik}\|$, являющейся обратной к $\boldsymbol{A}'$, где $\boldsymbol{A}'$ — матрица, транспонированная к матрице $\boldsymbol{A}$, задающей преобразование базиса.*

§ 2. Евклидово пространство

1. Определение евклидова пространства. В предыдущем параграфе линейное (аффинное) пространство было определено как множество элементов (векторов) с заданными в нем операциями умножения на числа и сложения.

С помощью этих операций можно сформулировать, что́ такое прямая, плоскость, число измерений пространства, что́ такое параллельные прямые и т. д.

Однако этих понятий недостаточно, чтобы охватить все многообразие фактов, составляющих содержание так называемой евклидовой геометрии. Например, в одних терминах сложения и умножения на число мы не сможем дать определение длины вектора, угла между векторами, скалярного произведения векторов и т. д. Ввести эти понятия проще всего следующим образом.

Выберем в качестве основного понятие скалярного произведения, которое определим аксиоматически.

В терминах сложения векторов, умножения их на числа и скалярного произведения векторов мы сможем развить всю евклидову геометрию.

Определение 1. *Будем говорить, что в вещественном пространстве R определено скалярное произведение, если каждой паре векторов $x, y \in R$ поставлено в соответствие действительное число, которое обозначим через (x, y) причем это соответствие обладает следующими свойствами* (удовлетворяет следующим аксиомам):

1° $(x, y) = (y, x)$, *т. е. скалярное произведение симметрично.*

2° $(\lambda x, y) = \lambda(x, y)$, *где λ — действительное число.*

3° $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ *(дистрибутивность скалярного произведения).*

4° *Скалярное произведение вектора с самим собой неотрицательно*: $(x, x) \geqslant 0$, *и обращается в нуль, лишь если $x = 0$.*

Аффинное пространство, в котором определено скалярное произведение, удовлетворяющее условиям 1°–4°, мы называем евклидовым.

Примеры. 1. Под векторами пространства R мы будем понимать векторы изучаемого в элементарной геометрии трехмерного пространства (пример 1 § 1). Скалярное произведение векторов определим как произведение их длин на косинус угла между ними. Можно проверить, что аксиомы 1°–4° действительно выполнены. Мы предоставляем эту проверку читателю.

2. Векторами пространства R мы назовем всякую систему n действительных чисел $x = (\xi_1, \xi_2, \ldots, \xi_n)$. Сложение векторов и умножение их на число определим так (пример 2 § 1):

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

где

$$x = (\xi_1, \xi_2, \ldots, \xi_n), \quad y = (\eta_1, \eta_2, \ldots, \eta_n).$$

Скалярное произведение векторов x и y определим формулой

$$(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n.$$

Легко проверить, что аксиомы 1°–3° действительно выполнены. Аксиома 4° также справедлива, так как $(x, x) = \sum \xi_i^2 \geqslant 0$ и $(x, x) = \sum \xi_i^2 = 0$ только при $\xi_1 = \xi_2 = \ldots = \xi_n = 0$.

3. Рассмотрим пример более общий, чем пример 2.

Вектор по-прежнему определим как совокупность n действительных чисел. Сложение векторов и умножение их на числа определим так же, как в примере 2.

Зададимся некоторой матрицей $\|a_{ik}\|$. Скалярное произведение векторов x и y определим формулой

$$\begin{aligned}(x, y) = a_{11}\xi_1\eta_1 + a_{12}\xi_1\eta_2 + \ldots + a_{1n}\xi_1\eta_n &+ \\ + a_{21}\xi_2\eta_1 + a_{22}\xi_2\eta_2 + \ldots + a_{2n}\xi_2\eta_n &+ \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ + a_{n1}\xi_n\eta_1 + a_{n2}\xi_n\eta_2 + \ldots + a_{nn}\xi_n\eta_n&. \end{aligned} \tag{1}$$

Посмотрим, какие условия нужно наложить на матрицу $\|a_{ik}\|$, чтобы выражение, определяемое формулой (1), действительно удовлетворяло всем аксиомам скалярного произведения.

Непосредственной проверкой убеждаемся в том, что аксиомы 2° и 3° выполнены для всякой матрицы $\|a_{ik}\|$. Для того чтобы была выполнена аксиома 1°, т. е. чтобы выражение (x, y) было симметричным относительно x и y, необходимо и достаточно, чтобы

$$a_{ik} = a_{ki}, \tag{2}$$

т. е. чтобы матрица $\|a_{ik}\|$ была симметричной.

Аксиома 4° требует, чтобы выражение

$$(x, x) = \sum_{i,k=1}^{n} a_{ik}\xi_i\xi_k \tag{3}$$

было неотрицательно для любых $\xi_1, \xi_2, \ldots, \xi_n$ и обращалось в нуль, лишь если $\xi_1 = \xi_2 = \ldots = \xi_n = 0$.

Однородный многочлен («квадратичная форма»), определяемый формулой (3), называется *положительно определенным,* если он принимает лишь неотрицательные значения и обращается в нуль, лишь когда все ξ_i равны нулю. Аксиома 4° требует, следовательно, чтобы квадратичная форма (3) была положительно определенной.

Итак, всякая матрица $\|a_{ik}\|$ задает скалярное произведение, определяемое формулой (1), если только эта матрица симметрична [условие (2)] и соответствующая ей квадратичная форма — положительно определенная.

Если в качестве матрицы $\|a_{ik}\|$ взять единичную матрицу, т. е. положить $a_{ii} = 1$ и $a_{ik} = 0$ $(i \neq k)$, то скалярное произведение (x, y) примет вид

$$(x, y) = \sum_{i=1}^{n} \xi_i\eta_i$$

и мы получим евклидово пространство, определенное в примере 2.

Упражнение. Показать, что матрица $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ непригодна для построения скалярного произведения (соответствующая ей квадратичная форма не является положительно определенной), а матрица $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ определяет скалярное произведение, удовлетворяющее аксиомам 1°–4°.

В дальнейшем (§ 6) будут указаны простые условия, дающие возможность проверить, будет ли данная квадратичная форма положительно определенной.

4. Векторами пространства R мы будем называть непрерывные функции, заданные на интервале (a, b); скалярное произведение таких функций определим как интеграл их произведения

$$\int_a^b f(t)g(t)\,dt.$$

Можно проверить, что при таком определении скалярного произведения аксиомы 1°–4° выполнены.

5. Будем считать векторами многочлены от t степени не выше $n-1$. Скалярное произведение двух многочленов определим как и в предыдущем примере:

$$(P, Q) = \int_a^b P(t)Q(t)\,dt.$$

Аксиомы 1°–4° проверяются как и в примере 4.

2. Длина вектора. Угол между векторами. Определим с помощью введенного понятия скалярного произведения длину вектора и угол между векторами.

Определение 2. *Длиной вектора x в евклидо-*

вом пространстве называется число

$$\sqrt{(x, x)}. \tag{4}$$

Длину вектора x будем обозначать через $|x|$.

Естественно пожелать, чтобы угол между векторами, длина вектора и скалярное произведение были связаны обычным соотношением: скалярное произведение векторов равно произведению их длин на косинус угла между ними. Так как в этой фразе смысл всех слов, кроме слов «угол между векторами», нам уже известен, то этим предписывается следующее

Определение 3. *Углом между векторами x и y мы назовем число*

$$\varphi = \arccos \frac{(x, y)}{|x||y|},$$

т. е. положим

$$\cos\varphi = \frac{(x, y)}{|x||y|}, \quad 0 \leqslant \varphi \leqslant \pi. \tag{5}$$

Векторы x и y называются *ортогональными*, если угол между ними равен $\frac{\pi}{2}$, т. е. если

$$(x, y) = 0.$$

С помощью введенных понятий можно перенести на евклидовы пространства ряд теорем элементарной геометрии *).

Рассмотрим один пример. Если x и y — ортогональные векторы, то $x + y$ естественно считать диагональю прямоугольника со сторонами x и y. Докажем, что

$$|x + y|^2 = |x|^2 + |y|^2,$$

*) Можно было бы, конечно, определить евклидово пространство иначе, чем в п. 1, введя аксиоматически понятия длины вектора и угла между векторами (а не скалярное произведение). Однако соответствующая система аксиом была бы более сложной.

т. е. *квадрат длины диагонали прямоугольника равен сумме квадратов длин двух его непараллельных сторон* (теорема Пифагора).

Доказательство. По определению квадрата длины вектора

$$|x+y|^2 = (x+y, x+y).$$

В силу дистрибутивности скалярного произведения (аксиома 3°)

$$(x+y, x+y) = (x, x) + (x, y) + (y, x) + (y, y).$$

В силу ортогональности векторов x и y

$$(x, y) = (y, x) = 0.$$

Следовательно,

$$|x+y|^2 = (x, x) + (y, y) = |x|^2 + |y|^2,$$

что и требовалось доказать.

Эту теорему можно обобщить: *если векторы* $x, y, z, \ldots$ *попарно ортогональны, то*

$$|x+y+z+\ldots|^2 = |x|^2 + |y|^2 + |z|^2 + \ldots$$

3. Неравенство Коши–Буняковского. В предыдущем пункте у нас остался пробел. Мы определили угол φ между векторами x и y формулой

$$\cos\varphi = \frac{(x, y)}{|x||y|}.$$

Для того чтобы можно было определить φ из этого равенства, нужно доказать, что

$$-1 \leqslant \frac{(x, y)}{|x||y|} \leqslant 1$$

или, что то же самое, что

$$\frac{(x, y)^2}{|x|^2|y|^2} \leqslant 1,$$

т. е.

$$(x, y)^2 \leqslant (x, x)(y, y). \tag{6}$$

Это неравенство называется *неравенством Коши–Буняковского.*

Итак, *для того чтобы иметь право определить угол между двумя векторами формулой* (5), *мы должны доказать неравенство Коши–Буняковского* *).

Чтобы доказать его, рассмотрим вектор $x - ty$, где t — произвольное действительное число. Согласно аксиоме 4° скалярного произведения

$$(x - ty, x - ty) \geqslant 0,$$

т. е. для любого t

$$t^2(y, y) - 2t(x, y) + (x, x) \geqslant 0.$$

Мы видим, что стоящий слева квадратный относительно t трехчлен принимает лишь неотрицательные значения. Следовательно, дискриминант уравнения

$$t^2(y, y) - 2t(x, y) + (x, x) = 0$$

не может быть положительным, т. е.

$$(x, y)^2 - (x, x)(y, y) \leqslant 0,$$

что и требовалось доказать.

Упражнение. Доказать, что знак равенства в (6) имеет место тогда и только тогда, когда векторы x и y линейно зависимы.

Примеры. Мы доказали неравенство (6) для аксиоматически заданного евклидова пространства. Разберем, как выглядит

---

*) В примере 1 п. 1 этого параграфа нет надобности доказывать это неравенство. Действительно, там в силу принятого в векторном исчислении определения скалярного произведения величина $\frac{(x, y)}{|x||y|}$ есть косинус некоторого уже заранее определенного угла между векторами и поэтому она по абсолютной величине не превосходит 1.

это неравенство в приведенных выше (п. 1) примерах евклидовых пространств.

1. В примере 1 неравенство (6) не означает ничего нового (см. сноску на стр. 41).

2. Так как в примере 2 скалярное произведение задается формулой

$$(x, y) = \sum_{i=1}^{n} \xi_i \eta_i,$$

то

$$(x, x) = \sum_{i=1}^{n} \xi_i^2, \quad (y, y) = \sum_{i=1}^{n} \eta_i^2;$$

поэтому неравенство (6) имеет здесь вид:

$$\left(\sum_{i=1}^{n} \xi_i \eta_i\right)^2 \leqslant \left(\sum_{i=1}^{n} \xi_i^2\right) \left(\sum_{i=1}^{n} \eta_i^2\right).$$

3. В примере 3 скалярное произведение имеет вид:

$$(x, y) = \sum_{i,k=1}^{n} a_{ik} \xi_i \eta_k,$$

где

$$a_{ik} = a_{ki} \tag{2}$$

и

$$\sum_{i,k=1}^{n} a_{ik} \xi_i \xi_k \geqslant 0 \tag{3}$$

для любых ξ_i. Поэтому неравенство (6) означает:

Если числа a_{ik} удовлетворяют условиям (2) *и* (3), *то имеем место неравенство*

$$\left(\sum_{i,k=1}^{n} a_{ik} \xi_i \eta_k\right)^2 \leqslant \left(\sum_{i,k=1}^{n} a_{ik} \xi_i \xi_k\right) \left(\sum_{i,k=1}^{n} a_{ik} \eta_i \eta_k\right).$$

Упражнение. Показать, что если числа a_{ik} удовлетворяют условиям (2) и (3), то $a_{ik}^2 \leqslant a_{ii} a_{kk}$. (Указание. Выбрать в только что выведенном неравенстве специальным образом числа $\xi_1, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$.)

4. В примере 4 скалярное произведение задается интегралом $\int\limits_a^b f(t) g(t)\, dt$, поэтому неравенство (6) имеет вид:

$$\left(\int_a^b f(t)g(t)\,dt\right)^2 \leqslant \int_a^b f^2(t)\,dt \cdot \int_a^b g^2(t)\,dt.$$

Это неравенство играет важную роль в разных вопросах анализа.

Приведем пример неравенства, являющегося следствием неравенства Коши–Буняковского.

Для любых векторов x и y в евклидовом пространстве R имеет место неравенство

$$|x + y| \leqslant |x| + |y|. \tag{7}$$

Доказательство.

$$|x + y|^2 = (x + y, x + y) = (x, x) + 2(x, y) + (y, y);$$

так как (в силу неравенства Коши–Буняковского) $2(x, y) \leqslant 2|x||y|$, то

$$|x+y|^2 = (x+y, x+y) \leqslant (x, x)+2|x||y|+(y, y) = (|x|+|y|)^2,$$

т. е. $|x + y| \leqslant |x| + |y|$, что и требовалось доказать. (См. также § 3, стр. 63.)

Упражнение. Написать неравенство (7) в каждом из примеров евклидовых пространств, разобранных в начале этого параграфа.

В геометрии расстояние между двумя точками x и y *) определяется как длина вектора $x - y$. В общем случае n-мерного евклидова пространства *определим расстояние между x и y формулой*

$$d = |x - y|.$$

---

*) Мы будем обозначать одной и той же буквой вектор и точку, являющуюся его концом (векторы мы проводим из начала координат, см. стр. 7).

§ 3. Ортогональный базис. Изоморфизм евклидовых пространств

1. Ортогональный базис. В § 1 мы ввели понятие базиса (системы координат) аффинного пространства. В аффинном пространстве у нас нет оснований предпочитать одни базисы другим — там все базисы равноправны *).

В евклидовом пространстве существуют наиболее удобные базисы, а именно ортогональные базисы. Они играют здесь ту же роль, что и прямоугольные системы координат в аналитической геометрии.

Определение 1. *Будем говорить, что n векторов $e_1, e_2, \ldots, e_n$, ни один из которых не равен нулю, образуют ортогональный базис в n-мерном евклидовом пространстве R, если они попарно ортогональны. Векторы $e_1, e_2, \ldots, e_n$ образуют ортогональный нормированный базис, если они попарно ортогональны и имеют каждый длину 1, т. е. если*

$$(e_i, e_k) = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k. \end{cases} \tag{1}$$

Для того чтобы данное нами определение ортогонального базиса было корректным, необходимо доказать, что входящие в определение векторы $e_1, e_2, \ldots, e_n$

*) Точный смысл этого утверждения таков. Если внимательно посмотреть приведенное в § 1 доказательство изоморфизма аффинных пространств, то легко заметить, что там доказано несколько больше, чем сформулировано, а именно, доказано, что между двумя n-мерными пространствами можно установить изоморфное соответствие так, чтобы заданный базис одного пространства перешел в заданный базис другого пространства. В частности, если в R заданы два базиса $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_n$, то существует изоморфное отображение пространства R на себя, при котором первый базис переходит во второй.

действительно образуют базис, т.е. линейно независимы.

Докажем это, т.е. покажем, что равенство

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0 \tag{2}$$

возможно лишь, если $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$. Умножим обе части равенства (2) скалярно на e_1. Получим:

$$\lambda_1(e_1, e_1) + \lambda_2(e_1, e_2) + \ldots + \lambda_n(e_1, e_n) = 0.$$

Но по определению ортогонального базиса

$$(e_1, e_1) \neq 0, \quad (e_1, e_k) = 0 \quad \text{при} \quad k \neq 1.$$

Следовательно, $\lambda_1 = 0$. Аналогично, умножая (2) скалярно на e_2, получим, что $\lambda_2 = 0$ и т.д. Мы доказали, таким образом, что $e_1, e_2, \ldots, e_n$ линейно независимы.

Чтобы доказать существование ортогональных базисов, воспользуемся так называемым *процессом ортогонализации*, который часто встречается в геометрии. Он состоит в том, что из данных линейно независимых векторов $f_1, \ldots, f_m$ строятся m попарно ортогональных векторов $e_1, \ldots, e_m$.

Опишем этот процесс. Пусть даны m линейно независимых векторов $f_1, \ldots, f_m$.

По этим векторам мы построим процессом ортогонализации m попарно ортогональных векторов.

Положим $e_1 = f_1$. Вектор e_2 будем искать в виде: $e_2 = f_2 + \alpha e_1$. Число α подберем так, чтобы $(e_2, e_1) = 0$, т.е. $(f_2 + \alpha e_1, e_1) = 0$. Отсюда

$$\alpha = -\frac{(f_2, e_1)}{(e_1, e_1)}.$$

Предположим, что попарно ортогональные и отличные от нуля векторы $e_1, e_2, \ldots, e_{k-1}$ уже построены. Вектор e_k ищем в виде:

$$e_k = f_k + \lambda_1 e_1 + \ldots + \lambda_{k-1} e_{k-1}, \tag{3}$$

т. е. вектор e_k мы получаем из вектора f_k «исправлением» его с помощью линейной комбинации уже построенных векторов $e_1, e_2, \ldots, e_{k-1}$.

Коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ находим из условия ортогональности вектора

$$e_k = f_k + \lambda_1 e_1 + \ldots + \lambda_{k-1} e_{k-1}$$

к векторам $e_1, e_2, \ldots, e_{k-1}$:

$$\begin{aligned}
(f_k + \lambda_1 e_1 + \ldots + \lambda_{k-1} e_{k-1},\ e_1) &= 0,\\
(f_k + \lambda_1 e_1 + \ldots + \lambda_{k-1} e_{k-1},\ e_2) &= 0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(f_k + \lambda_1 e_1 + \ldots + \lambda_{k-1} e_{k-1},\ e_{k-1}) &= 0.
\end{aligned}$$

Так как векторы $e_1, e_2, \ldots, e_{k-1}$ попарно ортогональны, то эти равенства записываются в виде:

$$\begin{aligned}
(f_k, e_1) + \lambda_1 (e_1, e_1) &= 0,\\
(f_k, e_2) + \lambda_2 (e_2, e_2) &= 0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\\
(f_k, e_{k-1}) + \lambda_{k-1} (e_{k-1}, e_{k-1}) &= 0.
\end{aligned}$$

Отсюда

$$\begin{gathered}
\lambda_1 = -\frac{(f_k, e_1)}{(e_1, e_1)}, \quad \lambda_2 = -\frac{(f_k, e_2)}{(e_2, e_2)}, \quad \ldots,\\
\lambda_{k-1} = -\frac{(f_k, e_{k-1})}{(e_{k-1}, e_{k-1})}.
\end{gathered} \tag{4}$$

До сих пор не было использовано то, что векторы $f_1, f_2, \ldots, f_m$ линейно независимы. Мы используем это при доказательстве того, что построенный вектор e_k отличен от нуля. Заметим предварительно, что вектор e_k есть линейная комбинация векторов $e_1, e_2, \ldots, e_{k-1}, f_k$. Но вектор e_{k-1} можно заменить линейной комбинацией вектора f_{k-1} и векторов $e_1, e_2, \ldots, e_{k-2}$ и т. д. Окончательно мы получаем, что вектор e_k записывается в виде

$$e_k = a_1 f_1 + a_2 f_2 + \ldots + a_{k-1} f_{k-1} + f_k. \qquad (5)$$

Теперь ясно, что $e_k \neq 0$. Действительно, в противном случае правая часть равенства (5) была бы нулем, что противоречит линейной независимости векторов $f_1, f_2, \ldots, f_k$, так как коэффициент при f_k равен 1. Итак, доказано, что $e_k \neq 0$.

Мы построили по векторам $e_1, e_2, \ldots, e_{k-1}$ и f_k вектор e_k. Таким же образом по $e_1, e_2, \ldots, e_k$ и f_{k+1} мы построим e_{k+1} и т. д.

Продолжая этот процесс до тех пор, пока не будут исчерпаны заданные векторы $f_1, f_2, \ldots, f_m$, получаем m отличных от нуля и попарно ортогональных векторов $e_1, e_2, \ldots, e_m$. Докажем теперь следующую теорему.

Теорема 1. *Во всяком n-мерном пространстве существуют ортогональные базисы.*

Доказательство. По определению n-мерного пространства (§ 1, п. 2) в нем существует какой-то базис $f_1, \ldots, f_n$. С помощью процесса ортогонализации из него можно построить ортогональный базис $e_1, \ldots, e_n$, что и доказывает теорему.

Если заменить векторы e_k векторами

$$e'_k = \frac{e_k}{|e_k|},$$

то это будут, как нетрудно видеть, попарно ортогональные векторы длины 1, т. е. мы получим ортогональный нормированный базис.

Нетрудно видеть, что таких базисов существует много. Действительно, уже из данного базиса $f_1, \ldots, f_n$ можно построить разные ортогональные базисы, если начинать построение с разных векторов f_k. Позднее, в гл. II, мы рассмотрим вопрос о том, как связаны между собой различные ортогональные базисы.

П р и м е р ы о р т о г о н а л и з а ц и и. 1. Пусть R — трехмерное пространство. Процесс ортогонализации в нем означает следующее: пусть даны три линейно независимых вектора f_1, f_2, f_3. Положим $e_1 = f_1$. Проводим затем плоскость через $e_1 = f_1$ и f_2 и в этой плоскости выбираем вектор e_2, ортогональный к e_1. Наконец, во всем пространстве находим вектор, ортогональный к e_1 и к e_2 (т. е. к построенной ранее плоскости).

2. Пусть R — трехмерное пространство, векторами в котором мы считаем многочлены степени не выше второй. Скалярное произведение зададим формулой

$$\int_{-1}^{1} P(t)Q(t)\,dt.$$

Векторы $1, t, t^2$ образуют базис в R. Применим к этому базису процесс ортогонализации: $e_1 = 1$; вектор e_2 ищем в виде: $t + \alpha \cdot 1$; из условия ортогональности

$$0 = (t + \alpha \cdot 1, 1) = \int_{-1}^{1} (t + \alpha)\,dt = 2\alpha$$

получаем $\alpha = 0$. Значит, $e_2 = t$. Вектор e_3 ищем в виде: $t^2 + \beta t + \gamma \cdot 1$. Из условий ортогональности получаем $\beta = 0$, $\gamma = -\frac{1}{3}$, т.е. $e_3 = t^2 - \frac{1}{3}$. Окончательно получаем ортогональный базис $1, t, t^2 - \frac{1}{3}$. Если разделить каждый вектор на его длину, то получим ортогональный нормированный базис.

3. Пусть R — пространство многочленов степени не выше чем $n - 1$. Скалярное произведение определим так же, как и в предыдущем примере.

Возьмем базис $1, t, t^2, \ldots, t^{n-1}$. Процесс ортогонализации приводит нас, как и в примере 2, к последовательности многочленов

$$1,\ t,\ t^2 - \frac{1}{3},\ t^3 - \frac{3}{5}t,\ \ldots$$

Эти многочлены с точностью до множителей совпадают с многочленами

$$\frac{1}{2^k \cdot k!} \frac{d^k (t^2 - 1)^k}{dt^k},$$

которые называются многочленами Лежандра. Многочлены Лежандра образуют ортогональный, но не нормированный базис в R. Умножая каждый из этих многочленов на соответствующий множитель, мы можем построить ортогональный и нормированный базис; его элементы будем обозначать через $P_k(t)$.

Пусть $e_1, e_2, \ldots, e_n$ — ортогональный нормированный базис евклидова пространства R. Найдем, как выражается скалярное произведение двух векторов через их координаты в этом базисе. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x, а $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора y в этом базисе, т. е.

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

Тогда

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

и так как

$$(e_i, e_k) = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k, \end{cases}$$

то

$$(x, y) = \xi_1 \eta_1 + \xi_2 \eta_2 + \ldots + \xi_n \eta_n, \tag{6}$$

т. е. *в нормированном ортогональном базисе скалярное произведение двух векторов равно сумме произведений их соответствующих координат* (ср. пример 2 § 2).

Упражнения. 1. Показать, что в произвольном базисе $f_1, f_2, \ldots, f_n$ скалярное произведение задается формулой

$$(x, y) = \sum_{i,k=1}^{n} a_{ik} \xi_i \eta_k,$$

где $a_{ik} = a_{ki}$, а $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$ — координаты векторов x и соответственно y.

2. Показать, что если в некотором базисе $f_1, f_2, \ldots, f_n$ скалярное произведение задается формулой

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

где $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$ — координаты векторов x и y, то этот базис является ортогональным и нормированным.

Найдем координаты вектора x в нормированном ортогональном базисе $e_1, e_2, \ldots, e_n$.

Пусть

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

Умножив обе части этого равенства скалярно на e_1, получим

$$(x, e_1) = \xi_1(e_1, e_1) + \xi_2(e_2, e_1) + \ldots + \xi_n(e_n, e_1) = \xi_1$$

и, аналогично,

$$\xi_2 = (x, e_2), \quad \ldots, \quad \xi_n = (x, e_n). \tag{7}$$

Итак: *координаты вектора в ортогональном нормированном базисе суть скалярные произведения этого вектора на соответствующие базисные векторы.*

Скалярное произведение вектора x на вектор e длины единица естественно назвать *проекцией* вектора x на вектор e. Доказанное утверждение означает, что, как и в аналитической геометрии, *координаты вектора в ортогональном нормированном базисе суть проекции этого вектора на базисные векторы* (на оси координат).

Примеры. 1. Пусть $P_0(t), P_1(t), \ldots, P_n(t)$ — нормированные многочлены Лежандра нулевой, первой, ..., n-й степени. Пусть, далее, $Q(t)$ — произвольный многочлен степени n. Представим $Q(t)$ в виде линейной комбинации многочленов Лежандра. Совокупность многочленов степени $\leqslant n$ образуют $(n+1)$-мерное линейное пространство, $P_0(t), P_1(t), \ldots, P_n(t)$ образуют ортогональный базис в нем. Поэтому всякий многочлен степени $\leqslant n$ представим в виде

$$Q(t) = c_0 P_0(t) + c_1 P_1(t) + \ldots + c_n P_n(t).$$

Коэффициенты c_i, как это следует из (7), вычисляются по формулам

$$c_i = \int_{-1}^{1} Q(t) P_i(t)\, dt.$$

2. Рассмотрим на интервале $(0, 2\pi)$ систему функций

$$1,\ \cos t,\ \sin t,\ \cos 2t,\ \sin 2t,\ \ldots,\ \cos nt,\ \sin nt. \tag{8}$$

Их линейная комбинация

$$P(t) = \frac{a_0}{2} + a_1 \cos t + b_1 \sin t + a_2 \cos 2t + \ldots + b_n \sin nt$$

называется тригонометрическим многочленом n-го порядка. Совокупность тригонометрических многочленов n-го порядка образует $(2n+1)$-мерное пространство R_1. Определим в R_1 скалярное произведение, как обычно, т. е. положим

$$(P, Q) = \int\limits_0^{2\pi} P(t)Q(t)\,dt.$$

Легко проверить, что система (8) будет ортогональным базисом. Действительно,

$$\int\limits_0^{2\pi} \cos kt\ \cos lt\,dt = 0, \quad \text{если} \quad k \neq l,$$

$$\int\limits_0^{2\pi} \sin kt\ \cos lt\,dt = 0,$$

$$\int\limits_0^{2\pi} \sin kt\ \sin lt\,dt = 0, \quad \text{если} \quad k \neq l.$$

Так как

$$\int\limits_0^{2\pi} \sin^2 kt\,dt = \int\limits_0^{2\pi} \cos^2 kt\,dt = \pi, \quad \text{а} \quad \int\limits_0^{2\pi} 1\,dt = 2\pi,$$

то функции

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}} \cos t,\ \frac{1}{\sqrt{\pi}} \sin t,\ \ldots,\ \frac{1}{\sqrt{\pi}} \cos nt,\ \frac{1}{\sqrt{\pi}} \sin nt \tag{8'}$$

образуют в R_1 ортогональный нормированный базис.

2. Перпендикуляр из точки на подпространство. Кратчайшее расстояние от точки до подпространства *).

Определение 2. *Пусть* R_1 — *подпространство евклидова пространства* R. Мы будем говорить, что вектор $h \in R$ ортогонален подпространству R_1, если он ортогонален любому вектору x из R_1.

Если вектор h ортогонален векторам $e_1, e_2, \ldots, e_m$, то он ортогонален любой их линейной комбинации. Действительно, из равенств

$$(h, e_i) = 0$$

следует, что для любых чисел $\lambda_1, \lambda_2, \ldots, \lambda_m$

$$(h,\ \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_m e_m) = 0.$$

Поэтому, для того чтобы вектор h был ортогонален m-мерному подпространству R_1, достаточно, чтобы он был ортогонален m линейно независимым векторам из R_1 (базису в R_1 **)).

Упражнение. Показать, что совокупность всех векторов $y \in R$, ортогональных к подпространству R_1, также образует подпространство пространства R. Это подпространство называется *ортогональным дополнением* к подпространству R_1 в пространстве R.

Рассмотрим в пространстве R некоторое m-мерное подпространство R_1 ***) и вектор f, не принадлежащий R_1. Поставим задачу: опустить перпендикуляр из точки f на R_1, *т. е. найти вектор* f_0 *из* R_1 *такой, чтобы вектор* $h = f - f_0$ *был ортогонален* R_1. Вектор f_0 называется при этом *ортогональной проекцией*

---

*) Этот пункт можно при первом чтении пропустить.

**) Для случая, когда R_1 есть плоскость, эта теорема в элементарной геометрии называется теоремой о двух перпендикулярах.

***) Размерность самого пространства R для нас несущественна. Оно может быть даже бесконечномерным.

вектора f на подпространство R_1. Несколько позже мы увидим, что эта задача всегда имеет решение, притом единственное. Сейчас мы покажем, что, как и в элементарной геометрии, перпендикуляр есть кратчайшее расстояние от точки до подпространства. Другими словами, покажем, что *если f_1 есть отличный от f_0 вектор из R_1, то*

$$|f - f_1| > |f - f_0|.$$

Действительно, вектор $f_0 - f_1$, как разность двух векторов из R_1, принадлежит R_1 и, следовательно, ортогонален вектору $h = f - f_0$. По теореме Пифагора имеем:

$$|f - f_0|^2 + |f_0 - f_1|^2 = |f - f_0 + f_0 - f_1|^2 = |f - f_1|^2$$

и, значит,

$$|f - f_1| > |f - f_0|.$$

Покажем теперь, как фактически вычислить по f его ортогональную проекцию f_0 на подпространство R_1 (т. е. опустить перпендикуляр из f на R_1). Пусть базис подпространства R_1 состоит из векторов $e_1, e_2, \ldots, e_m$. Будем искать вектор f_0 в виде

$$f_0 = c_1 e_1 + c_2 e_2 + \ldots + c_m e_m, \tag{9}$$

где коэффициенты c_k найдем из условия ортогональности $f - f_0$ к R_1. Для того чтобы эта ортогональность имела место, необходимо и достаточно, чтобы выполнялись m равенств $(f - f_0, e_k) = 0$ $(k = 1, 2, \ldots, m)$, т. е.

$$(f_0, e_k) = (f, e_k). \tag{10}$$

Подставляя сюда вместо f_0 его выражение (9), получаем систему m уравнений

$$c_1(e_1, e_k) + c_2(e_2, e_k) + \ldots + c_m(e_m, e_k) = (f, e_k) \\ (k = 1, 2, \ldots, m) \tag{11}$$

относительно чисел c_i.

Рассмотрим сначала отдельно часто встречающийся случай, когда базис $e_1, e_2, \ldots, e_m$ — ортогональный и нормированный. В этом случае задача решается особенно просто. Действительно, система (11) превращается в таком базисе в систему равенств

$$c_i = (f, e_i), \tag{12}$$

сразу определяющих нужные коэффициенты.

Так как в каждом m-мерном подпространстве можно выбрать ортогональный нормированный базис, то мы доказали, таким образом, что *у каждого вектора f существует, и притом только одна, ортогональная проекция f_0 на подпространство R_1.*

Вернемся теперь к случаю произвольного базиса. В этом случае система (11) также должна иметь единственное решение. Действительно, вектор f_0, по доказанному, существует и притом только один. В базисе $e_1, e_2, \ldots, e_m$ вектор f_0 имеет вполне определенные координаты $c_1, c_2, \ldots, c_m$. Так как эти числа удовлетворяют системе (11), то эта система имеет, следовательно, единственное решение. Система m уравнений с неизвестными может иметь единственное решение, лишь если ее определитель отличен от нуля. Отсюда следует, что определитель системы (11)

$$\begin{vmatrix} (e_1, e_1) & (e_2, e_1) & \ldots & (e_m, e_1) \\ (e_1, e_2) & (e_2, e_2) & \ldots & (e_m, e_2) \\ \ldots & \ldots & \ldots & \ldots \\ (e_1, e_m) & (e_2, e_m) & \ldots & (e_m, e_m) \end{vmatrix}$$

отличен от нуля. Этот определитель называется *определителем Грама векторов $e_1, e_2, \ldots, e_m$.*

Итак, пусть задано подпространство R_1 с базисом $e_1, \ldots, e_m$ и произвольный вектор f пространства R.

Ортогональная проекция f_0 вектора f на R_1 имеет вид

$$f_0 = c_1 e_1 + \ldots + c_m e_m.$$

При этом, если базис $e_1, e_2, \ldots, e_m$ ортогонален, то

$$c_i = (f, e_i).$$

Если же базис $e_1, \ldots, e_m$ произволен, то коэффициенты c_i определяются как решение системы (11).

Пример 1. Способ наименьших квадратов. Предположим, что величина y есть линейная функция величин $x_1, \ldots, x_m$, т.е. что

$$y_0 = c_1 x_1 + \ldots + c_m x_m,$$

где $c_1, \ldots, c_m$ — постоянные, неизвестные нам коэффициенты. Часто коэффициенты $c_1, \ldots, c_m$ определяются экспериментально. Для этого производится ряд измерений величин $x_1, \ldots, x_m$ и y. Обозначим результаты k-го измерения через $x_{1k}, \ldots, x_{mk}$ и, соответственно, y_k. Коэффициенты $c_1, \ldots, c_m$ можно было бы попытаться определить из системы уравнений

$$\left.\begin{array}{r} x_{11}c_1 + x_{21}c_2 + \ldots + x_{m1}c_m = y_1, \\ x_{12}c_1 + x_{22}c_2 + \ldots + x_{m2}c_m = y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{1n}c_1 + x_{2n}c_2 + \ldots + x_{mn}c_m = y_n. \end{array}\right\} \qquad (13)$$

Здесь число уравнений n равно числу произведенных измерений и обычно превосходит число неизвестных ($n > m$). Так как измерение величин $x_1, \ldots, x_m, y$ неизбежно связано с погрешностями, то система (13), вообще говоря, противоречива и о ее точном решении говорить бессмысленно. Поэтому уравнениям (13) можно удовлетворить лишь приближенно. Таким образом, ставится задача разыскать такие значения неизвестных $c_1, \ldots, c_m$, при которых левые части уравнения (13) были бы возможно более близки к соответствующим правым частям. В качестве «меры близости» берется так называемое *квадратичное уклонение* левых частей уравнений от свободных членов, т. е. величина

$$\sum_{k=1}^{n}(x_{1k}c_1 + x_{2k}c_2 + \ldots + x_{mk}c_m - y_k)^2. \qquad (14)$$

Нам нужно найти числа $c_1, c_2, \ldots, c_m$, при которых квадратичное уклонение имеет наименьшее значение. Эту задачу на минимум можно решить непосредственно. Однако ее решение можно сразу получить из результатов, изложенных выше.

В самом деле, рассмотрим n-мерное евклидово пространство и в нем векторы $e_1 = (x_{11}, x_{12}, \ldots, x_{1n})$, $e_2 = (x_{21}, \ldots, x_{2n})$, $\ldots$, $e_m = (x_{m1}, \ldots, x_{mn})$ и $f = (y_1, \ldots, y_n)$. Правые части уравнений системы (13) являются компонентами вектора f, левые части — вектора

$$c_1e_1 + c_2e_2 + \ldots + c_me_m.$$

Выражение (14) есть, следовательно, квадрат расстояния вектора $c_1e_1 + c_2e_2 + \ldots + c_me_m$ от вектора f. Таким образом, условие, чтобы квадратичное уклонение было минимальным, равносильно следующей задаче: выбрать числа $c_1, c_2, \ldots, c_m$ так, чтобы расстояние от вектора f до вектора $f_0 = c_1e_1 + c_2e_2 + \ldots + c_me_m$ было наименьшим. Если обозначить через R_1 подпространство n-мерного пространства, состоящее из линейных комбинаций векторов $e_1, e_2, \ldots, e_m$ *), то задача состоит в нахождении проекции вектора f на это подпространство. Как мы видели (формула (11)), числа $c_1, c_2, \ldots, c_m$,

*) Мы предполагаем,что ранг матрицы системы (13) равен m и, следовательно, векторы $e_1, e_2, \ldots, e_m$ линейно независимы.

решающие эту задачу, находятся из системы уравнений

$$\left.\begin{array}{l}(e_1,e_1)c_1+(e_2,e_1)c_2+\ldots+(e_m,e_1)c_m=(f,e_1),\\ (e_1,e_2)c_1+(e_2,e_2)c_2+\ldots+(e_m,e_2)c_m=(f,e_2),\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ (e_1,e_m)c_1+(e_2,e_m)c_2+\ldots+(e_m,e_m)c_m=(f,e_m),\end{array}\right\} \qquad (15)$$

где

$$(f,e_k)=\sum_{j=1}^{n}x_{kj}y_j; \quad (e_i,e_k)=\sum_{j=1}^{n}x_{ij}x_{kj}.$$

Система (15) называется в этом случае *системой нормальных уравнений.* Итак, приближенное решение системы (13) состоит в замене ее нормальной системой (15) m уравнений с m неизвестными. Изложенный метод называется *способом наименьших квадратов.*

Упражнение. Решить по способу наименьших квадратов систему уравнений

$$\begin{aligned}2c&=3,\\ 3c&=4,\\ 4c&=5.\end{aligned}$$

Решение. $e_1=(2,3,4)$, $f=(3,4,5)$. Нормальная система сводится в этом случае к одному уравнению:

$$(e_1,e_1)c=(e_1,f),$$

т. е.

$$29c=40; \qquad c=\frac{40}{29}.$$

Для случая, когда система (13) есть система n уравнений с одним неизвестным

$$\begin{array}{l}x_1c=y_1,\\ x_2c=y_2,\\ \ldots\ldots\\ x_nc=y_n,\end{array} \qquad (13')$$

решение запишется следующим образом:

$$c = \frac{(x,y)}{(x,x)} = \frac{\sum\limits_{k=1}^{n} x_k y_k}{\sum\limits_{k=1}^{n} x_k^2}.$$

Приближенное решение системы $(13')$ может быть истолковано геометрически как проведение через начало координат прямой, проходящей «возможно более близко» от совокупности точек $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Число c представляет тогда угловой коэффициент такой прямой.

Пример 2. Приближение функций тригонометрическими многочленами. Пусть $f(t)$ — некоторая непрерывная функция, заданная на интервале $(0, 2\pi)$. Часто бывает нужно подобрать тригонометрический многочлен $P(t)$ данного порядка, возможно меньше отличающийся от $f(t)$. В качестве меры отклонения $P(t)$ от $f(t)$ мы возьмем квадратичное уклонение, которое задается формулой

$$\int\limits_0^{2\pi} [f(t) - P(t)]^2\, dt. \tag{16}$$

Итак, точная постановка рассматриваемой задачи такова: *среди всех тригонометрических многочленов порядка n*

$$P(t) = \frac{a_0}{2} + a_1 \cos t + b_1 \sin t + \ldots + a_n \cos nt + b_n \sin nt \tag{17}$$

найти тот, квадратичное уклонение которого от заданной функции $f(t)$ минимально.

Введем в рассмотрение пространство R непрерывных функций на отрезке $(0, 2\pi)$. Скалярное произведение в этом пространстве зададим, как обычно, интегралом

$$(f, g) = \int\limits_0^{2\pi} f(t) g(t)\, dt.$$

Длина вектора выражается тогда формулой

$$|f| = \sqrt{\int_0^{2\pi} [f(t)]^2\, dt},$$

и следовательно, квадратичное уклонение (16) есть в нашем пространстве просто квадрат расстояния от $f(t)$ до $P(t)$. Тригонометрические многочлены вида (17) образуют в R подпространство R_1 размерности $2n+1$. Нам нужно найти элемент из R_1, находящийся на минимальном расстоянии от $f(t)$. Эта задача снова решается опусканием перпендикуляра из точки $f(t)$ на подпространство R_1.

Так как функции

$$e_0 = \frac{1}{\sqrt{2\pi}},\ e_1 = \frac{\cos t}{\sqrt{\pi}},\ e_2 = \frac{\sin t}{\sqrt{\pi}},\ \ldots,\ e_{2n-1} = \frac{\cos nt}{\sqrt{\pi}},\ e_{2n} = \frac{\sin nt}{\sqrt{\pi}}$$

образуют ортогональный и нормированный базис в этом подпространстве (см. пример 2 предыдущего пункта), то решением этой задачи служит линейная комбинация базисных векторов

$$P(t) = \sum_{k=0}^{2n} c_k e_k, \tag{18}$$

где

$$c_k = (f, e_k),$$

т. е., вспоминая определение скалярного произведения, имеем:

$$c_0 = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t)\,dt; \quad c_{2k-1} = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(t) \cos kt\,dt;$$

$$c_{2k} = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(t) \sin kt\,dt \qquad (k = 1, \ldots, n).$$

Подставляя c_k и e_k $(k = 1, \ldots, n)$ в формулу (18), мы приходим, таким образом, к следующему результату: *для того чтобы определить тригонометрический многочлен*

$$P(t) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt),$$

квадратичное уклонение которого от заданной функции $f(t)$ минимально, надо определить коэффициенты a_k, b_k по формулам

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(t)\,dt,\ a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt\,dt,\ b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt\,dt.$$

Так определенные числа a_k, b_k называются *коэффициентами Фурье* функции $f(t)$.

3. Изоморфизм евклидовых пространств. Мы рассмотрели ряд примеров n-мерных евклидовых пространств. Эти пространства отличались одно от другого во всяком случае способом задания векторов (так, в примере 2 § 2 вектор есть совокупность n чисел, в примере 5 § 2 — многочлен и т. д.).

Возникает вопрос: какие из этих пространств действительно различны и для каких различие является лишь чисто внешним, т. е. различны лишь способы задания этих пространств?

Для того чтобы вопрос был точно поставлен, нужно определить, какие два евклидовых пространства мы будем считать лишь несущественно различающимися (изоморфными).

Определение 3. *Два евклидовых пространства R и R_1 называются изоморфными, если между их элементами можно установить взаимно однозначное соответствие $x \leftrightarrow x'$ ($x \in R$, $x' \in R'$) так, что*:

1° *Если $x \leftrightarrow x'$ и $y \leftrightarrow y'$, то $x + y \leftrightarrow x' + y'$, т. е. если вектору $x \in R$ соответствует вектор $x' \in R'$, а вектору $y \in R$ соответствует вектор $y' \in R'$, то сумме $x + y$ соответствует сумма $x' + y'$.*

2° *Если $x \leftrightarrow x'$, то $\lambda x \leftrightarrow \lambda x'$.*

3° *Если $x \leftrightarrow x'$ и $y \leftrightarrow y'$, то $(x, y) = (x', y')$, т. е. скалярные произведения соответствующих пар векторов равны между собой.*

Таким образом, евклидовы пространства R_1 и R_2 изоморфны, если они изоморфны как линейные пространства и этот изоморфизм таков, что он сохраняет скалярное произведение соответствующих векторов.

Если в каком-нибудь n-мерном евклидовом пространстве R доказана теорема, сформулированная в терминах сложения, умножения на числа и скалярного произведения векторов, то эта же теорема верна и в любом изоморфном ему пространстве. В самом деле, если как в формулировке, так и в доказательстве такой теоремы заменить векторы из R соответствующими им векторами из R', то в силу свойств 1°, 2°, 3° определения изоморфизма все рассуждения останутся справедливыми, т. е. соответствующая теорема верна и в R'.

Вернемся к вопросу, поставленному ранее. Оказывается, что имеет место следующая

Теорема 2. *Все евклидовы пространства данной размерности изоморфны между собой.*

Доказательство. Докажем, что все n-мерные евклидовы пространства изоморфны специально выбранному «стандартному» n-мерному пространству. Тем самым будет доказано, что все n-мерные евклидовы пространства изоморфны между собой.

В качестве такого стандартного n-мерного пространства R' мы возьмем рассмотренное в § 2 (пример 2) пространство, в котором вектор определяется как совокупность действительных чисел $x' = (\xi_1, \xi_2, \ldots, \xi_n)$, а скалярное произведение векторов $x' = (\xi_1, \xi_2, \ldots, \xi_n)$ и $y' = (\eta_1, \eta_2, \ldots, \eta_n)$ задается формулой

$$(x', y') = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n.$$

Пусть нам дано какое-либо n-мерное евклидово пространство R; выберем в нем нормированный ортогональный базис $e_1, e_2, \ldots, e_n$ (мы доказали ранее, что во всяком евклидовом пространстве такой базис существует). Поставим в соответствие вектору

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

совокупность n чисел $\xi_1, \xi_2, \ldots, \xi_n$, т. е. вектор

$$x' = (\xi_1, \xi_2, \ldots, \xi_n)$$

из R'.

Покажем, что установленное соответствие есть изоморфизм.

Это соответствие взаимно однозначно. Нужно проверить, что выполнены условия 1°–3° определения изоморфизма.

Свойства 1° и 2° очевидны.

Проверим свойство 3°, т. е. равенство скалярных произведений соответствующих друг другу пар векторов. Пользуясь выведенной выше (стр. 49) формулой для скалярного произведения в ортогональном нормированном базисе, имеем:

$$(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n.$$

С другой стороны, по определению скалярного произведения в пространстве R' (пример 2 § 2) имеем:

$$(x', y') = \xi_1\eta_1 + \xi_2\eta_2 + \ldots + \xi_n\eta_n.$$

Таким образом,

$$(x, y) = (x', y'),$$

т. е. равенство скалярных произведений доказано.

Теорема полностью доказана.

Упражнение. Доказать эту теорему методом, аналогичным доказательству в п. 4 § 1.

Из теоремы об изоморфизме можно вывести интересное следствие: любое геометрическое утверждение о двух или трех векторах достаточно проверить в известном из элементарной геометрии трехмерном пространстве *). В самом деле, линейная комбинация этих векторов образует подпространство нашего пространства не более трех измерений. В силу теоремы о изоморфизме это подпространство изоморфно обычному трехмерному пространству (либо его подпространству)

*) Геометрической теоремой мы будем называть утверждение о векторах, которое может быть сформулировано в терминах сложения векторов, умножения их на числа и скалярного произведения.

и, следовательно, наше утверждение достаточно проверить в последнем пространстве.

В частности, справедливость неравенства Коши–Буняковского (являющегося утверждением о паре векторов) следует из того, что оно верно в элементарной геометрии (см. пример 1, стр. 42). Мы получаем, таким образом, новое доказательство неравенства Коши–Буняковского.

Еще один пример. В § 2 мы доказали неравенство (7)

$$|x + y| \leqslant |x| + |y|.$$

В элементарной геометрии это неравенство означает, что длина диагонали параллелограмма не превосходит суммы длин двух смежных сторон, и доказывается в любом учебнике элементарной геометрии. Следовательно, в силу сказанного ранее, это неравенство справедливо в любом евклидовом пространстве. Теорема о изоморфизме дает нам, таким образом, возможность получить, например, неравенство

$$\sqrt{\int_a^b (f(t) + g(t))^2\,dt} \leqslant \sqrt{\int_a^b f^2(t)\,dt} + \sqrt{\int_a^b g^2(t)\,dt},$$

являющееся неравенством (7) § 2 в пространстве функций (пример 4 § 2) как непосредственное следствие только что сформулированной теоремы из элементарной геометрии.

§ 4. Билинейные и квадратичные формы

В этом параграфе мы будем опять заниматься аффинным пространством, а именно, будем изучать простейшие числовые функции от векторов в аффинном пространстве.

1. Линейная функция. Простейшей функцией в аффинном пространстве является линейная функция.

Определение 1. *Говорят, что в аффинном пространстве задана линейная функция (линейная форма), если каждому вектору x поставлено в соответствие число $f(x)$, так что при этом выполнены условия:*

1° $f(x+y) = f(x) + f(y)$.

2° $f(\lambda x) = \lambda f(x)$.

Выберем в n-мерном пространстве произвольный базис $e_1, e_2, \ldots, e_n$. Так как каждый вектор x можно представить в виде

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

то в силу свойств линейной функции имеем:

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

Итак: в *n-мерном пространстве с заданным базисом линейная функция может быть представлена в виде*

$$f(x) = a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n, \tag{1}$$

где $a_i = f(e_i)$ — постоянные, зависящие лишь от выбора базиса, а $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора в этом базисе.

Таким образом, данное выше определение линейной функции совпадает, по существу, с принятым в алгебре определением линейной функции (линейной формы); надо лишь иметь в виду, что в нашем случае коэффициенты зависят от выбора базиса.

Выясним, как меняются коэффициенты линейной функции при замене одного базиса другим.

Пусть $e_1, e_2, \ldots, e_n$ и $e_1', e_2', \ldots, e_n'$ — два базиса в R. Пусть, далее, векторы e_i' выражаются через базис e_1,

$e_2, \ldots, e_n$ формулами

$$\begin{aligned} e_1' &= \alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n, \\ e_2' &= \alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n' &= \alpha_{1n}e_1 + \alpha_{2n}e_2 + \ldots + \alpha_{nn}e_n. \end{aligned}$$

Пусть в базисе $e_1, e_2, \ldots, e_n$ линейная функция выражается формулой

$$f(x) = a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n,$$

а в базисе $e_1', e_2', \ldots, e_n'$ — формулой

$$f(x) = a_1'\xi_1' + a_2'\xi_2' + \ldots + a_n'\xi_n'.$$

Так как $a_i = f(e_i)$, а $a_k' = f(e_k')$, то

$$\begin{aligned} a_k' &= f(\alpha_{1k}e_1 + \alpha_{2k}e_2 + \ldots + \alpha_{nk}e_n) = \\ &= \alpha_{1k}f(e_1) + \alpha_{2k}f(e_2) + \ldots + \alpha_{nk}f(e_n) = \\ &= \alpha_{1k}a_1 + \alpha_{2k}a_2 + \ldots + \alpha_{nk}a_n. \end{aligned}$$

Мы видим, следовательно, что *коэффициенты линейной формы преобразуются при переходе к другому базису так же, как векторы базиса* (или, как иногда говорят, *когредиентно* векторам базиса).

Пример 1. В пространстве, векторами которого являются непрерывные функции $\varphi(t)$, заданные на отрезке $[a, b]$, рассмотрим функцию $f(\varphi)$, заданную формулой

$$f(\varphi) = \int_a^b \varphi(t)\, dt.$$

Эта функция линейна, так как выполняются условия 1° и 2°.

Действительно, первое из них означает, что интеграл суммы равен сумме интегралов, а второе означает, что постоянный множитель можно выносить за знак интеграла.

П р и м е р 2. В том же пространстве рассмотрим функцию $f(\varphi)$, определенную следующим образом. Выберем на отрезке $[a,b]$ некоторое значение $t=t_0$ и положим

$$f(\varphi)=\varphi(t_0).$$

Проверьте, что эта функция $f(\varphi)$ также линейна.

2. Билинейные формы. Существенную роль в дальнейшем будут играть билинейные и квадратичные функции (формы).

О п р е д е л е н и е 2. *Мы говорим, что $A(x;y)$ есть билинейная функция (билинейная форма) от векторов x и y, если:*

1° *при фиксированном y $A(x;y)$ есть линейная функция от x,*

2° *при фиксированном x $A(x;y)$ есть линейная функция от y.*

Иными словами, в силу определения линейной функции условия 1° и 2° означают соответственно

$$\begin{aligned}
&1^\circ &\quad A(x_1+x_2;y)&=A(x_1;y)+A(x_2;y),\\
&&A(\lambda x;y)&=\lambda A(x;y).\\
&2^\circ &\quad A(x;y_1+y_2)&=A(x;y_1)+A(x;y_2),\\
&&A(x;\mu y)&=\mu A(x;y).
\end{aligned}$$

П р и м е р ы. 1. Рассмотрим n-мерное пространство, в котором вектор есть совокупность n чисел. Положим

$$\begin{aligned}
A(x;y)=\;&a_{11}\xi_1\eta_1+a_{12}\xi_1\eta_2+\ldots+a_{1n}\xi_1\eta_n+\\
&+a_{21}\xi_2\eta_1+a_{22}\xi_2\eta_2+\ldots+a_{2n}\xi_2\eta_n+\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&+a_{n1}\xi_n\eta_1+a_{n2}\xi_n\eta_2+\ldots+a_{nn}\xi_n\eta_n,
\end{aligned}\tag{2}$$

где x есть вектор $(\xi_1,\xi_2,\ldots,\xi_n)$, а y — вектор $(\eta_1,\eta_2,\ldots,\eta_n)$. Формула (2) определяет билинейную функцию. В самом деле, если зафиксировать y, т. е. считать $\eta_1, \eta_2, \ldots, \eta_n$ постоянными, то $\sum\limits_{i,k=1}^{n} a_{ik}\xi_i\eta_k$ зависит от ξ_i линейно, т. е. есть линейная функция от $x = (\xi_1, \xi_2, \ldots, \xi_n)$, а при постоянных $\xi_1, \xi_2, \ldots, \xi_n$ форма $A(x; y)$ — линейная функция от y.

2. В пространстве, в котором векторами являются непрерывные функции $f(t)$, рассмотрим следующий пример билинейной функции. Пусть $K(s, t)$ — некоторая непрерывная функция переменных s и t. Положим

$$A(f; g) = \int\limits_a^b \int\limits_a^b K(s, t) f(s) g(t)\, ds\, dt.$$

$A(f; g)$ есть билинейная функция векторов f и g.

Действительно, условия 1° и 2° проверяются так же, как и в примере 1 предыдущего пункта.

Если $K(s, t) \equiv 1$, то

$$A(f; g) = \int\limits_a^b \int\limits_a^b f(s) g(t)\, ds\, dt = \int\limits_a^b f(s)\, ds \int\limits_a^b g(t)\, dt,$$

т. е. $A(f, g)$ есть произведение линейных функций $\int\limits_a^b f(s)\, ds$ и $\int\limits_a^b g(t)\, dt$.

Упражнение. Показать, что если $f(x)$ и $g(y)$ — линейные функции, то их произведение $f(x) \cdot g(y)$ есть билинейная функция.

Определение 3. *Билинейная функция* (форма) *называется симметрической, если для любых векторов x и y имеет место равенство*

$$A(x; y) = A(y; x).$$

В приведенном выше примере 1 определенная формулой (2) билинейная форма $A(x; y)$ симметрична тогда и только тогда, когда $a_{ik} = a_{ki}$ для любых i и k.

Скалярное произведение (x, y) в евклидовом пространстве является примером симметрической билинейной формы.

В самом деле, аксиомы 1°, 2°, 3° скалярного произведения (§ 2) как раз и означают, что скалярное произведение есть симметрическая билинейная форма.

3. Матрица билинейной формы. Мы определили билинейную форму аксиоматически. Выберем теперь в n-мерном пространстве какой-либо базис $e_1, e_2, \ldots, e_n$ и выразим билинейную форму $A(x; y)$ через координаты $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$ векторов x и y в этом базисе. Мы имеем:

$$A(x; y) = A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n;\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n).$$

В силу свойств 1° и 2° билинейной формы

$$\begin{aligned} A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n;\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n) = \\ = \xi_1\eta_1 A(e_1; e_1) + \xi_1\eta_2 A(e_1; e_2) + \ldots + \xi_1\eta_n A(e_1; e_n) + \\ + \xi_2\eta_1 A(e_2; e_1) + \xi_2\eta_2 A(e_2; e_2) + \ldots + \xi_2\eta_n A(e_2; e_n) + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ + \xi_n\eta_1 A(e_n; e_1) + \xi_n\eta_2 A(e_n; e_2) + \ldots + \xi_n\eta_n A(e_n; e_n), \end{aligned}$$

или, короче

$$A(x, y) = \sum_{i,k=1}^{n} A(e_i; e_k)\xi_i\eta_k.$$

Обозначим постоянные $A(e_i; e_k)$ через a_{ik}. Тогда имеем:

при заданном базисе $e_1, e_2, \ldots, e_n$ всякая билинейная форма в n-мерном пространстве может быть записана в виде

$$A(x; y) = \sum_{i,k=1}^{n} a_{ik}\xi_i\eta_k, \tag{3}$$

где $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x, а $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора y в данном базисе.

Числа a_{ik} зависят от выбора базиса и вычисляются по формулам

$$a_{ik} = A(e_i; e_k). \tag{4}$$

Матрица $\boldsymbol{A} = \|a_{ik}\|$ называется *матрицей билинейной формы $A(x; y)$ в базисе $e_1, e_2, \ldots, e_n$*.

Таким образом, в каждом базисе билинейная форма $A(x; y)$ определяется своей матрицей $\boldsymbol{A} = \|a_{ik}\|$.

Пример. Пусть R — трехмерное пространство, векторами которого являются тройки чисел (ξ_1, ξ_2, ξ_3). Зададим в R билинейную форму $A(x; y)$ формулой

$$A(x; y) = \xi_1\eta_1 + 2\xi_2\eta_2 + 3\xi_3\eta_3.$$

Возьмем в R в качестве базиса три вектора

$$e_1 = (1, 1, 1); \quad e_2 = (1, 1, -1); \quad e_3 = (1, -1, -1).$$

Найдем матрицу $\boldsymbol{A}$ билинейной формы $A(x; y)$ в этом базисе. В силу (4) получим:

$$\begin{aligned}
&a_{11} = 1 \cdot 1 + 2 \cdot 1 \cdot 1 + 3 \cdot 1 \cdot 1 = 6,\\
&a_{12} = a_{21} = 1 \cdot 1 + 2 \cdot 1 \cdot 1 + 3 \cdot 1 \cdot (-1) = 0,\\
&a_{22} = 1 \cdot 1 + 2 \cdot 1 \cdot 1 + 3 \cdot (-1) \cdot (-1) = 6,\\
&a_{13} = a_{31} = 1 \cdot 1 + 2 \cdot 1 \cdot (-1) + 3 \cdot 1 \cdot (-1) = -4,\\
&a_{23} = a_{32} = 1 \cdot 1 + 2 \cdot 1 \cdot (-1) + 3 \cdot (-1) \cdot (-1) = 2,\\
&a_{33} = 1 \cdot 1 + 2 \cdot (-1) \cdot (-1) + 3 \cdot (-1) \cdot (-1) = 6,
\end{aligned}$$

т. е.

$$\boldsymbol{A} = \begin{pmatrix} 6 & 0 & -4 \\ 0 & 6 & 2 \\ -4 & 2 & 6 \end{pmatrix}.$$

Таким образом, если обозначить через (ξ_1', ξ_2', ξ_3') и $\eta_1', \eta_2', \eta_3'$ координаты векторов x и y в базисе e_1, e_2, e_3, то

$$A(x; y) = 6\xi_1'\eta_1' - 4\xi_1'\eta_3' + 6\xi_2'\eta_2' + 2\xi_2'\eta_3' - 4\xi_3'\eta_1' + 2\xi_3'\eta_2' + 6\xi_3'\eta_3'.$$

4. Преобразование матрицы билинейной формы при изменении базиса. Пусть даны в n-мерном пространстве два базиса: $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_n$. Пусть векторы $f_1, f_2, \ldots, f_n$ выражаются через векторы базиса

$e_1, e_2, \ldots, e_n$ формулами

$$\left.\begin{aligned} f_1 &= c_{11}e_1 + c_{21}e_2 + \ldots + c_{n1}e_n, \\ f_2 &= c_{12}e_1 + c_{22}e_2 + \ldots + c_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n &= c_{1n}e_1 + c_{2n}e_2 + \ldots + c_{nn}e_n. \end{aligned}\right\} \tag{5}$$

Таким образом, $c_{1k}, c_{2k}, \ldots, c_{nk}$ — координаты вектора f_k в базисе $e_1, e_2, \ldots, e_n$. Матрицу

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix}$$

назовем матрицей перехода от базиса $e_1, e_2, \ldots, e_n$ к базису $f_1, f_2, \ldots, f_n$.

Пусть $\boldsymbol{A} = \|a_{ik}\|$ есть матрица билинейной формы $A(x;y)$ в базисе $e_1, e_2, \ldots, e_n$, а $\boldsymbol{B} = \|b_{ik}\|$ — матрица той же билинейной формы в базисе $f_1, f_2, \ldots, f_n$. Наша задача состоит в том, чтобы по матрице $\|a_{ik}\|$ найти матрицу $\|b_{ik}\|$.

По определению [формула (4)] $b_{pq} = A(f_p; f_q)$, т. е. b_{pq} — значение билинейной формы $A(x;y)$ при $x = f_p$, $y = f_q$; для того чтобы найти его, воспользуемся формулой (3), подставив в нее вместо $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$ координаты векторов f_p и f_q в базисе $e_1, e_2, \ldots, e_n$, т. е. числа $c_{1p}, c_{2p}, \ldots, c_{np}$ и $c_{1q}, c_{2q}, \ldots, c_{nq}$. Получим:

$$b_{pq} = A(f_p; f_q) = \sum_{i,k=1}^{n} a_{ik} c_{ip} c_{kq}. \tag{6}$$

Это есть искомая формула.

Запишем ее в матричной форме. Для этого положим $c'_{pi} = c_{ip}$; таким образом, c'_{pi} являются элементами матрицы $\boldsymbol{C}'$, транспонированной к матрице $\boldsymbol{C}$. Тогда

$$b_{pq} = \sum_{i,k=1}^{n} c'_{pi} a_{ik} c_{kq}.$$

В матричной форме это означает *):

$$\boldsymbol{B} = \boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}. \tag{7}$$

Итак: *если* $\boldsymbol{A}$ *и* $\boldsymbol{B}$ *суть матрицы билинейной формы* $A(x;y)$ *соответственно в базисах* $e_1, e_2, \ldots, e_n$ *и* $f_1, f_2, \ldots, f_n$*, то* $\boldsymbol{B} = \boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$*, где* $\boldsymbol{C}$ *— матрица перехода от базиса* $e_1, e_2, \ldots, e_n$ *к базису* $f_1, f_2, \ldots, f_n$*, а* $\boldsymbol{C}'$ *— матрица, транспонированная к матрице* $\boldsymbol{C}$*.*

5. Квадратичные формы.

Определение 4. *Пусть* $A(x;y)$ *— симметрическая билинейная форма. Функция* $A(x;x)$*, которая получается из* $A(x;y)$*, если положить* $y = x$*, называется квадратичной формой.*

---

*) Произведение двух матриц определяется, как известно, так: элемент произведения матриц, стоящий на пересечении i-й строки и k-го столбца, равен сумме произведений элементов i-й строки первой матрицы на соответствующие элементы k-го столбца второй матрицы. Удобнее это правило записывать в виде формулы

$$c_{ik} = \sum_{\alpha=1}^{n} a_{i\alpha} b_{\alpha k},$$

где $a_{i\alpha}$ — элементы первой матрицы, а $b_{\alpha k}$ — элементы второй матрицы. Отсюда, применяя это правило дважды, можно получить, что произведение трех матриц вычисляется так: если $\boldsymbol{D} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{C}$, то

$$d_{ik} = \sum_{\alpha,\beta=1}^{n} a_{i\alpha} b_{\alpha\beta} c_{\beta k}.$$

Если $\boldsymbol{A}'$ — матрица, транспонированная к матрице $\boldsymbol{A}$, то элементы матрицы $\boldsymbol{A}'\boldsymbol{B}\boldsymbol{C}$ имеют вид $\sum\limits_{\alpha,\beta=1}^{n} a_{\alpha i} b_{\alpha\beta} c_{\beta k}$.

$A(x;y)$ называется билинейной формой, *полярной* к квадратичной форме $A(x;x)$.

Требование симметричности формы $A(x;y)$ в определении квадратичной формы оправдывается следующим предложением, которое без этого было бы неверно.

Полярная форма $A(x;y)$ однозначно определяется своей квадратичной формой $A(x;x)$.

Доказательство. Из определения билинейной формы легко следует, что

$$A(x+y;x+y) = A(x;x) + A(x;y) + A(y;x) + A(y;y).$$

Отсюда в силу симметрии (т. е. равенства $A(x;y) = A(y;x)$) получаем:

$$A(x;y) = \frac{1}{2}[A(x+y;x+y) - A(x;x) - A(y;y)].$$

В правой части этого равенства стоят значения квадратичной формы; следовательно, мы доказали, что билинейная форма $A(x;y)$ определяется своей квадратичной формой *).

Выше мы уже доказали, что всякая симметрическая билинейная форма $A(x;y)$ записывается через коорди-

---

*) Функция $A(x;x)$, полученная из произвольной (не обязательно симметрической) билинейной формы $A(x;y)$, может быть получена и из симметрической билинейной формой. Действительно, пусть $A(x;y)$ — произвольная билинейная форма; тогда

$$A_1(x;y) = \frac{1}{2}[A(x;y) + A(y;x)]$$

есть снова билинейная форма и притом, как можно видеть, симметрическая. Но

$$A_1(x;x) = \frac{1}{2}[A(x;x) + A(x;x)] = A(x;x),$$

т. е. $A_1(x;y)$ приводит нас (при $y = x$) к той же квадратичной форме, что и $A(x;y)$.

наты векторов x и y в виде

$$A(x;y) = \sum_{i,k=1}^{n} a_{ik}\xi_i\eta_k,$$

где $a_{ik} = a_{ki}$. Поэтому:

всякая квадратичная форма $A(x;x)$ при заданном базисе выражается формулой

$$A(x;x) = \sum_{i,k=1}^{n} a_{ik}\xi_i\xi_k,$$

где $a_{ik} = a_{ki}$. Введем еще одно важное

Определение 5. *Квадратичная форма $A(x;x)$ называется положительно определенной, если для любого вектора $x \neq 0$*

$$A(x;x) > 0.$$

Пример. $A(x;x) = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2$ является, очевидно, положительно определенной квадратичной формой.

Пусть $A(x;x)$ — положительно определенная квадратичная форма и $A(x;y)$ — ее полярная форма. В силу сформулированных выше определений это означает:

1° $A(x;y) = A(y;x)$.

2° $A(x_1 + x_2;y) = A(x_1;y) + A(x_2;y)$.

3° $A(\lambda x;y) = \lambda A(x;y)$.

4° $A(x;x) \geqslant 0$ и $A(x;x) > 0$ при $x \neq 0$.

Мы видим, что эти условия совпадают с аксиомами скалярного произведения, сформулированными в § 2. Следовательно,

скалярное произведение есть билинейная форма, соответствующая положительно определенной квадратичной форме, и любая такая форма может быть принята за скалярное произведение.

Поэтому мы можем определить евклидово пространство следующим образом.

Евклидовым пространством называется аффинное пространство, в котором выбрана какая-нибудь фиксированная положительно определенная квадратичная форма $A(x;x)$. Значение $A(x;y)$ соответствующей *) *ей билинейной формы считается при этом скалярным произведением* **) *векторов x и y.*

§ 5. Приведение квадратичной формы к сумме квадратов

Мы знаем уже, что выражение квадратичной формы $A(x;x)$ через координаты вектора x зависит от выбора базиса. В этом параграфе будет показано, как привести квадратичную форму к сумме квадратов, т. е. выбрать такой базис (систему координат), в котором квадратичная форма имеет простой вид

$$A(x;x) = \lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \ldots + \lambda_n\xi_n^2. \tag{1}$$

Пусть в некотором базисе $f_1, f_2, \ldots, f_n$ имеем равенство

$$A(x,x) = \sum_{i,k=1}^{n} a_{ik}\eta_i\eta_k, \tag{2}$$

где $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора x в этом базисе. Будем постепенно преобразовывать базис так, чтобы в формуле (2) пропадали произведения координат с различными индексами. Так как каждому преобразованию базиса отвечает определенное преобразование координат (см. § 1, п. 6) и обратно, то мы можем писать формулы преобразования координат.

Для приведения формы $A(x;x)$ к сумме квадратов нам нужно будет, чтобы хоть один из коэффициентов

---

*) Мы уже доказали, что билинейная форма $A(x;y)$ однозначно определяется своей квадратичной формой $A(x;x)$.

**) Выше оно обозначалось (x,y), а не $A(x;y)$.

a_{kk} (коэффициент при η_k^2) был отличен от нуля. Этого всегда можно добиться. Действительно, предположим, что форма $A(x;x)$, не равная тождественно нулю, не содержит ни одного квадрата переменного; тогда она содержит хотя бы одно произведение, например, $2a_{12}\eta_1\eta_2$. Заменим координаты η_1 и η_2 по формулам

$$\begin{aligned} \eta_1 &= \eta_1' + \eta_2', \\ \eta_2 &= \eta_1' - \eta_2', \end{aligned}$$

не изменяя остальных переменных. При этом преобразовании член $2a_{12}\eta_1\eta_2$ перейдет в $2a_{12}(\eta_1'^2 - \eta_2'^2)$, и так как, по предположению, $a_{11} = a_{22} = 0$, то он ни с чем не может сократиться, т. е. коэффициент при $\eta_1'^2$ отличен от нуля.

Будем теперь считать, что уже в формуле (2) коэффициент $a_{11} \neq 0$ *). Выделим в нашей квадратичной форме члены, содержащие η_1:

$$a_{11}\eta_1^2 + 2a_{12}\eta_1\eta_2 + \ldots + 2a_{1n}\eta_1\eta_n.$$

Дополним эту сумму до полного квадрата, т. е. запишем ее в виде

$$a_{11}\eta_1^2 + 2a_{12}\eta_1\eta_2 + \ldots + 2a_{1n}\eta_1\eta_n = \\ = \frac{1}{a_{11}}(a_{11}\eta_1 + \ldots + a_{1n}\eta_n)^2 - B, \quad (3)$$

где через B мы обозначили члены, содержащие лишь квадраты и попарные произведения членов $a_{12}\eta_2, \ldots$ $\ldots, a_{1n}\eta_n$. После подстановки выражения (3) в (2) рассматриваемая квадратичная форма примет вид

$$A(x;x) = \frac{1}{a_{11}}(a_{11}\eta_1 + \ldots + a_{1n}\eta_n)^2 + \ldots,$$

---

*) Если $a_{11} = 0$, но отличен от нуля коэффициент при квадрате какой-нибудь другой координаты, то к рассматриваемому случаю можно прийти, иначе занумеровав векторы $e_1, e_2, \ldots, e_n$, что также является некоторым преобразованием этого базиса.

где невыписанные члены содержат только переменные $\eta_2, \ldots, \eta_n$.

Положим

$$\begin{aligned} \eta_1^* &= a_{11}\eta_1 + a_{12}\eta_2 + \ldots + a_{1n}\eta_n, \\ \eta_2^* &= \qquad\qquad \eta_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \eta_n^* &= \qquad\qquad\qquad\qquad \eta_n. \end{aligned}$$

Тогда квадратичная форма примет вид

$$A(x;x) = \frac{1}{a_{11}}\eta_1^{*2} + \sum_{i,k=2}^{n} a_{ik}^*\eta_i^*\eta_k^*.$$

Выражение $\sum\limits_{i,k=1}^{n} a_{ik}^*\eta_i^*\eta_k^*$ вполне аналогично правой части формулы (2) с той только разницей, что оно не содержит первой координаты. Предположим, что коэффициент $a_{22}^* \neq 0$ (этого, как мы видели, всегда можно добиться простыми вспомогательными преобразованиями). Тогда можно произвести новое, аналогичное первому, преобразование переменных по формулам

$$\begin{aligned} \eta_1^{**} &= \eta_1^*, \\ \eta_2^{**} &= \qquad a_{22}^*\eta_2^* + a_{23}^*\eta_3^* + \ldots + a_{2n}^*\eta_n^*, \\ \eta_3^{**} &= \qquad\qquad \eta_3^*, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \eta_n^{**} &= \qquad\qquad\qquad\qquad \eta_n^*. \end{aligned}$$

В новых переменных форма примет вид

$$A(x;x) = \frac{1}{a_{11}}\eta_1^{**2} + \frac{1}{a_{22}^*}\eta_2^{**2} + \sum_{i,k=3}^{n} a_{ik}^{**}\eta_i^{**}\eta_k^{**}.$$

Продолжая этот процесс, мы после конечного числа шагов придем к переменным $\xi_1, \xi_2, \ldots, \xi_n$, в которых форма $A(x;x)$ будет иметь вид

$$A(x;x) = \lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \ldots + \lambda_m\xi_m^2,$$

причем $m \leqslant n$.

Мы предоставляем читателю выписать преобразование базиса, отвечающее каждому из произведенных преобразований координат (см. п. 6 § 1), и убедиться, что наши преобразования действительно переводят базис снова в базис, т.е. что полученные из базиса в результате преобразования векторы снова линейно независимы.

Полагая в случае $m < n$, что $\lambda_{m+1} = \ldots = \lambda_n = 0$, мы приходим, таким образом, к следующей теореме:

Теорема. *Пусть в n-мерном пространстве задана произвольная квадратичная форма $A(x;x)$. Тогда в R существует базис $e_1, e_2, \ldots, e_n$, в котором эта квадратичная форма имеет вид*

$$A(x;x) = \lambda_1\xi_1^2 + \lambda_2\xi_2^2 + \ldots + \lambda_n\xi_n^2,$$

где $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x в базисе $e_1, e_2, \ldots, e_n$.

Приведем пример приведения квадратичной формы к сумме квадратов по описанному методу. Пусть в трехмерном пространстве с некоторым базисом f_1, f_2, f_3 задана квадратичная форма

$$A(x;x) = 2\eta_1\eta_2 + 4\eta_1\eta_3 - \eta_2^2 - 8\eta_3^2.$$

Положим

$$\begin{aligned} \eta_1 &= \eta_2', \\ \eta_2 &= \eta_1', \\ \eta_3 &= \eta_3'. \end{aligned}$$

Тогда получим:

$$A(x;x) = -\eta_1'^2 + 2\eta_1'\eta_2' + 4\eta_2'\eta_3' - 8\eta_3'^2.$$

Дальше, полагая

$$\begin{aligned} \eta_1^* &= -\eta_1' + \eta_2', \\ \eta_2^* &= \qquad\ \ \eta_2', \\ \eta_3^* &= \qquad\qquad \eta_3', \end{aligned}$$

мы получим новое выражение для квадратичной формы:

$$A(x;x) = -\eta_1^{*\,2} + \eta_2^{*\,2} + 4\eta_2^*\eta_3^* - 8\eta_3^{*\,2}.$$

Преобразование

$$\begin{aligned} \xi_1 &= \eta_1^*, \\ \xi_2 &= \quad \eta_2^* + 2\eta_3^*, \\ \xi_3 &= \qquad\quad \eta_3^* \end{aligned}$$

выделит из нашей квадратичной формы еще один полный квадрат, после чего форма примет канонический вид:

$$A(x;x) = -\xi_1^2 + \xi_2^2 - 12\xi_3^2.$$

Имея формулы, выражающие $\eta_1^*, \eta_2^*, \ldots, \eta_n^*$ через $\eta_1, \eta_2, \ldots, \eta_n$, затем $\eta_1^{**}, \ldots, \eta_n^{**}$ через $\eta_1^*, \eta_2^*, \ldots, \eta_n^*$ и т. д., мы можем получить выражение координат $\xi_1, \xi_2, \ldots, \xi_n$ через первоначальные координаты $\eta_1, \eta_2, \ldots, \eta_n$:

$$\begin{aligned} \xi_1 &= c_{11}\eta_1 + c_{12}\eta_2 + \ldots + c_{1n}\eta_n, \\ \xi_2 &= c_{21}\eta_1 + c_{22}\eta_2 + \ldots + c_{2n}\eta_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \xi_n &= c_{n1}\eta_1 + c_{n2}\eta_2 + \ldots + c_{nn}\eta_n. \end{aligned}$$

Так, в приведенном выше примере эти формулы имеют вид

$$\begin{aligned} \xi_1 &= \eta_1 - \eta_2, \\ \xi_2 &= \eta_1 \qquad + 2\eta_3, \\ \xi_3 &= \qquad\qquad \eta_3. \end{aligned}$$

Вспоминая (§ 1, п. 6), что матрица, дающая преобразование координат, является обратной и транспонированной к матрице преобразования базиса, мы можем выразить векторы нового базиса $e_1, e_2, \ldots, e_n$ через векторы старого базиса $f_1, f_2, \ldots, f_n$:

$$\begin{aligned} e_1 &= d_{11}f_1 + d_{12}f_2 + \ldots + d_{1n}f_n, \\ e_2 &= d_{21}f_1 + d_{22}f_2 + \ldots + d_{2n}f_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n &= d_{n1}f_1 + d_{n2}f_2 + \ldots + d_{nn}f_n. \end{aligned}$$

Если в процессе приведения нам ни разу не приходилось производить преобразования, меняющего сразу две координаты (такое преобразование, как мы помним, приходится совершать, когда в преобразуемой форме отсутствуют квадраты координат, либо если приходилось менять нумерацию), то формулы преобразования имеют вид

$$\begin{aligned} \xi_1 &= c_{11}\eta_1 + c_{12}\eta_2 + \ldots + c_{1n}\eta_n, \\ \xi_2 &= \qquad\quad\; c_{22}\eta_2 + \ldots + c_{2n}\eta_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \xi_n &= \qquad\qquad\qquad\qquad\quad c_{nn}\eta_n, \end{aligned}$$

т. е. матрица преобразования является так называемой треугольной матрицей. Легко проверить, что матрица преобразования базиса будет в этом случае также треугольной матрицей вида

$$\begin{aligned} e_1 &= d_{11}f_1, \\ e_2 &= d_{21}f_1 + d_{22}f_2, \\ &\ldots\ldots\ldots\ldots \\ e_n &= d_{n1}f_1 + d_{n2}f_2 + \ldots + d_{nn}f_n. \end{aligned}$$

Здесь $d_{\alpha\beta}$ — алгебраическое дополнение элемента $c_{\alpha\beta}$ матрицы $\|c_{ik}\|$, деленное на определитель этой матрицы.

§ 6. Приведение квадратичной формы к сумме квадратов треугольным преобразованием

1. В этом параграфе мы укажем еще один способ построения базиса, в котором квадратичная форма приводится к сумме квадратов. В отличие от предыдущего параграфа мы дадим формулы, выражающие искомый базис $e_1, e_2, \ldots, e_n$ непосредственно через исходный базис (а не в несколько шагов, как в § 5).

При этом, однако, мы должны будем на форму $A(x;y)$ и исходный базис $f_1, f_2, \ldots, f_n$ наложить следующее ограничение: пусть $\|a_{ik}\|$ — матрица билинейной формы $A(x;y)$ в базисе $f_1, f_2, \ldots, f_n$. Мы предположим, что следующие миноры матрицы $\|a_{ik}\|$ все отличны от нуля *):

$$\Delta_1 = a_{11} \neq 0; \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0; \quad \ldots;$$

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \neq 0. \tag{1}$$

В каждом базисе $f_1, f_2, \ldots, f_n$ квадратичная форма $A(x;x)$ имеет вид

$$A(x,x) = \sum_{i,k=1}^{n} a_{ik}\xi_i\xi_k, \quad \text{где} \quad a_{ik} = A(f_i, f_k).$$

Наша цель — определить векторы $e_1, e_2, \ldots, e_n$ так, чтобы

$$A(e_i; e_k) = 0 \quad \text{при} \quad i \neq k \quad (i, k = 1, 2, \ldots, n). \tag{2}$$

Процесс, с помощью которого это будет сделано, совпадает с процессом ортогонализации, описанным в п. 1 § 3, если заменить в этом процессе скалярное произведение векторов произвольной билинейной формой $A(x;y)$, удовлетворяющей условиям (1).

---

*) Можно показать, что это требование равносильно тому, что при приведении квадратичной формы к сумме квадратов по методу, описанному в § 5, $a_{11} \neq 0$, $a_{22}^* \neq 0$ и т. д.

Будем искать векторы $e_1, e_2, \ldots, e_n$ в виде

$$\left.\begin{aligned} &e_1 = \alpha_{11} f_1, \\ &e_2 = \alpha_{21} f_1 + \alpha_{22} f_2, \\ &\ldots\ldots\ldots\ldots\ldots \\ &e_n = \alpha_{n1} f_1 + \alpha_{n2} f_2 + \ldots + \alpha_{nn} f_n. \end{aligned}\right\} \tag{3}$$

Коэффициенты α_{ik} можно было бы найти из условий (2), подставив в эти условия вместо $e_1, e_2, \ldots, e_n$ их выражения из (3). Однако это неудобно для вычислений, так как пришлось бы решать уравнения второй степени относительно α_{ik}. Поступим поэтому несколько иначе.

Если

$$A(e_k; f_i) = 0 \quad \text{для} \quad i = 1, 2, \ldots, k-1,$$

то и

$$A(e_k; e_i) = 0 \quad \text{для} \quad i = 1, 2, \ldots, k-1.$$

Действительно, подставляя вместо e_i выражение

$$\alpha_{i1} f_1 + \alpha_{i2} f_2 + \ldots + \alpha_{ii} f_i,$$

получаем:

$$\begin{aligned} A(e_k; e_i) = A(e_k;\ &\alpha_{i1} f_1 + \alpha_{i2} f_2 + \ldots + \alpha_{ii} f_i) = \\ &= \alpha_{i1} A(e_k; f_1) + \alpha_{i2} A(e_k; f_2) + \ldots + \alpha_{ii} A(e_k; f_i). \end{aligned}$$

Таким образом, если $A(e_k; f_i) = 0$ для любого k и для любого $i < k$, то и $A(e_k; e_i) = 0$ для $i < k$, и следовательно, в силу симметрии билинейной формы и для $i > k$, т. е. $e_1, e_2, \ldots, e_n$ — требуемый базис. Наша задача сведена, таким образом, к следующей:

определить коэффициенты $\alpha_{k1}, \alpha_{k2}, \ldots, \alpha_{kk}$ так, чтобы вектор

$$e_k = \alpha_{k1} f_1 + \alpha_{k2} f_2 + \ldots + \alpha_{kk} f_k$$

удовлетворял условиям

$$A(e_k; f_i) = 0, \quad i = 1, 2, \ldots, k-1. \tag{4}$$

Этими условиями вектор e_k определяется с точностью до постоянного множителя. Зафиксируем этот множитель с помощью требования

$$A(e_k; f_k) = 1. \tag{5}$$

Мы увидим сейчас, что условиями (4) и (5) вектор e_k определен уже однозначно.

Подставив в (4) и (5) выражение для e_k, мы получим следующую систему уравнений первой степени относительно α_{ki}:

$$\left.\begin{aligned}
&\alpha_{k1}A(f_1; f_1) + \alpha_{k2}A(f_1; f_2) + \ldots + \alpha_{kk}A(f_1; f_k) = 0,\\
&\alpha_{k1}A(f_2; f_1) + \alpha_{k2}A(f_2; f_2) + \ldots + \alpha_{kk}A(f_2; f_k) = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\alpha_{k1}A(f_{k-1}; f_1) + \alpha_{k2}A(f_{k-1}; f_2) + \ldots\\
&\qquad\qquad\qquad\ldots + \alpha_{kk}A(f_{k-1}; f_k) = 0,\\
&\alpha_{k1}A(f_k; f_1) + \alpha_{k2}A(f_k; f_2) + \ldots + \alpha_{kk}A(f_k; f_k) = 1.
\end{aligned}\right\} \tag{6}$$

Определитель этой системы уравнений равен

$$\Delta_k = \begin{vmatrix} A(f_1; f_1) & A(f_1; f_2) & \ldots & A(f_1; f_k)\\ A(f_2; f_1) & A(f_2; f_2) & \ldots & A(f_2; f_k)\\ \ldots & \ldots & \ldots & \ldots\\ A(f_k; f_1) & A(f_k; f_2) & \ldots & A(f_k; f_k) \end{vmatrix} \tag{7}$$

и по условию (1) отличен от нуля. Поэтому решение системы (6) существует и единственно. Таким образом, задача нахождения вектора e_k нами решена для любого k.

Найдем теперь коэффициенты b_{ik} квадратичной формы $A(x; x)$ в базисе $e_1, e_2, \ldots, e_n$. Как нам уже известно,

$$b_{ik} = A(e_i; e_k).$$

По построению этого базиса, $A(e_i; e_k) = 0$ при $i \neq k$, т. е. $b_{ik} = 0$ при $i \neq k$.

Вычислим $b_{kk} = A(e_k; e_k)$:

$$A(e_k; e_k) = A(e_k;\ \alpha_{k1} f_1 + \alpha_{k2} f_2 + \ldots + \alpha_{kk} f_k) =$$
$$= \alpha_{k1} A(e_k; f_1) + \alpha_{k2} A(e_k; f_2) + \ldots + \alpha_{kk} A(e_k; f_k)\ ^{*)}$$

и в силу условий (4) и (5)

$$A(e_k; e_k) = \alpha_{kk}.$$

Число α_{kk} можно найти из системы (6); согласно правилу Крамера

$$\alpha_{kk} = \frac{\Delta_{k-1}}{\Delta_k},$$

где Δ_{k-1} — определитель, аналогичный (7) порядка $k-1$, и где положено $\Delta_0 = 1$.

Таким образом,

$$b_{kk} = A(e_k; e_k) = \frac{\Delta_{k-1}}{\Delta_k}.$$

Итак, доказана следующая

Теорема 1. *Пусть в базисе $f_1, f_2, \ldots, f_n$ квадратичная форма имеет вид*

$$A(x; x) = \sum a_{ik} \eta_i \eta_k, \quad \text{где} \quad a_{ik} = A(f_i; f_k).$$

Пусть, далее, определители

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \ldots,$$

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

отличны от нуля. Тогда существует базис $e_1, e_2, \ldots$ $\ldots, e_n$, в котором $A(x; x)$ записывается в виде суммы

---

*) Выкладка сильно усложнилась бы, если бы мы заменили e_k его выражением через f_i и на первом, и на втором месте.

квадратов следующим образом:

$$A(x;x) = \frac{\Delta_0}{\Delta_1}\xi_1^2 + \frac{\Delta_1}{\Delta_2}\xi_2^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n}\xi_n^2,$$

где ξ_k — *координаты вектора* x *в базисе* $e_1, e_2, \ldots, e_n$.

Этот способ приведения квадратичной формы к сумме квадратов обычно называется *методом Якоби.*

З а м е ч а н и е. В процессе доказательства теоремы мы пришли к некоторому вполне определенному базису $e_1, e_2, \ldots, e_n$. Это, конечно, не означает, что базис, в котором квадратичная форма приводится к сумме квадратов, вообще единствен. Действительно, если взять другой исходный базис $f_1, f_2, \ldots, f_n$ (даже просто, если занумеровать его векторы в другом порядке), то описанный выше процесс приведет нас, вообще говоря, к другому базису $e_1, e_2, \ldots, e_n$ (не говоря уже о том, что базис $e_1, e_2, \ldots, e_n$ не обязательно искать в виде (3)).

П р и м е р. Рассмотрим квадратичную форму

$$2\xi_1^2 + 3\xi_1\xi_2 + 4\xi_1\xi_3 + \xi_2^2 + \xi_3^2$$

в трехмерном пространстве с базисом

$$f_1 = (1, 0, 0), \quad f_2 = (0, 1, 0), \quad f_3 = (0, 0, 1).$$

Соответствующая ей билинейная форма имеет вид

$$A(x;y) = 2\xi_1\eta_1 + \frac{3}{2}\xi_1\eta_2 + 2\xi_1\eta_3 + \frac{3}{2}\xi_2\eta_1 + \xi_2\eta_2 + 2\xi_3\eta_1 + \xi_3\eta_3.$$

Вычислив определители Δ_1, Δ_2 и Δ_3, получим, что они равны соответственно 2, $-\frac{1}{4}$ и $-4\frac{1}{4}$, т. е. ни один из них не нуль. Условия теоремы, таким образом, выполнены. Положим

$$\begin{aligned} e_1 &= \alpha_{11}f_1 &&= (\alpha_{11}, 0, 0),\\ e_2 &= \alpha_{21}f_1 + \alpha_{22}f_2 &&= (\alpha_{21}, \alpha_{22}, 0),\\ e_3 &= \alpha_{31}f_1 + \alpha_{32}f_2 + \alpha_{33}f_3 &&= (\alpha_{31}, \alpha_{32}, \alpha_{33}). \end{aligned}$$

Коэффициент α_{11} находим из условия

$$A(e_1; f_1) = 1,$$

т. е. $2\alpha_{11} = 1$, или $\alpha_{11} = \frac{1}{2}$ и, значит,

$$e_1 = \frac{1}{2} f_1 = \left(\frac{1}{2}, 0, 0\right).$$

Для α_{21} и α_{22} имеем уравнения

$$A(e_2; f_1) = 0 \quad \text{и} \quad A(e_2; f_2) = 1,$$

или

$$2\alpha_{21} + \frac{3}{2}\alpha_{22} = 0; \quad \frac{3}{2}\alpha_{21} + \alpha_{22} = 1,$$

откуда

$$\alpha_{21} = 6, \quad \alpha_{22} = -8,$$

т. е. $e_2 = 6f_1 - 8f_2 = (6, -8, 0)$.

Наконец, для $\alpha_{31}, \alpha_{32}, \alpha_{33}$ имеем систему уравнений

$$A(e_3; f_1) = 0, \quad A(e_3; f_2) = 0, \quad A(e_3; f_3) = 1,$$

т. е.

$$\begin{aligned} 2\alpha_{31} + \frac{3}{2}\alpha_{32} + 2\alpha_{33} &= 0, \\ \frac{3}{2}\alpha_{31} + \alpha_{32} \phantom{+ 2\alpha_{33}} &= 0, \\ 2\alpha_{31} \phantom{+ \frac{3}{2}\alpha_{32}} + \alpha_{33} &= 1, \end{aligned}$$

откуда

$$\alpha_{31} = \frac{8}{17}; \quad \alpha_{32} = -\frac{12}{17}; \quad \alpha_{33} = \frac{1}{17},$$

т. е.

$$e_3 = \frac{8}{17} f_1 - \frac{12}{17} f_2 + \frac{1}{17} f_3 = \left(\frac{8}{17}, -\frac{12}{17}, \frac{1}{17}\right).$$

Квадратичная форма в базисе e_1, e_2, e_3 имеет вид

$$A(x; x) = \frac{1}{\Delta_1}\zeta_1^2 + \frac{\Delta_1}{\Delta_2}\zeta_2^2 + \frac{\Delta_2}{\Delta_3}\zeta_3^2 = \frac{1}{2}\zeta_1^2 - 8\zeta_2^2 + \frac{1}{17}\zeta_3^2,$$

где $\zeta_1, \zeta_2, \zeta_3$ — координаты вектора x в базисе e_1, e_2, e_3.

2. Выше при доказательстве теоремы 1 мы не только построили базис, в котором данная квадратичная форма записывается как сумма квадратов координат, но и получили вполне определенные выражения для коэффициентов при этих квадратах, а именно:

$$\frac{1}{\Delta_1}, \ \frac{\Delta_1}{\Delta_2}, \ \ldots, \ \frac{\Delta_{n-1}}{\Delta_n},$$

так что квадратичная форма имеет вид

$$\frac{1}{\Delta_1}\xi_1^2 + \frac{\Delta_1}{\Delta_2}\xi_2^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n}\xi_n^2. \tag{8}$$

Это дает нам возможность найти число положительных и отрицательных коэффициентов при квадратах. Именно, если Δ_{i-1} и Δ_i имеют одинаковые знаки, то коэффициент при ξ_i^2 положителен, если же их знаки различны, то этот коэффициент отрицателен, т. е. число отрицательных коэффициентов при квадратах равно числу перемен знака в ряду

$$1, \Delta_1, \Delta_2, \ldots, \Delta_n.$$

Итак, доказана следующая теорема.

Т е о р е м а 2. *Число отрицательных коэффициентов при квадратах координат в каноническом виде* (8) *квадратичной формы равно числу перемен знака в последовательности определителей*

$$1, \Delta_1, \Delta_2, \ldots, \Delta_n \ ^{*}).$$

Пусть, в частности, $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_n > 0$. Тогда существует базис $e_1, e_2, \ldots, e_n$, в котором квадратичная форма имеет вид

$$A(x; x) = \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \ldots + \lambda_n \xi_n^2,$$

причем все $\lambda_i > 0$. Следовательно, $A(x; x) \geqslant 0$ для всякого x, и притом равенство

$$A(x; x) = \sum \lambda_i \xi_i^2 = 0$$

возможно, лишь если

$$\xi_1 = \xi_2 = \ldots = \xi_n = 0.$$

Иначе говоря:

Если $\Delta_1 > 0$, $\Delta_2 > 0$, ..., $\Delta_n > 0$, то квадратичная форма $A(x; x)$ — положительно определенная.

---

*) Мы показали, как найти число положительных и отрицательных квадратов при определенном способе приведения квадратичной формы к сумме квадратов. В следующем параграфе будет показано, что это число — одно и то же при всех способах приведения формы к сумме квадратов.

Обратно, пусть $A(x;x)$ — положительно определенная квадратичная форма. Покажем, что в этом случае

$$\Delta_k > 0 \quad (k = 1, 2, \ldots, n);$$

для этого покажем раньше, что $\Delta_k \neq 0$. Предположим противное, т. е. что

$$\Delta_k = \begin{vmatrix} A(f_1;f_1) & A(f_1;f_2) & \ldots & A(f_1;f_k) \\ A(f_2;f_1) & A(f_2;f_2) & \ldots & A(f_2;f_k) \\ \ldots & \ldots & \ldots & \ldots \\ A(f_k;f_1) & A(f_k;f_2) & \ldots & A(f_k;f_k) \end{vmatrix} = 0;$$

тогда одна из строк этого определителя есть линейная комбинация остальных, т. е.

$$\mu_1 A(f_1;f_i) + \mu_2 A(f_2;f_i) + \ldots + \mu_k A(f_k;f_i) = 0,$$

$i = 1, 2, \ldots, k$, где не все μ_j равны нулю. Но тогда

$$A(\mu_1 f_1 + \mu_2 f_2 + \ldots + \mu_k f_k;\ f_i) = 0 \quad (i = 1, 2, \ldots, k),$$

а следовательно,

$$A(\mu_1 f_1 + \mu_2 f_2 + \ldots + \mu_k f_k;\ \mu_1 f_1 + \mu_2 f_2 + \ldots + \mu_k f_k) = 0,$$

в то время как

$$\mu_1 f_1 + \mu_2 f_2 + \ldots + \mu_k f_k \neq 0,$$

что противоречит определению положительно определенной квадратичной формы.

Следовательно, согласно теореме 1, $A(x;x)$ можно привести к виду

$$A(x;x) = \lambda_1 \xi_1^2 + \ldots + \lambda_n \xi_n^2,$$

где

$$\lambda_k = \frac{\Delta_{k-1}}{\Delta_k}.$$

Так как для положительно определенной квадратичной формы все $\lambda_k > 0$, то и все $\Delta_k > 0$. (Напомним, что $\Delta_0 = 1$.)

Итак, нами доказана

Теорема 3. *Пусть $A(x;y)$ — симметрическая билинейная форма и $f_1, f_2, \ldots, f_n$ — базис в n-мерном пространстве R. Для того чтобы квадратичная форма $A(x;x)$ была положительно определенной, необходимо и достаточно, чтобы*

$$\Delta_1 > 0, \quad \Delta_2 > 0, \quad \ldots, \quad \Delta_n > 0.$$

Эта теорема называется *условием Сильвестра* положительной определенности квадратичной формы.

Мы могли бы взять вместо $f_1, f_2, \ldots, f_n$ какой-либо другой базис и написать условия положительной определенности формы $A(x;x)$ через векторы этого нового базиса. В частности, если мы в качестве нового базиса возьмем те же самые векторы $f_1, f_2, \ldots, f_n$, но только в другом порядке, то новыми минорами $\Delta_1, \Delta_2, \ldots, \Delta_n$ будут различные главные миноры *) матрицы $\|a_{ik}\|$. Отсюда вытекает интересное

Следствие. *Если все главные миноры $\Delta_1, \Delta_2, \ldots, \Delta_n$ матрицы $\|a_{ik}\|$ квадратичной формы $A(x;x)$ в данном базисе положительны, то вообще все главные миноры этой матрицы положительны.*

В самом деле, если все миноры Δ_k матрицы $\|a_{ik}\|$ положительны, то форма $A(x;x)$ положительно определенная. Пусть Δ — какой-либо главный минор матрицы $\|a_{ik}\|$ и пусть $p_1, p_2, \ldots, p_k$ — номера входящих в него строк и столбцов этой матрицы (так как минор — главный, то эти номера для строк и столбцов одни и те же). Переставив в исходном базисе векторы с номерами $p_1, p_2, \ldots, p_k$ на первое, второе и т. д., k-е место и записав в этом новом базисе условия положительной определенности формы, получим $\Delta > 0$.

3. Определители Грама. Результаты этого параграфа мы изложим сейчас для случая, когда в качестве квадратичной формы выбрано скалярное произведение в евклидовом пространстве, т. е.

$$A(x;x) \equiv (x,x).$$

*) Главными минорами называются те, при составлении которых выделяются столбцы с теми же номерами, что и строки.

Мы знаем, что скалярное произведение вектора с собой есть положительно определенная квадратичная форма, и обратно, каждая симметрическая билинейная форма, которой соответствует положительно определенная квадратичная форма, может быть принята за скалярное произведение. Поэтому всякая теорема о положительно определенных квадратичных формах является одновременно некоторой теоремой о векторах в евклидовом пространстве.

Пусть $e_1, e_2, \ldots, e_k$ — векторы в евклидовом пространстве.

Определитель

$$\begin{vmatrix} (e_1, e_1) & (e_1, e_2) & \ldots & (e_1, e_k) \\ (e_2, e_1) & (e_2, e_2) & \ldots & (e_2, e_k) \\ \ldots & \ldots & \ldots & \ldots \\ (e_k, e_1) & (e_k, e_2) & \ldots & (e_k, e_k) \end{vmatrix}$$

называется *определителем Грама* этой системы векторов.

Теорема 4. *Определитель Грама любой системы векторов всегда больше или равен нулю. Он равен нулю тогда и только тогда, когда векторы $e_1, e_2, \ldots, e_k$ линейно зависимы.*

Доказательство. Пусть векторы $e_1, e_2, \ldots, e_k$ линейно независимы. Рассмотрим билинейную форму

$$A(x; y) \equiv (x, y),$$

где (x, y) — скалярное произведение векторов x и y. Тогда определитель Грама есть определитель Δ_k, рассмотренный в этом параграфе [см. формулу (7)]. Так как $A(x; x)$ — положительно определенная квадратичная форма, то, в силу теоремы 3, $\Delta_k > 0$.

Докажем, что в случае линейно зависимых векторов определитель Грама равен нулю. Действительно, если $e_1, e_2, \ldots, e_k$ линейно зависимы, то хоть один из них, например e_k, есть линейная комбинация остальных:

$$e_k = \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_{k-1} e_{k-1}.$$

Поэтому последняя строка в определителе Грама есть линейная комбинация остальных. Значит, он равен нулю. Теорема полностью доказана.

В качестве примера рассмотрим определитель Грама двух векторов x и y:

$$\Delta_2 = \begin{vmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{vmatrix}.$$

Утверждение, что $\Delta_2 \geqslant 0$, превращается в этом случае в неравенство Коши–Буняковского.

Примеры. 1. В евклидовом трехмерном пространстве (или на плоскости) определитель

$$\Delta_2 = \begin{vmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{vmatrix}$$

имеет следующий геометрический смысл: Δ_2 равно квадрату площади параллелограмма, построенного на векторах x и y. В самом деле, по определению скалярного произведения

$$(x,y) = (y,x) = |x||y|\cos\varphi,$$

где φ — угол между векторами x и y. Поэтому

$$\Delta_2 = |x|^2|y|^2 - |x|^2|y|^2\cos^2\varphi = |x|^2|y|^2(1-\cos^2\varphi) = |x|^2|y|^2\sin^2\varphi,$$

т. е. Δ_2 равно квадрату площади параллелограмма, построенного на векторах x и y.

2. В трехмерном пространстве объем параллелепипеда, построенного на векторах x, y, z, как показывается в аналитической геометрии, равен абсолютной величине определителя

$$v = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix},$$

где x_i, y_i, z_i — координаты векторов x, y, z в ортогональном базисе. Вычислим квадрат этого определителя, умножая строки на строки. Мы получим:

$$v^2 = \begin{vmatrix} x_1^2+x_2^2+x_3^2 & x_1y_1+x_2y_2+x_3y_3 & x_1z_1+x_2z_2+x_3z_3 \\ y_1x_1+y_2x_2+y_3x_3 & y_1^2+y_2^2+y_3^2 & y_1z_1+y_2z_2+y_3z_3 \\ z_1x_1+z_2x_2+z_3x_3 & z_1y_1+z_2y_2+z_3y_3 & z_1^2+z_2^2+z_3^2 \end{vmatrix} = \begin{vmatrix} (x,x) & (x,y) & (x,z) \\ (y,x) & (y,y) & (y,z) \\ (z,x) & (z,y) & (z,z) \end{vmatrix}.$$

Таким образом, определитель Грама векторов x, y, z равен квадрату объема параллелепипеда, построенного на этих векторах.

Аналогично можно показать, что определитель Грама k векторов $x, y, \ldots, w$ в k-мерном пространстве *) равен квадрату определителя

$$\begin{vmatrix} x_1 & x_2 & \ldots & x_k \\ y_1 & y_2 & \ldots & y_k \\ \ldots & \ldots & \ldots & \ldots \\ w_1 & w_2 & \ldots & w_k \end{vmatrix}, \tag{9}$$

где x_i, соответственно y_i и т. д. — координаты вектора x, соответственно y и т. д. в каком-нибудь ортогональном базисе.

По аналогии с трехмерным пространством модуль определителя (9) называют объемом k-мерного параллелепипеда, определяемого векторами $x, y, \ldots, w$.

3. В пространстве функций (пример 4 § 2) определитель Грама пишется так:

$$\Delta = \begin{vmatrix} \int\limits_a^b f_1^2(t)\,dt & \int\limits_a^b f_1(t)f_2(t)\,dt & \ldots & \int\limits_a^b f_1(t)f_k(t)\,dt \\ \int\limits_a^b f_2(t)f_1(t)\,dt & \int\limits_a^b f_2^2(t)\,dt & \ldots & \int\limits_a^b f_2(t)f_k(t)\,dt \\ \ldots & \ldots & \ldots & \ldots \\ \int\limits_a^b f_k(t)f_1(t)\,dt & \int\limits_a^b f_k(t)f_2(t)\,dt & \ldots & \int\limits_a^b f_k^2(t)\,dt \end{vmatrix},$$

*) Для нас, конечно, несущественно, что размерность пространства равна k. В действительности пространство R может иметь произвольное (даже бесконечное) число измерений, поскольку наши рассуждения могут быть отнесены к подпространству, порожденному векторами $x, y, \ldots, w$.

и доказанная нами теорема означает:

Определитель Грама системы функций $\geqslant 0$. *Для линейной зависимости системы функций необходимо и достаточно, чтобы их определитель Грама был равен нулю.*

§ 7. Закон инерции

1. Закон инерции. Приводя квадратичную форму $A(x;x)$ к сумме квадратов, можно по-разному выбирать тот базис, в котором эта форма приводится к сумме квадратов, т.е. к виду

$$A(x;x) = \sum_{i=1}^{n} \lambda_i \xi_i^2. \tag{1}$$

Все те λ_i, которые отличны от нуля, можно, заменяя векторы базиса им пропорциональными, сделать равными ± 1. Таким образом, канонический вид формы $A(x;x)$ в некотором соответствующим образом подобранном базисе вполне можно характеризовать количеством коэффициентов, равных соответственно нулю, $+1$ и -1. Так как мы можем по-разному выбирать тот базис, в котором квадратичная форма записывается в виде суммы квадратов, то возникает вопрос, зависит ли количество коэффициентов, равных нулю, $+1$ и -1, от выбора базиса или же эти числа зависят лишь от квадратичной формы $A(x;x)$ (являются ее инвариантами).

Например, если квадратичная форма $A(x;x)$ в некотором базисе $e_1, e_2, \ldots, e_n$ имеет матрицу

$$\|a_{ik}\|,$$

где $a_{ik} = A(e_i; e_k)$ и все определители

$$\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \ldots,\ \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

отличны от нуля, то, как мы показали в п. 2 предыдущего параграфа, все λ_i в формуле (1) отличны от нуля и при приведении $A(x;x)$ к сумме квадратов по описанному там способу число отрицательных коэффициентов равно числу перемен знака в ряду определителей $1, \Delta_1, \Delta_2, \ldots, \Delta_n$.

Но мы могли взять другой исходный базис $e'_1, e'_2, \ldots$ $\ldots, e'_n$ (например, хотя бы взять те же самые векторы, но в другом порядке); при этом получается другая матрица $\|a'_{ik}\|$ и другие определители

$$\Delta'_1,\ \Delta'_2,\ \ldots,\ \Delta'_n,$$

и заранее совершенно неясно, почему число перемен знака в обоих случаях должно быть одно и то же.

В этом параграфе будет доказана следующая теорема, называемая *законом инерции квадратичной формы*:

Т е о р е м а 1. *Если квадратичная форма приведена двумя различными способами (т. е. в двух различных базисах) к сумме квадратов, то число положительных коэффициентов, так же как и число отрицательных, в обоих случаях одно и то же.*

Так как общее число коэффициентов λ_i в каноническом виде квадратичной формы равно n, то отсюда непосредственно следует, *что число коэффициентов λ_i, равных нулю, также есть инвариант квадратичной формы.*

Д о к а з а т е л ь с т в о. Пусть в базисе $e_1, e_2, \ldots, e_n$ квадратичная форма $A(x;x)$ имеет вид *)

$$A(x;x) = \xi_1^2 + \xi_2^2 + \ldots + \xi_p^2 - \xi_{p+1}^2 - \ldots - \xi_{p+q}^2; \qquad (2)$$

*) Коэффициенты λ_i в формуле (1) можно, как мы знаем, сделать равными ± 1 или 0. Те члены, для которых $\lambda_i = 0$, мы в формулах (2) и (3) опускаем.

при этом $\xi_1, \xi_2, \ldots, \xi_n$ — координаты вектора x, т. е.

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_p e_p + \xi_{p+1} e_{p+1} + \ldots \\ \ldots + \xi_{p+q} e_{p+q} + \ldots + \xi_n e_n.$$

Пусть в базисе $f_1, f_2, \ldots, f_n$ эта же квадратичная форма имеет вид

$$A(x; x) = \eta_1^2 + \eta_2^2 + \cdots + \eta_{p'}^2 - \eta_{p'+1}^2 - \ldots - \eta_{p'+q'}^2, \quad (3)$$

где $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора в базисе $f_1, f_2, \ldots, f_n$. Нам нужно доказать, что $p = p'$ и $q = q'$. Предположим, что это не так, например, пусть $p > p'$.

Рассмотрим подпространство R', состоящее из линейных комбинаций векторов $e_1, e_2, \ldots, e_p$. Оно имеет p измерений. Подпространство R'', состоящее из линейных комбинаций векторов $f_{p'+1}, f_{p'+2}, \ldots, f_n$, имеет $n - p'$ измерений. Так как $n - p' + p > n$ (ибо мы предположили, что $p > p'$), то существует вектор $x \neq 0$, лежащий на пересечении R' и R'', т. е. такой, что

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_p e_p$$

и

$$x = \eta_{p'+1} f_{p'+1} + \ldots + \eta_{p'+q'} f_{p'+q'} + \ldots + \eta_n f_n.$$

В базисе $e_1, e_2, \ldots, e_n$ этот вектор имеет координаты $\xi_1, \xi_2, \ldots, \xi_p, 0, \ldots, 0$, в базисе $f_1, f_2, \ldots, f_n$ он имеет координаты $0, 0, \ldots, 0, \eta_{p'+1}, \ldots, \eta_n$. Подставляя эти координаты в формулы (2) и (3), мы получим, с одной стороны,

$$A(x; x) = \xi_1^2 + \xi_2^2 + \ldots + \xi_p^2 > 0 \quad (4)$$

(так как не все числа $\xi_1, \xi_2, \ldots, \xi_p$ равны нулю), а с другой стороны,

$$A(x;x) = -\eta_{p'+1}^2 - \eta_{p'+2}^2 - \ldots - \eta_{p'+q'}^2 \leqslant 0 \;^{*}). \qquad (5)$$

Мы пришли к противоречию, следовательно, неравенство $p > p'$ невозможно. Точно так же доказывается невозможность неравенств $p < p'$, $q > q'$ и $q < q'$. Таким образом, закон инерции для квадратичных форм доказан.

2. Ранг квадратичной формы.

Определение 1. *Число отличных от нуля коэффициентов λ_i в каноническом виде квадратичной формы называется рангом квадратичной формы.*

Как уже было указано выше, из доказанного нами закона инерции непосредственно следует, что ранг квадратичной формы зависит только от самой формы, а не от способа ее приведения к каноническому виду. Посмотрим, как фактически найти ранг квадратичной формы. Для этого мы определим ранг квадратичной формы, не прибегая к ее каноническому виду. Попутно мы получим определение одного подпространства, тесно связанного с данной билинейной формой.

Определение 2. *Нулевым подпространством данной билинейной формы $A(x;y)$ мы называем совокупность R_0 векторов y, удовлетворяющих условию $A(x;y) = 0$ для любого вектора $x \in R$.*

Легко видеть, что R_0 действительно есть подпространство. В самом деле, пусть $y_1, y_2 \in R_0$, т. е. $A(x;y_1) = 0$ и $A(x;y_2) = 0$ для любого $x \in R$. Тогда $A(x;y_1+y_2) = 0$ и $A(x;\lambda y_1) = 0$ для любых x и λ, т. е. $y = y_1 + y_2 \in R_0$ и $\lambda y_1 \in R_0$.

---

*) В формуле (5) нельзя вместо знака $\leqslant$ поставить знак $<$, так как, хотя среди чисел $\eta_{p'+1}, \ldots, \eta_n$ есть отличные от нуля, но возможно, что $\eta_{p'+1} = \eta_{p'+2} = \ldots = \eta_{p'+q'} = 0$.

Поставим вопрос: как найти подпространство R_0? Пусть $f_1, f_2, \ldots, f_n$ — какой-либо базис в R. Для того чтобы вектор

$$y = \eta_1 f_1 + \eta_2 f_2 + \ldots + \eta_n f_n \tag{6}$$

принадлежал нулевому подпространству, достаточно, чтобы

$$A(f_i; y) = 0 \quad \text{для} \quad i = 1, 2, \ldots, n. \tag{7}$$

Действительно, если эти равенства выполнены, то и для любого x имеем $A(x; y) = 0$, так как всякий вектор x есть линейная комбинация базисных векторов.

Подставляя в (7) вместо y его выражение (6), мы приходим к следующей системе уравнений:

$$\begin{gathered}
A(f_1;\ \eta_1 f_1 + \eta_2 f_2 + \ldots + \eta_n f_n) = 0, \\
A(f_2;\ \eta_1 f_1 + \eta_2 f_2 + \ldots + \eta_n f_n) = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
A(f_n;\ \eta_1 f_1 + \eta_2 f_2 + \ldots + \eta_n f_n) = 0,
\end{gathered}$$

или, если положить $A(f_i; f_k) = a_{ik}$, к системе

$$\begin{gathered}
a_{11}\eta_1 + a_{12}\eta_2 + \ldots + a_{1n}\eta_n = 0, \\
a_{21}\eta_1 + a_{22}\eta_2 + \ldots + a_{2n}\eta_n = 0, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{n1}\eta_1 + a_{n2}\eta_2 + \ldots + a_{nn}\eta_n = 0.
\end{gathered}$$

Совокупность векторов y, координаты $\eta_1, \eta_2, \ldots, \eta_n$ которых являются решениями этой системы, и образует нулевое подпространство R_0. Как известно из теории линейных уравнений, размерность этого подпространства равна $n - r$, где r — ранг матрицы $\|a_{ik}\|$.

Мы можем теперь сделать следующий вывод:

Ранг матрицы $\|a_{ik}\|$ билинейной формы $A(x; y)$ в некотором базисе не зависит от выбора этого базиса (хотя сама матрица $\|a_{ik}\|$, как мы знаем из § 5, зависит от выбора базиса).

В самом деле, ранг этой матрицы равен $n - r_0$, где r_0 — размерность нулевого подпространства. Нулевое же подпространство ни от какой системы координат вообще не зависит.

Свяжем ранг матрицы квадратичной формы с рангом самой квадратичной формы. Рангом квадратичной формы мы назвали число отличных от нуля квадратов в каноническом виде квадратичной формы. Но в каноническом базисе матрица квадратичной формы имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

и ранг этой матрицы равен r, где r — число коэффициентов, отличных от нуля, т. е. равен рангу квадратичной формы. Так как ранг матрицы квадратичной формы, как мы доказали, не зависит от системы координат, то и в любой другой системе координат ранг матрицы квадратичной формы равен рангу самой квадратичной формы *).

Итак, нами доказана следующая

Т е о р е м а 2. *Матрицы квадратичной формы в различных системах координат имеют один и тот же ранг r. Этот ранг равен числу квадратов в каноническом виде формы, коэффициенты при которых отличны от нуля.*

Таким образом, для того чтобы найти ранг квадратичной формы, нужно вычислить ранг ее матрицы в какой-нибудь одной системе координат.

---

*) Пользуясь тем известным из теории матриц фактом, что ранг матрицы не меняется при умножении ее на любую неособенную матрицу, этот результат можно получить и непосредственно из выведенной в § 4 формулы преобразования матрицы билинейной формы при изменении базиса $\boldsymbol{B} = \boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$.

§ 8. Комплексное n-мерное пространство

Во всех предыдущих параграфах мы всюду, кроме тех случаев, когда это особо оговаривалось, имели дело с пространством над полем вещественных чисел. Ряд изложенных выше результатов справедлив для любого основного поля. Для дальнейшего особое значение, кроме пространства над полем вещественных чисел, будет иметь пространство над полем комплексных чисел. Разберем содержание предыдущих параграфов применительно к этому случаю.

1. Комплексное линейное пространство. Как указывалось в § 1, все изложенные там результаты справедливы для пространства над любым полем и, значит, в частности для пространства над полем комплексных чисел.

2. Комплексное евклидово пространство. Комплексным евклидовым пространством называется комплексное линейное пространство, в котором введено скалярное произведение, т. е. каждой паре векторов x и y поставлено в соответствие комплексное число (x, y), причем выполнены следующие аксиомы:

1° $(x, y) = \overline{(y, x)}$ (под $\overline{(y, x)}$ мы понимаем число, комплексно сопряженное с (y, x));

2° $(\lambda x, y) = \lambda(x, y)$;

3° $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$;

4° (x, x) есть вещественное неотрицательное число, равное нулю лишь при $x = 0$.

Из аксиом 1° и 2° следует, что $(x, \lambda y) = \overline{\lambda}(x, y)$. Действительно,

$$(x, \lambda y) = \overline{(\lambda y, x)} = \overline{\lambda}\,\overline{(y, x)} = \overline{\lambda}(x, y).$$

Далее, справедливо равенство $(x, y_1 + y_2) = (x, y_1) + (x, y_2)$. В самом деле,

$$(x, y_1 + y_2) = \overline{(y_1 + y_2, x)} = \overline{(y_1, x)} + \overline{(y_2, x)} = (x, y_1) + (x, y_2).$$

Аксиома 1° отличается от соответствующей аксиомы 1° для вещественного евклидова пространства; при переходе к комплексному пространству мы не могли бы сохранить аксиомы 1°, 2°, 4° вещественного евклидова пространства без изменений. В самом деле, если бы

$$(x, y) = (y, x),$$

то

$$(x, \lambda y) = \lambda(x, y).$$

Но тогда

$$(\lambda x, \lambda x) = \lambda^2(x, x);$$

следовательно, в частности

$$(ix, ix) = -(x, x),$$

т. е. числа (x, x) и (y, y), где $y = ix$, были бы разных знаков, что противоречит аксиоме 4°.

Примеры комплексных евклидовых пространств. 1. Вектором пространства R мы назовем систему n комплексных чисел. Сложение векторов и умножение их на числа определим обычным образом. Скалярное произведение векторов

$$x = (\xi_1, \xi_2, \ldots, \xi_n) \quad \text{и} \quad y = (\eta_1, \eta_2, \ldots, \eta_n)$$

зададим формулой

$$(x, y) = \xi_1\overline{\eta}_1 + \xi_2\overline{\eta}_2 + \ldots + \xi_n\overline{\eta}_n.$$

Мы предоставляем читателю проверить, что аксиомы 1°–4° выполнены. В частности, скалярное произведение вектора с самим собой задается формулой

$$(x, x) = \xi_1\overline{\xi}_1 + \xi_2\overline{\xi}_2 + \ldots + \xi_n\overline{\xi}_n = |\xi_1|^2 + |\xi_2|^2 + \ldots + |\xi_n|^2.$$

2. Векторы в пространстве R определим, как и в примере 1. Скалярное произведение задаем формулой

$$(x, y) = \sum_{i,k=1}^{n} a_{ik}\xi_i\overline{\eta}_k,$$

где a_{ik} — заданные комплексные числа, удовлетворяющие условиям:

α) $a_{ik} = \overline{a}_{ki}$,

β) $\sum a_{ik}\xi_i\overline{\xi}_k \geqslant 0$ для любых $\xi_1, \xi_2, \ldots, \xi_n$ и обращается в нуль лишь при $\xi_1 = \xi_2 = \ldots = \xi_n = 0$.

3. Векторами пространства R мы будем считать функции от t, заданные на отрезке $[a, b]$ и принимающие комплексные значения. Скалярное произведение двух таких функций определим формулой

$$(f(t), g(t)) = \int_a^b f(t)\overline{g(t)}\,dt.$$

Можно проверить, что все аксиомы скалярного произведения при этом выполнены.

Длиной вектора x назовем $\sqrt{(x, x)}$. Из аксиомы 4° следует, что длина вектора неотрицательна и обращается в нуль лишь для нулевого вектора. Так как скалярное произведение двух векторов, вообще говоря, комплексно, то мы не будем определять угла между векторами, а введем лишь понятие ортогональности двух векторов.

Векторы x и y называются ортогональными, если $(x, y) = 0$.

3. Ортогональный базис. Изоморфизм комплексных евклидовых пространств. Ортогональным базисом в n-мерном комплексном евклидовом пространстве называется совокупность n попарно ортогональных не равных нулю векторов $e_1, e_2, \ldots, e_n$. Так же, как в § 3, доказывается, что векторы $e_1, e_2, \ldots, e_n$ линейно независимы, т. е. образуют базис.

Существование ортогонального базиса в комплексном n-мерном евклидовом пространстве доказывается процессом ортогонализации, в точности совпадающим с приведенным в § 3.

Выразим скалярное произведение двух векторов x и y через их координаты $\xi_1, \xi_2, \ldots, \xi_n$ и $\eta_1, \eta_2, \ldots, \eta_n$ в ортогональном нормированном базисе. Мы имеем:

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n \quad \text{и} \quad y = \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n.$$

Тогда

$$(x, y) = (\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n,\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n) = \\ = \xi_1 \overline{\eta}_1 + \xi_2 \overline{\eta}_2 + \ldots + \xi_n \overline{\eta}_n.$$

(Ср. пример 1 этого параграфа.)

Выразим координаты ξ_i вектора x в ортогональном нормированном базисе через векторы базиса и сам вектор x. Имеем:

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

Умножая скалярно обе части равенства на e_i, получим:

$$(x, e_i) = \xi_1 (e_1, e_i) + \xi_2 (e_2, e_i) + \ldots + \xi_i (e_i, e_i) + \ldots + \xi_n (e_n, e_i)$$

или

$$(x, e_i) = \xi_i.$$

Так же, как и в § 3, доказывается, что все комплексные евклидовы пространства данного числа измерений n изоморфны между собой.

4. Билинейные и квадратичные формы. Все определения (линейной функции, квадратичной формы и т. д.), введенные в § 4 (за исключением понятия положительной определенности), имеют смысл для линейного пространства над любым полем, в том числе и над полем комплексных чисел. Однако в случае комплексного линейного пространства можно еще по-другому ввести эти понятия; для нас именно этот второй способ будет даже более существенным.

Линейные функции первого и второго рода. Функция, ставящая в соответствие каждому вектору комплексное число, называется *линейной функцией первого рода*, если она удовлетворяет следующим условиям:

$$1^\circ \qquad f(x+y) = f(x) + f(y),$$
$$2^\circ \qquad f(\lambda x) = \lambda f(x).$$

Это совпадает с определением линейной функции в § 4.

Линейной функцией второго рода называется функция, удовлетворяющая условиям

$$1^\circ \qquad f(x+y) = f(x) + f(y),$$
$$2^\circ \qquad f(\lambda x) = \overline{\lambda} f(x).$$

Так же, как и в § 4, можно доказать, что всякая линейная функция первого рода может быть записана в виде

$$f(x) = a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n,$$

где ξ_i — координаты вектора x в базисе $e_1, e_2, \ldots, e_n$, а a_i — постоянные, $a_i = f(e_i)$.

Всякая же линейная функция второго рода может быть записана в виде

$$f(x) = b_1\overline{\xi}_1 + b_2\overline{\xi}_2 + \ldots + b_n\overline{\xi}_n.$$

Очевидно, что если $f(x)$ — линейная функция первого рода, то $\overline{f(x)}$ — линейная функция второго рода.

Как было определено выше (п. 2, § 4), билинейной функцией называется функция двух векторов $A(x;y)$, линейная по каждому из аргументов. Наличие в комплексном пространстве двух типов линейных функций приводит к существованию целых четырех типов билинейных функций — линейных первого рода и по x и по y, первого рода по x и второго рода по y, второго рода по x и первого по y и второго рода по обоим аргументам. Но третий и четвертый типы комплексно сопряжены соответственно ко второму и первому, а билинейные функции первого типа определяются в комплексном пространстве буквально так же, как и в вещественном. Поэтому мы остановимся подробнее лишь на билинейных формах второго типа. Для краткости будем называть их просто билинейными. Итак, введем следующее определение:

Определение 1. *Будем говорить, что $A(x;y)$ есть билинейная функция (форма) от векторов x и y, если*

1° *при фиксированном y $A(x;y)$ есть линейная функция первого рода от x;*

2° *при фиксированном x $A(x;y)$ есть линейная функция второго рода от y.* Или, иначе:

1°
$$A(x_1+x_2;y) = A(x_1;y) + A(x_2;y),$$
$$A(\lambda x;y) = \lambda A(x;y),$$

2°
$$A(x;y_1+y_2) = A(x;y_1) + A(x;y_2),$$
$$A(x;\mu y) = \overline{\mu} A(x;y).$$

Примером билинейной функции является скалярное произведение в комплексном евклидовом пространстве
$$A(x;y) = (x,y),$$
рассматриваемое как функция векторов x и y. Другим примером билинейной формы в комплексном пространстве является выражение
$$A(x;y) = \sum_{i,k=1}^{n} a_{ik}\xi_i\overline{\eta}_k,$$
рассматриваемое как функция векторов
$$x = \xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n \quad \text{и} \quad y = \eta_1e_1+\eta_2e_2+\ldots+\eta_ne_n.$$
Легко проверить, что условия, определяющие билинейную функцию, при этом выполнены.

Пусть $e_1, e_2, \ldots, e_n$ — некоторый базис в n-мерном комплексном пространстве. Пусть $A(x;y)$ — билинейная форма, x и y можно записать в виде
$$x = \xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n, \quad y = \eta_1e_1+\eta_2e_2+\ldots+\eta_ne_n.$$

Тогда

$$A(x;y)=A(\xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n;\ \eta_1e_1+\eta_2e_2+\ldots+\eta_ne_n)=\\ =\sum_{i,k=1}^{n}\xi_i\overline{\eta}_kA(e_i;e_k).$$

Матрица $\|a_{ik}\|$ из чисел

$$a_{ik}=A(e_i;e_k)$$

называется *матрицей билинейной формы* $A(x;y)$ в базисе $e_1,e_2,\ldots,e_n$.

Если в билинейной форме $A(x;y)$ положить $y=x$, то получится функция $A(x;x)$, называемая *квадратичной формой* (в комплексном пространстве). Справедливо следующее утверждение:

Всякая билинейная форма однозначно определяется своей квадратичной формой *).

Доказательство. Пусть $A(x;x)$ — квадратичная форма, а x и y — произвольные векторы. Легко проверить, что имеет место тождество **)

$$A(x;y)=\frac{1}{4}\{A(x+y;x+y)+iA(x+iy;x+iy)-\\ -A(x-y;x-y)-iA(x-iy;x-iy)\}. \quad (1)$$

Выражение, стоящее справа в формуле (1), представляет собой комбинацию значений квадратичной формы для векторов $x+y$, $x-y$, $x+iy$ и $x-iy$. Слева стоит значение билинейной формы для произвольных векторов x и y. Таким образом, билинейная форма однозначно определяется своей квадратичной формой.

*) В отличие от определенных в § 4 форм в вещественном пространстве, для которых соответствующее утверждение справедливо лишь для симметрических билинейных форм.

**) Читатель должен помнить, что $A(x;\lambda y)=\overline{\lambda}A(x;y)$ и, следовательно, в частности, $A(x;iy)=-iA(x;y)$.

Определение 2. *Билинейная форма называется эрмитовой, если*

$$A(x;y) = \overline{A(y;x)}.$$

Это понятие является аналогом понятия симметрической билинейной формы в вещественном евклидовом пространстве.

Для того чтобы форма $A(x;y)$ была эрмитовой, необходимо и достаточно, чтобы ее матрица $\|a_{ik}\|$ в каком-либо базисе удовлетворяла условию

$$a_{ik} = \overline{a}_{ki}.$$

Действительно, если форма $A(x;y)$ эрмитова, то

$$a_{ik} = A(e_i;e_k) = \overline{A(e_k;e_i)} = \overline{a}_{ki}.$$

Обратно, если $a_{ik} = \overline{a}_{ki}$, то

$$A(x;y) = \sum a_{ik}\xi_i\overline{\eta}_k = \overline{\sum a_{ki}\eta_k\overline{\xi}_i} = \overline{A(y;x)}.$$

Замечание. Если в каком-либо базисе матрица билинейной формы удовлетворяет условию $a_{ik} = \overline{a}_{ki}$, то это же условие выполнено для матрицы этой билинейной формы и в любом другом базисе. В самом деле, если в каком-либо базисе равенство $a_{ik} = \overline{a}_{ki}$ имеет место, то $A(x;y)$ является эрмитовой билинейной формой; но тогда и в любом другом базисе $a_{ik} = \overline{a}_{ki}$.

Если билинейная форма эрмитова, то соответствующая ей квадратичная форма тоже называется эрмитовой.

Для того чтобы билинейная форма $A(x;y)$ была эрмитовой, необходимо и достаточно, чтобы $A(x;x)$ было вещественно для любого вектора x.

Доказательство. Пусть форма $A(x;y)$ эрмитова, т. е. $A(x;y) = \overline{A(y;x)}$. Тогда, полагая $x = y$, получаем:

$$A(x;x) = \overline{A(x;x)},$$

т.е. число $A(x;x)$ равно своему сопряженному и, значит, вещественно. Обратно, пусть $A(x;x)$ вещественно для любого вектора x. Тогда $A(x+y;x+y)$, $A(x+iy;x+iy)$, $A(x-y;x-y)$, $A(x-iy;x-iy)$ вещественны, и поэтому из формулы (1) непосредственно видно, что выражения $A(x;y)$ и $A(y;x)$ являются комплексно сопряженными.

С л е д с т в и е. *Квадратичная форма эрмитова в том и только в том случае, когда она принимает только вещественные значения.*

Действительно, только что было доказано, что для эрмитовости билинейной формы $A(x;y)$ необходимо и достаточно, чтобы $A(x;x)$ была вещественна для всех x.

Примером эрмитовой квадратичной формы является форма

$$A(x;x) = (x,x),$$

где (x,x) означает скалярное произведение вектора x с самим собой. Действительно, аксиомы 1°–3° скалярного произведения в комплексном евклидовом пространстве означают, что (x,y) есть эрмитова билинейная форма, и поэтому (x,x) есть эрмитова квадратичная форма.

Если, как и в § 4, назвать положительно определенной квадратичную форму, удовлетворяющую условию

$$A(x;x) > 0 \quad \text{при} \quad x \neq 0,$$

то комплексное евклидово пространство можно определить как комплексное линейное пространство, в котором задана положительно определенная эрмитова квадратичная форма.

Аналогично тому, как это сделано в вещественном пространстве, можно показать, что если $\boldsymbol{A}$ и $\boldsymbol{B}$ суть матрицы билинейной формы $A(x;y)$ соответственно в базисах $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_n$, то

$$\boldsymbol{B} = \boldsymbol{C}^* \boldsymbol{A} \boldsymbol{C},$$

где $\boldsymbol{C}$ — матрица перехода от базиса $e_1, e_2, \ldots, e_n$ к базису $f_1, f_2, \ldots, f_n$, а $\boldsymbol{C}^*$ — матрица, транспонированная и комплексно-сопряженная к матрице $\boldsymbol{C}$.

5. Приведение квадратичной формы к сумме квадратов.

Теорема 1. *Пусть $A(x; x)$ — эрмитова квадратичная форма в комплексном аффинном пространстве R. Тогда в R существует базис $e_1, e_2, \ldots, e_n$, в котором эта квадратичная форма имеет вид*

$$A(x; x) = \lambda_1 \xi_1 \overline{\xi}_1 + \lambda_2 \xi_2 \overline{\xi}_2 + \ldots + \lambda_n \xi_n \overline{\xi}_n,$$

где λ_i — вещественны.

Доказательство можно получить, перенося почти дословно доказательство соответствующей теоремы в вещественном пространстве.

Однако ввиду того, что в § 5 это доказательство изложено без уяснения его геометрической стороны, мы здесь вкратце повторим это доказательство в ином, более геометрическом, изложении. Для этого мы будем один за другим выбирать векторы того базиса, в котором форма приводится к сумме квадратов.

Выберем вектор e_i так, что $A(e_1; e_1) \neq 0$; это возможно, так как в противном случае мы имели бы $A(x; x) = 0$ для любого x, а следовательно, в силу формулы (1), и $A(x; y) \equiv 0$. В $(n-1)$-мерном пространстве $R^{(1)}$, состоящем из векторов x, удовлетворяющих условию $A(e_1; x) = 0$, выберем вектор e_2 такой, что $A(e_2; e_2) \neq 0$, и т. д. Этот процесс продолжим до тех пор, пока мы не придем к подпространству $R^{(r)}$, в котором $A(x; y) \equiv 0$ ($R^{(r)}$ может оказаться состоящем лишь из нуля). Если $R^{(r)}$ отлично от нуля, то выберем в нем произвольный базис $e_{r+1}, e_{r+2}, \ldots, e_n$. Вместе с построенными векторами $e_1, e_2, \ldots, e_r$ они образуют базис $e_1, e_2, \ldots, e_n$ всего R.

По построению

$$A(e_i; e_k) = 0 \quad \text{для} \quad i < k,$$

а значит, в силу эрмитовости формы $A(x; y)$,

$$A(e_i; e_k) = 0 \quad \text{и для} \quad i > k,$$

т. е.

$$A(e_i; e_k) = 0 \quad \text{для} \quad i \neq k.$$

Поэтому, если

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

— произвольный вектор, то

$$A(x; x) = \xi_1 \overline{\xi}_1 A(e_1; e_1) + \xi_2 \overline{\xi}_2 A(e_2; e_2) + \ldots + \xi_n \overline{\xi}_n A(e_n; e_n).$$

При этом числа $A(e_i; e_i)$ вещественны, как значения эрмитовой квадратичной формы. Обозначая $A(e_i; e_i)$ через λ_i, имеем:

$$A(x; x) = \lambda_1 \xi_1 \overline{\xi}_1 + \lambda_2 \xi_2 \overline{\xi}_2 + \ldots + \lambda_n \xi_n \overline{\xi}_n = \\ = \lambda_1 |\xi_1|^2 + \lambda_2 |\xi_2|^2 + \ldots + \lambda_n |\xi_n|^2.$$

6. Приведение эрмитовой квадратичной формы к сумме квадратов треугольным преобразованием. Пусть $A(x; x)$ — эрмитова квадратичная форма в комплексном линейном пространстве и $e_1, e_2, \ldots, e_n$ — базис. Мы будем предполагать, что определители

$$\Delta_1 = a_{11}, \ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ \ldots, \ \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

где $a_{ik} = A(e_i; e_k)$, отличны от нуля. Тогда, так же как и в § 6, мы можем написать формулу для нахождения базисов, в которых квадратичная форма приводится к сумме квадратов. Эти формулы в точности совпадают с формулами (3) и (6) § 6. При этом сама квадратичная форма в новом базисе имеет вид

$$A(x;x) = \frac{\Delta_0}{\Delta_1}|\xi_1|^2 + \frac{\Delta_1}{\Delta_2}|\xi_2|^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n}|\xi_n|^2, \qquad (2)$$

где $\Delta_0 = 1$. Отсюда, в частности, следует, что определители $\Delta_1, \Delta_2, \ldots, \Delta_n$ вещественны; действительно, если эрмитова квадратичная форма приведена к каноническому виду (2), то коэффициенты λ_i равны $A(e_i;e_i)$ и вещественны.

Упражнение. Доказать непосредственно, что если квадратичная форма $A(x;x)$ эрмитова, то определители $\Delta_1, \Delta_2, \ldots, \Delta_n$ вещественны.

Так же, как и в § 6, мы получаем, что *для того чтобы эрмитова квадратичная форма $A(x;x)$ была положительно определенной, необходимо и достаточно, чтобы построенные по ней определители $\Delta_1, \Delta_2, \ldots, \Delta_n$ были положительны.*

Число отрицательных коэффициентов при квадратах в каноническом виде эрмитовой квадратичной формы равно числу перемен знака в последовательности определителей

$$1,\ \Delta_1,\ \Delta_2,\ \ldots,\ \Delta_n.$$

7. Закон инерции. Имеет место теорема, доказательство которой ничем не отличается от доказательства соответствующей теоремы в § 7.

Теорема 2. *Если эрмитова квадратичная форма имеет в двух базисах канонический вид, то число положительных, отрицательных и нулевых коэффициентов в обоих случаях одно и то же.*

Понятие ранга квадратичной формы, введенное нами в § 7 для случая вещественного пространства, переносится без изменений и на комплексный случай.

ГЛАВА II

ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ

§ 9. Линейные преобразования и операции над ними

1. Основные определения. В предыдущей главе мы изучали функции в n-мерном линейном пространстве, принимающие численные значения (линейные функции, квадратичные и т. д.). Но в ряде случаев возникает потребность рассматривать функции другого вида, а именно, функции, которые точкам пространства ставят в соответствие снова точки того же пространства (а не числа). Простейшими среди функций такого рода являются линейные преобразования.

Определение 1. *Пусть каждому вектору x n-мерного пространства поставлен в соответствие вектор y этого же пространства. Функцию $y = A(x)$ мы назовем преобразованием пространства R.*

Преобразование A называется линейным, если выполнены следующие условия:

$$1^\circ \qquad A(x_1 + x_2) = A(x_1) + A(x_2),$$
$$2^\circ \qquad A(\lambda x) = \lambda A(x).$$

Там, где это не сможет привести к недоразумениям, вместо $A(x)$ мы будем писать Ax.

Примеры. 1. Рассмотрим трехмерное евклидово пространство R и в нем преобразование, состоящее в повороте R вокруг какой-либо оси, проходящей через нуль. Каждому вектору x ставится в соответствие вектор Ax, полученный из него данным поворотом. Условия 1° и 2° проверяются без труда. Проверим, например, условие 1°. $A(x_1 + x_2)$ означает, что векторы x_1 и x_2 сначала складываются, а затем полученный вектор поворачивается. $Ax_1 + Ax_2$ означает, что векторы x_1 и x_2 сперва поворачиваются, а затем складываются. Ясно, что в обоих случаях результат один и тот же.

2. Пусть R' — некоторая плоскость в трехмерном пространстве R, проходящая через нуль. Поставим в соответствие каждому вектору x его проекцию $x' = Ax$ на эту плоскость. Условия 1° и 2° опять легко проверяются. Например, 1° означает, что проекция суммы равна сумме проекций.

3. Рассмотрим аффинное n-мерное пространство, в котором вектор определен как совокупность n чисел.

Пусть $\|a_{ik}\|$ — некоторая матрица. Поставим в соответствие каждому вектору

$$x = (\xi_1, \xi_2, \ldots, \xi_n)$$

вектор

$$y = Ax = (\eta_1, \eta_2, \ldots, \eta_n),$$

где η_i вычисляются по формулам

$$\eta_i = \sum_{k=1}^{n} a_{ik}\xi_k.$$

Условия 1° и 2°, определяющие линейное преобразование, проверяются без труда.

4. Рассмотрим n-мерное пространство, элементами которого являются многочлены степени $\leqslant n-1$.

Положим

$$AP(t) = P'(t),$$

где $P'(t)$ — производная многочлена $P(t)$.

Это преобразование — линейное. Действительно,

1° $$(P_1(t) + P_2(t))' = P_1'(t) + P_2'(t),$$

2° $$(\lambda P(t))' = \lambda P'(t).$$

5. Рассмотрим пространство, в котором векторами являются непрерывные функции $f(t)$, $0 \leqslant t \leqslant 1$. Положим

$$Af(t) = \int\limits_0^t f(\tau)\, d\tau.$$

Преобразование A — линейное. Действительно,

1° $$A(f_1 + f_2) = \int\limits_0^t [f_1(\tau) + f_2(\tau)]\, d\tau = \int\limits_0^t f_1(\tau)\, d\tau + \int\limits_0^t f_2(\tau)\, d\tau = Af_1 + Af_2,$$

2° $$A(\lambda f) = \int\limits_0^t \lambda f(\tau)\, d\tau = \lambda \int\limits_0^t f(\tau)\, d\tau = \lambda Af.$$

6. Рассмотрим то же пространство, что и в примере 5. Пусть $k(t,s)$ — непрерывная функция, заданная в квадрате $0 \leqslant t \leqslant 1$, $0 \leqslant s \leqslant 1$. Положим

$$\varphi(t) \equiv Af(t) = \int\limits_0^1 k(t,s) f(s)\, ds.$$

Проверьте сами, что это преобразование линейно.

Среди линейных преобразований особую роль играют следующие простые преобразования:

единичное преобразование E, ставящее в соответствие каждому вектору этот же самый вектор, т. е.

$$Ex = x;$$

нулевое преобразование O, ставящее в соответствие каждому вектору x нулевой вектор:

$$Ox = 0.$$

2. Связь между матрицами и линейными преобразованиями. Пусть $e_1, e_2, \ldots, e_n$ — некоторый базис в n-мерном пространстве R и A — линейное преобразование в R.

Для любых n векторов $g_1, g_2, \ldots, g_n$ существует одно и только одно линейное преобразование A, такое что

$$Ae_1 = g_1, \quad Ae_2 = g_2, \quad \ldots, \quad Ae_n = g_n.$$

Докажем это. Покажем сначала, что преобразование A однозначно определяется векторами $Ae_1, Ae_2, \ldots$ $\ldots, Ae_n$. Действительно, пусть

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n \tag{1}$$

— произвольный вектор из R. Тогда

$$\begin{aligned} Ax = A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) &= \\ = \xi_1 Ae_1 + \xi_2 Ae_2 + \ldots + \xi_n Ae_n& \end{aligned} \tag{2}$$

и, следовательно, Ax однозначно определяется по Ae_1, $Ae_2, \ldots, Ae_n$.

Теперь покажем, что для всяких векторов g_1, $g_2, \ldots, g_n$ существует линейное преобразование A, такое, что $Ae_i = g_i$. Для этого поставим в соответствие векторам e_i векторы g_i; произвольному же вектору $x = \xi_1 e_1 + \ldots + \xi_n e_n$ поставим в соответствие вектор $\xi_1 g_1 + \ldots + \xi_n g_n$. Так как вектор x выражается через e_1 однозначно, то ему ставится в соответствие вполне определенный вектор Ax. Легко проверить, что так определенное преобразование A линейно.

Обозначим координаты вектора g_k в базисе $e_1, e_2, \ldots, e_n$ через $a_{1k}, a_{2k}, \ldots, a_{nk}$, т. е. положим

$$g_k = Ae_k = \sum_{i=1}^{n} a_{ik} e_i. \tag{3}$$

Совокупность чисел a_{ik} $(i, k = 1, 2, \ldots, n)$ образует матрицу

$$\boldsymbol{A} = \|a_{ik}\|,$$

которую мы назовем *матрицей линейного преобразования A в базисе $e_1, e_2, \ldots, e_n$.*

Итак, мы доказали, что при *заданном базисе $e_1, e_2, \ldots, e_n$ каждому линейному преобразованию A однозначно соответствует матрица $\|a_{ik}\|$,* и, обратно, *каждой матрице $\|a_{ik}\|$ однозначно отвечает линейное преобразование, определяемое формулами* (3), (1), (2).

Мы видим, таким образом, что линейные преобразования можно описывать с помощью матриц и матрицы являются тем аналитическим аппаратом, с помощью которого изучаются линейные преобразования в конечномерных пространствах.

Заметим, что при изменении базиса матрица, соответствующая данному линейному преобразованию, вообще говоря, изменится.

Примеры. 1. Пусть R — трехмерное пространство, A — линейное преобразование, состоящее в проектировании каждого вектора на плоскость XY. Примем за базис единичные векторы e_1, e_2, e_3, направленные по осям координат. Тогда

$$Ae_1 = e_1, \quad Ae_2 = e_2, \quad Ae_3 = 0,$$

т. е. матрица преобразования A в этом базисе имеет вид

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Упражнение. Найти матрицу того же преобразования в базисе e_1', e_2', e_3', где

$$e_1' = e_1, \quad e_2' = e_2, \quad e_3' = e_1 + e_2 + e_3.$$

2. Пусть E — единичное преобразование и $e_1, e_2, \ldots$ $\ldots e_n$ — базис в R. Тогда

$$Ae_i = e_i \quad (i = 1, 2, \ldots, n),$$

т. е. матрица единичного преобразования в любом базисе имеет вид

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Легко также видеть, что матрица нулевого преобразования в любом базисе состоит сплошь из нулей.

3. Пусть R — пространство многочленов степени $\leqslant n-1$. Преобразование A — дифференцирование, т. е.

$$AP(t) = P'(t).$$

Выберем в R базис следующим образом:

$$e_1 = 1, \quad e_2 = t, \quad e_3 = \frac{t^2}{2!}, \quad \ldots, \quad e_n = \frac{t^{n-1}}{(n-1)!}.$$

Тогда

$$Ae_1 = 1' = 0, \; Ae_2 = t' = 1, \; Ae_3 = \left(\frac{t^2}{2}\right)' = t = e_2, \; \ldots$$
$$\ldots, \; Ae_n = \left(\frac{t^{n-1}}{(n-1)!}\right)' = \frac{t^{n-2}}{(n-2)!} = e_{n-1}.$$

Таким образом, матрица преобразования A в этом базисе имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Пусть A — линейное преобразование, $e_1, e_2, \ldots, e_n$ — базис в R и $\|a_{ik}\|$ — матрица преобразования A в этом базисе. Пусть

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n, \tag{4}$$

$$Ax = \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_e. \tag{4'}$$

Найдем выражение координат $\eta_1, \eta_2, \ldots, \eta_n$ вектора Ax через координаты $\xi_1, \xi_2, \ldots, \xi_n$ вектора x. Имеем:

$$\begin{aligned} Ax = A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) = \\ &= \xi_1(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n) + \\ &+ \xi_2(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n) + \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ \xi_n(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n) = \\ &= (a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n)e_1 + \\ &+ (a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n)e_2 + \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ (a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n)e_n. \end{aligned}$$

Следовательно, сравнивая с (4′), получаем:

$$\begin{aligned} \eta_1 &= a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n, \\ \eta_2 &= a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \eta_n &= a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n, \end{aligned}$$

или короче:

$$\eta_i = \sum_{k=1}^{n} a_{ik}\xi_k. \tag{5}$$

Таким образом: *если линейное преобразование A имеет в данном базисе $e_1, e_2, \ldots, e_n$ матрицу $\|a_{ik}\|$, то базисные векторы преобразуются с помощью столбцов этой матрицы [формула (3)], а координаты произвольного вектора — с помощью ее строк [формула (5)].*

3. Сложение и умножение линейных преобразований. Линейные преобразования можно складывать и умножать.

Определение 2. *Произведением линейных преобразований A и B называется преобразование C, состоящее в последовательном выполнении сначала преобразования B, а затем преобразования A.*

Другими словами: $C = AB$ означает, что для любого x $Cx = A(Bx)$.

Произведение линейных преобразований есть линейное преобразование, т.е. удовлетворяет условиям 1° и 2° определения 1. Действительно,

$$C(x_1 + x_2) = A[B(x_1 + x_2)] = A(Bx_1 + Bx_2) = \\ = ABx_1 + ABx_2 = Cx_1 + Cx_2.$$

Первое равенство написано на основании определения произведения, второе на основании свойства 1° для B, третье в силу того же свойства для A и, наконец, четвертое опять-таки в силу определения произведения.

Аналогично показывается, что $C(\lambda x) = \lambda Cx$.

Если E — единичное преобразование, а A — произвольное, то легко проверить, что

$$AE = EA = A.$$

Как обычно, определяем степени преобразования A:

$$A^2 = A \cdot A, \quad A^3 = A^2 \cdot A, \quad \ldots \quad \text{и т. д.}$$

Как и для чисел, полагаем, по определению, $A^0 = E$. Очевидно, что

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

П р и м е р. R — пространство многочленов $P(t)$ степени не выше $n-1$. Определим в нем преобразование D формулой

$$DP(t) = P'(t),$$

где $P'(t)$ — производная многочлена $P(t)$. Тогда для любого $P(t)$

$$D^2P(t) = D(DP(t)) = (P'(t))' = P''(t).$$

Это равенство определяет преобразование D^2. Аналогично можно определить преобразование $D^3P(t) = P'''(t), \ldots$ Заметим, что в данном случае $D^n = 0$. Действительно, так как векторами пространства являются многочлены степени $\leqslant n-1$, то

$$D^nP(t) = P^{(n)}(t) = 0.$$

У п р а ж н е н и е. Выберем в пространстве многочленов степени не выше, чем $n-1$, базис, указанный в примере 3 п. 2 этого параграфа. Найти в этом базисе матрицы преобразований $D, D^2, D^3, \ldots$

Мы знаем, что при заданном базисе $e_1, e_2, \ldots, e_n$ каждому линейному преобразованию отвечает матрица. Пусть преобразованию A отвечает матрица $\|a_{ik}\|$, преобразованию B — матрица $\|b_{ik}\|$; найдем матрицу $\|c_{ik}\|$, отвечающую преобразованию $C = AB$. В силу определения матрицы преобразования C мы имеем:

$$Ce_k = \sum_i c_{ik}e_i. \tag{6}$$

Далее,

$$ABe_k = A\left(\sum_{j=1}^n b_{jk}e_j\right) = \sum_j b_{jk}Ae_j = \sum_{j,i} b_{jk}a_{ij}e_i. \tag{7}$$

Сравнивая коэффициенты при e_i в равенствах (6) и (7), получаем:

$$c_{ik} = \sum_j a_{ij} b_{jk}. \tag{8}$$

Мы видим, что элемент c_{ik} матрицы $\boldsymbol{C}$ есть сумма произведений элементов i-й строки матрицы $\boldsymbol{A}$ на соответствующие элементы k-го столбца матрицы $\boldsymbol{B}$. Так определенная матрица называется произведением матрицы $\boldsymbol{A}$ на матрицу $\boldsymbol{B}$. Итак, если преобразованию A отвечает матрица $\|a_{ik}\|$, а преобразованию B — матрица $\|b_{ik}\|$, то произведению этих преобразований отвечает матрица $\|c_{ik}\|$, являющаяся произведением матриц $\|a_{ik}\|$ и $\|b_{ik}\|$. Произведение матриц вычисляется по формуле (8).

Определение 3. *Суммой линейных преобразований A и B называется такое преобразование C, которое каждому вектору x ставит в соответствие вектор $Ax + Bx$*; иначе говоря, $C = A + B$ означает, что $Cx = Ax + Bx$ для любого x.

Пусть преобразование C есть сумма преобразований A и B. Тогда, зная матрицы преобразований A и B, легко найти матрицу преобразования C. Действительно, пусть $\|a_{ik}\|$, соответственно $\|b_{ik}\|$, суть матрицы преобразования A, соответственно B, т. е.

$$Ae_k = \sum_i a_{ik} e_i, \quad Be_k = \sum_i b_{ik} e_i,$$

и $\|c_{ik}\|$ — матрица преобразования C, т. е.

$$Ce_k = \sum_i c_{ik} e_i.$$

Так как $C = A + B$, то

$$Ce_k = Ae_k + Be_k = \sum_i (a_{ik} + b_{ik}) e_i$$

и, следовательно,

$$c_{ik} = a_{ik} + b_{ik}.$$

Матрица $\|a_{ik} + b_{ik}\|$ называется суммой матриц $\|a_{ik}\|$ и $\|b_{ik}\|$. Итак: *матрица суммы линейных преобразований равна сумме матриц, соответствующих отдельным слагаемым.*

Операции сложения и умножения линейных преобразований удовлетворяют обычным для сложения и умножения условиям, а именно:

1° $$A + B = B + A;$$

2° $$(A + B) + C = A + (B + C);$$

3° $$A(BC) = (AB)C;$$

4° $$\begin{cases} (A + B)C = AC + BC, \\ C(A + B) = CA + CB. \end{cases}$$

Заметим, что умножение линейных преобразований, вообще говоря, некоммутативно. Действительно, возьмем линейное преобразование A с матрицей $\begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix}$ и линейное преобразование B с матрицей $\begin{Vmatrix} 1 & 0 \\ 1 & 1 \end{Vmatrix}$. Так как

$$\begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix} \begin{Vmatrix} 1 & 0 \\ 1 & 1 \end{Vmatrix} = \begin{Vmatrix} 2 & 1 \\ 1 & 1 \end{Vmatrix},$$

а

$$\begin{Vmatrix} 1 & 0 \\ 1 & 1 \end{Vmatrix} \begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix} = \begin{Vmatrix} 1 & 1 \\ 1 & 2 \end{Vmatrix},$$

то

$$AB \neq BA.$$

Мы могли бы без большого труда доказать равенства 1°–4° непосредственно. Но в этом нет необходимости.

В самом деле, между линейными преобразованиями и матрицами установлено взаимно однозначное соответствие, причем сумме соответствует сумма, а произведению — произведение. Для матриц формулы 1°–4° доказываются в курсе алгебры; в силу установленного соответствия они автоматически переносятся на линейные преобразования.

Определим еще произведение линейного преобразования A на число λ; под преобразованием λA мы будем понимать преобразование, которое каждому вектору x ставит в соответствие вектор $\lambda(Ax)$. Ясно, что если линейному преобразованию A отвечает матрица $\|a_{ik}\|$, то преобразованию λA отвечает матрица $\|\lambda a_{ik}\|$.

Упражнение. Проверить, что множество всех линейных преобразований пространства R с введенными здесь операциями сложения и умножения на число образуют линейное пространство. Какова его размерность?

Умея находить сумму и произведение линейных преобразований, можно теперь найти любой многочлен от преобразования A. Пусть $P(t) = a_0t^m + a_1t^{m-1} + \ldots$ $\ldots + a_m$ — произвольный многочлен. Тогда под $P(A)$ мы понимаем линейное преобразование, определенное формулой

$$P(A) = a_0A^m + a_1A^{m-1} + \ldots + a_mE.$$

Пример. Рассмотрим пространство R, элементами которого являются функции, определенные на интервале (a, b) и имеющие производные всех порядков. В этом пространстве рассмотрим линейное преобразование D, которое каждой функции ставит в соответствие ее производную, т. е.

$$Df(t) = f'(t).$$

Пусть теперь нам задан многочлен $P(t) = a_0t^m + a_1t^{m-1} + \ldots + a_m$. Тогда $P(D)$ есть линейное преобразование, которое переводит функцию $f(t)$ в

$$P(D)f(t) = a_0f^{(m)}(t) + a_1f^{(m-1)}(t) + \ldots + a_mf(t).$$

В соответствии с введенными выше определениями сложения и умножения матриц многочлен от матрицы $\boldsymbol{A}$ мы задаем формулой

$$P(\boldsymbol{A}) = a_0\boldsymbol{A}^m + a_1\boldsymbol{A}^{m-1} + \ldots + a_m\boldsymbol{E}.$$

П р и м е р. Предположим, что $\boldsymbol{A}$ — так называемая диагональная матрица, т. е. матрица, у которой на всех местах, кроме главной диагонали, стоят нули. Найдем $p(\boldsymbol{A})$. Мы имеем

$$\boldsymbol{A} = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix};$$

тогда

$$\boldsymbol{A}^2 = \begin{pmatrix} \lambda_1^2 & 0 & \ldots & 0 \\ 0 & \lambda_2^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n^2 \end{pmatrix}, \quad \ldots, \quad \boldsymbol{A}^m = \begin{pmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n^m \end{pmatrix}.$$

Отсюда следует, что если $P(t) = a_0t^m + \ldots + a_{m-1}t + a_m$, то

$$P(\boldsymbol{A}) = \begin{pmatrix} P(\lambda_1) & 0 & \ldots & 0 \\ 0 & P(\lambda_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & P(\lambda_n) \end{pmatrix}.$$

У п р а ж н е н и е. Пусть матрица $\boldsymbol{A}$ имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Найти $P(\boldsymbol{A})$.

Можно определить не только многочлен от матрицы, но и вообще функцию от матрицы, например $e^{\boldsymbol{A}}$, $\sin\boldsymbol{A}$, и т. д.

Совокупность матриц n-го порядка образует, как мы уже упоминали в § 1 (стр. 9, пример 5), линейное пространство, если определить сумму и произведение матриц на число, как обычно. Это пространство имеет

n^2 измерений (каждая матрица задается системой n^2 чисел); поэтому всякие n^2+1 матриц линейно зависимы. Рассмотрим последовательность степеней некоторой матрицы $\boldsymbol{A}$:

$$\boldsymbol{E},\ \boldsymbol{A},\ \boldsymbol{A}^2,\ \ldots,\ \boldsymbol{A}^{n^2}.$$

Так как их n^2+1, то они линейно зависимы, т. е. существуют числа $a_0, a_1, a_2, \ldots, a_{n^2}$ такие, что

$$a_0\boldsymbol{E} + a_1\boldsymbol{A} + a_2\boldsymbol{A}^2 + \ldots + a_{n^2}\boldsymbol{A}^{n^2} = 0.$$

Мы получили следующий интересный вывод: для каждой матрицы порядка n существует многочлен степени n^2 такой, что $P(\boldsymbol{A}) = 0$. Указанный здесь очень простой вывод существования многочлена $P(t)$, для которого $P(\boldsymbol{A}) = 0$, обладает двумя недостатками. Во-первых, не указан способ вычисления такого многочлена и, во-вторых, степень такого многочлена завышена. В действительности мы несколько позже покажем, что для каждой матрицы $\boldsymbol{A}$ существует многочлен степени n, очень просто связанный с матрицей и обращающийся в нуль при подстановке в него этой матрицы.

4. Обратное преобразование. Ядро и образ преобразования.

Определение 4. *Преобразование B называется обратным к A, если $AB = BA = E$, где E — единичное преобразование.*

В силу определения E это означает, что для любого x $B(Ax) = x$, т. е. если A переводит x в вектор Ax, то обратное преобразование B переводит вектор Ax обратно в вектор x. Преобразование, обратное преобразованию A, обозначается A^{-1}.

Не для всякого преобразования существует обратное. Например, преобразование, состоящее в проектировании трехмерного пространства на плоскость XY (см. пример 1 п. 1), очевидно, не имеет обратного.

С понятием обратного преобразования связано понятие обратной матрицы. Как известно, для каждой матрицы $\boldsymbol{A}$, удовлетворяющей условию $\mathrm{Det}(\boldsymbol{A}) \neq 0$, можно определить матрицу $\boldsymbol{A}^{-1}$, удовлетворяющую условию

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}. \tag{9}$$

Эта матрица $\boldsymbol{A}^{-1}$ называется обратной к матрице $\boldsymbol{A}$. Ее можно найти, решая систему линейных уравнений, эквивалентную матричному равенству (9). Элементы ее k-го столбца окажутся равными минорам k-й строки матрицы $\boldsymbol{A}$, деленным на ее определитель. Легко проверить, что так составленная матрица $\boldsymbol{A}^{-1}$ удовлетворяет условиям (9).

Так как при заданном базисе между матрицами и линейными преобразованиями имеется взаимно однозначное соответствие, сохраняющее операцию умножения, то, *для того чтобы преобразование A имело обратное, необходимо и достаточно, чтобы его матрица в каком-нибудь базисе имела определитель, отличный от нуля, т. е. имела бы ранг n*. Преобразование, имеющее обратное, называют *невырожденным.*

С произвольным линейным преобразованием A связаны два важных подпространства — *ядро* и *образ* этого преобразования.

Определение 5. *Совокупность M векторов вида Ax, где x пробегает все R, называется образом пространства R при преобразовании A.*

Другими словами, образ пространства — это множество тех векторов y, для которых уравнение $Ax = y$ имеет хотя бы одно решение. Ясно, что у обратимого преобразования образ есть все пространство.

Покажем, что *M есть подпространство пространства R.* Действительно, пусть $y_1 \in M$ и $y_2 \in M$.

Это значит, что существуют x_1 и x_2 такие, что $y_1 = Ax_1$ и $y_2 = Ax_2$. Но тогда $y_1 + y_2 = Ax_1 + Ax_2 =$

$= A(x_1 + x_2)$ и, значит, $y_1 + y_2 \in M$. Аналогично, если $y = Ax$, то $\lambda y = \lambda Ax = A\lambda x$, т. е. $\lambda y \in M$.

Следовательно, M является подпространством. Размерность этого подпространства называется рангом преобразования A.

П р и м е р. Рассмотрим преобразование A, состоящее в проектировании трехмерного пространства R в плоскость XY (пример 1, п. 2). Очевидно, что образ этого преобразования есть плоскость XY.

У п р а ж н е н и е. Написать матрицу произвольного преобразования A в базисе, первые k векторов которого являются базисом в образе пространства при этом преобразовании.

Другим важным подпространством является ядро преобразования A, состоящее из всех векторов, переходящих при этом преобразовании в нуль.

О п р е д е л е н и е 6. *Совокупность N векторов x таких, что $Ax = 0$, называется ядром преобразования A.*

Ясно, что ядро также есть подпространство пространства R. Действительно, если $Ax_1 = 0$ и $Ax_2 = 0$, то $A(x_1 + x_2) = Ax_1 + Ax_2 = 0$. Точно так же, если $Ax = 0$, то $A\lambda x = \lambda Ax = 0$, т. е. N есть подпространство.

Очевидно, что если A — невырожденное преобразование, то его ядро состоит из нуля (т. е. система однородных уравнений с отличным от нуля определителем имеет только нулевое решение).

У п р а ж н е н и е. Написать матрицу преобразования A в базисе, первые k векторов которого есть базис ядра.

П р и м е р. Пусть R — пространство многочленов степени $\leqslant n - 1$ и преобразование A — дифференцирование, т. е.

$$AP(t) = P'(t).$$

Ядро этого преобразования состоит из многочленов $P(t)$, для которых $P'(t) = 0$, т.е. из констант. Таким образом, ядро N здесь одномерно.

Образ A состоит из многочленов вида $P'(t)$, где $P(t)$ имеет степень $\leqslant n-1$, т.е. M состоит из всех многочленов степени $\leqslant n-2$. Размерность M равна $n-1$.

Рассмотрим теперь преобразование A^2, которое задается формулой

$$A^2 P(t) = P''(t).$$

Для преобразования A^2 ядро N состоит из всех многочленов не выше первой степени, а образ из всех многочленов степени $\leqslant n-3$ (проверьте!), т.е. N двумерно, а M имеет размерность $n-2$.

Аналогично у преобразования A^3 ядро трехмерно, а образ имеет размерность $n-3$ и т. д.

Наконец, преобразование A^n в этом случае есть нулевое преобразование. Его ядро $N = R$, а образ состоит только из нуля.

На этом примере видно, что при возведении преобразования в степень его ядро расширяется, а образ, наоборот, уменьшается. При этом размерность ядра как бы характеризует степень вырожденности преобразования. Чем больше ядро, тем меньше образ и тем «более вырожденным» является преобразование. Крайними случаями являются нулевое преобразование, ядром которого является все R, а образ равен нулю, и, с другой стороны, обратимое преобразование, образом которого является все пространство, а ядро равно нулю.

При этом сумма размерностей ядра и образа всегда остается равной размерности всего пространства.

Имеет место общая теорема.

Теорема. *Пусть A — произвольное линейное преобразование n-мерного пространства R. Сумма* *размерностей ядра и образа преобразования A равна размерности всего пространства.*

Доказательство. Предположим, что ядро N преобразования A имеет размерность k. Выберем в N базис из векторов $e_1, \ldots, e_k$ и дополним его до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ во всем пространстве R.

Рассмотрим векторы $Ae_{k+1}, \ldots, Ae_n$. Множество линейных комбинаций этих векторов образует подпространство, которое совпадает с M — образом преобразования A.

Действительно, пусть y — произвольный вектор из M. Тогда, по определению, существует вектор x такой, что $y = Ax$. Так как $e_1, \ldots, e_n$ — базис в R, то $x = \gamma_1 e_1 + \ldots + \gamma_n e_n$. Но так как $Ae_1 = \ldots = Ae_k = 0$ ($e_1, \ldots, e_k$ — базис в ядре), то $y = Ax = \gamma_{k+1} Ae_{k+1} + \ldots + \gamma_n Ae_n$.

Покажем, что $n-k$ векторов $Ae_{k+1}, \ldots, Ae_n$ линейно независимы.

Действительно, пусть существуют числа α_j, не равные одновременно нулю и такие, что $\alpha_1 Ae_{k+1} + \ldots + \alpha_{n-k} Ae_n = 0$. Рассмотрим вектор $x = \alpha_1 e_{k+1} + \ldots + \alpha_{n-k} e_n$. Тогда $Ax = A(\alpha_1 e_{k+1} + \ldots + \alpha_{n-k} e_n) = \alpha_1 Ae_{k+1} + \ldots + \alpha_{n-k} Ae_n = 0$, т. е. x принадлежит ядру. Мы пришли к противоречию, поскольку, с одной стороны, x как элемент ядра представим как линейная комбинация первых k базисных векторов, а, с другой стороны, $x = \alpha_1 e_{k+1} + \ldots + \alpha_{n-k} e_n$ был задан как линейная комбинация $e_{k+1}, \ldots, e_n$. Это противоречит единственности представления вектора x через векторы базиса. Следовательно, векторы $Ae_{k+1}, \ldots, Ae_n$ линейно независимы.

Мы показали, что существует $n-k$ линейно независимых векторов таких, что любой вектор образа есть их линейная комбинация, т. е. размерность образа равна $n-k$, что и требовалось доказать.

5. Связь между матрицами линейного преобразования в различных базисах. Одно и то же линейное преобразование может в различных базисах иметь различные матрицы (см., например, упражнение к примеру 1 п. 3 этого параграфа). Выясним, как изменяется матрица линейного преобразования A при переходе от одного базиса к другому.

Пусть в R даны два базиса: $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots \ldots, f_n$. Матрицу перехода от базиса $e_1, e_2, \ldots, e_n$ к базису $f_1, f_2, \ldots, f_n$ обозначим через $\boldsymbol{C}$, т. е. положим

$$\left.\begin{array}{l} f_1 = c_{11}e_1 + c_{21}e_2 + \ldots + c_{n1}e_n, \\ f_2 = c_{12}e_1 + c_{22}e_2 + \ldots + c_{n2}e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n = c_{1n}e_1 + c_{2n}e_2 + \ldots + c_{nn}e_n. \end{array}\right\} \tag{10}$$

Введем вспомогательное линейное преобразование C, положив

$$Ce_i = f_i.$$

Его матрица в базисе $e_1, e_2, \ldots, e_n$ согласно формулам (2) и (3) п. 3 будет $\boldsymbol{C}$.

Обозначим матрицу линейного преобразования A в базисе $e_1, e_2, \ldots, e_n$ через $\boldsymbol{A} = \|a_{ik}\|$, а в базисе $f_1, f_2, \ldots \ldots, f_n$ через $\boldsymbol{B} = \|b_{ik}\|$. Иначе говоря,

$$Ae_k = \sum_{i=1}^{n} a_{ik}e_i, \tag{10$'$}$$

$$Af_k = \sum_{i=1}^{n} b_{ik}f_i. \tag{10$''$}$$

Наша цель — выразить матрицу $\boldsymbol{B}$ через матрицы $\boldsymbol{A}$ и $\boldsymbol{C}$. Заменим для этого в правой и левой частях формулы ($10''$) f_k через Ce_k и f_i через Ce_i. Мы будем иметь:

$$ACe_k = \sum_{i=1}^{n} b_{ik} Ce_i.$$

Применим к обеим частям этого равенства преобразование C^{-1} (оно существует, так как векторы $f_1, f_2, \ldots, f_n$ линейно независимы). Мы получим:

$$C^{-1}ACe_k = \sum_{i=1}^{n} b_{ik} e_i.$$

Мы видим, что интересующая нас матрица $\|b_{ik}\|$ есть также матрица преобразования $C^{-1}AC$ в базисе $e_1, e_2, \ldots, e_n$. При перемножении преобразований их матрицы в данном базисе $e_1, e_2, \ldots, e_n$ перемножаются. Поэтому

$$\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}. \tag{11}$$

Матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$, связанные соотношением (11), называются *подробными.*

Итак, *матрица* $\boldsymbol{B}$ *преобразования* A *в базисе* $f_1, f_2, \ldots, f_n$ *получается из матрицы* $\boldsymbol{A}$ *преобразования* A *в базисе* $e_1, e_2, \ldots, e_n$ *по формуле* (11), *где* $\boldsymbol{C}$ — *матрица перехода от базиса* $e_1, e_2, \ldots, e_n$ *к базису* $f_1, f_2, \ldots, f_n$ (*формула* (10)).

6. Линейное преобразование пространства R_1 в пространство R_2. Определяя линейное преобразование A, мы фактически нигде не пользовались тем, что векторы x и Ax принадлежат одному и тому же пространству. Поэтому, повторяя дословно определение 1 п. 1, можно определить также линейное преобразование пространства R_1 в другое пространство R_2.

Все сказанное в этом параграфе без каких-либо существенных изменений переносится на такие преобразования. Остановимся на операциях сложения и умножения линейных преобразований.

Пусть A и B — линейные преобразования пространства R_1 в пространство R_2. Тогда, как и в п. 3, можно определить их сумму $A + B$: $C = A + B$ означает, что $Cx = Ax + Bx$ для любого $x \in R_1$.

Произведение AB в этом случае смысла уже не имеет. Однако мы можем определить произведение AB в том случае, когда B — линейное преобразование пространства R_1 в R_2, а A — линейное преобразование пространства R_2 в R_3. В этом случае AB есть, по определению, линейное преобразование пространства R_1 в R_3, состоящее в последовательном выполнении сначала преобразования B, отображающего R_1 в R_2, а затем преобразования A, отображающего R_2 в R_3.

Введенные операции сложения и умножения линейных преобразований удовлетворяют ассоциативному и дистрибутивному законам.

З а д а ч а. Установить, как изменяется матрица линейного преобразования R_1 в R_2 при замене базисов в R_1 и R_2.

§ 10. Инвариантные подпространства, собственные векторы и собственные значения линейного преобразования

1. Инвариантные подпространства. Пусть R_1 — подпространство пространства R и A — линейное преобразование в R. Вообще говоря, для произвольного $x \in R_1$, $Ax \notin R_1$ *). Например, если R — евклидова плоскость, R_1 — произвольная прямая и A — поворот на угол $\varphi = \dfrac{\pi}{6}$, то очевидно, что для любого $x \neq 0$ и принадлежащего R_1, $Ax \notin R_1$. Однако может случиться,

*) Запись $Ax \notin R_1$ означает, что вектор Ax не принадлежит подпространству R_1.

что некоторые подпространства переходят сами в себя при линейном преобразовании A. Введем следующие определения.

Определение 1. *Пусть A — линейное преобразование пространства R. Линейное подпространство R_1 называется инвариантным относительно A, если для каждого вектора x из R_1 вектор Ax также принадлежит R_1.*

При изучении линейного преобразования A в инвариантном подпространстве R_1 можно, таким образом, рассматривать это преобразование только в R_1.

Тривиальными инвариантными подпространствами являются подпространство, состоящее лишь из нуля, и все пространство.

Примеры. 1. Пусть R — трехмерное пространство и A — поворот вокруг некоторой оси, проходящей через нуль. Инвариантными подпространствами при этом являются: а) ось вращения (одномерное инвариантное подпространство), б) плоскость, проходящая через начало координат и ортогональная к этой оси (двумерное инвариантное подпространство).

2. R — плоскость. Преобразование A заключается в растяжении плоскости в λ_1 раз вдоль оси X и в λ_2 раз вдоль оси Y. Иначе говоря, если вектор z равен $\xi_1 e_1 + \xi_2 e_2$, то $Az = \lambda_1 \xi_1 e_1 + \lambda_2 \xi_2 e_2$, где e_1, e_2 — единичные векторы на осях. Координатные оси X и Y являются в этом случае одномерными инвариантными подпространствами. Если $\lambda_1 = \lambda_2 = \lambda$, то A является преобразованием подобия с коэффициентом подобия λ. В этом случае каждая прямая, проходящая через начало координат, является инвариантным подпространством.

Упражнение. Показать, что если $\lambda_1 \neq \lambda_2$, то в примере 2 нет никаких других одномерных инвариантных подпространств, кроме указанных выше.

3. R — совокупность многочленов степени не выше $n-1$. Линейное преобразование A — дифференцирование, т. е.

$$AP(t) = P'(t).$$

Совокупность многочленов, степень которых меньше или равна k, где $k \leqslant n-1$, образует инвариантное подпространство. Действительно, дифференцируя многочлен степени $\leqslant k$, мы получим многочлен, степень которого снова не превосходит k.

Упражнение. Доказать, что в примере 3 никаких инвариантных подпространств, кроме указанных, нет.

4. R — произвольное n-мерное пространство. Линейное преобразование A задается в некотором базисе $e_1, e_2, \ldots, e_n$ матрицей вида

$$\begin{pmatrix} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ 0 & \ldots & 0 & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & a_{n,k+1} & \ldots & a_{nn} \end{pmatrix}.$$

В этом случае подпространство R_1, порожденное векторами $e_1, e_2, \ldots, e_k$, инвариантно. Доказательство этого мы предоставляем читателю. Если, кроме того,

$$a_{i,k+1} = \ldots = a_{in} = 0 \quad (1 \leqslant i \leqslant k),$$

то подпространство, порожденное векторами e_{k+1}, $e_{k+2}, \ldots, e_n$, также будет инвариантным.

5. R — произвольное n-мерное пространство, A — произвольное линейное преобразование в этом пространстве.

Тогда *образ M и ядро N преобразования A являются инвариантными подпространствами.* Действительно, пусть $y \in M$. Тогда $Ay \in M$ в силу определения M.

Точно так же, если $x \in N$, то $Ax = 0 \in N$.

Этот простой факт будет использован в дальнейшем при приведении произвольного преобразования к простейшему виду.

Пусть дано пространство R и линейное преобразование A в этом пространстве. Предположим, что R разложимо в прямую сумму двух инвариантных подпространств R_1 размерности k и R_2 размерности $n-k$ (см. стр. 27). Тогда в базисе $e_1, \ldots, e_n$, первые k векторов которого лежат в R_1, а последние $(n-k)$ — в R_2, матрица преобразования A состоит из двух клеток размерностей k и $n-k$, стоящих на диагонали, а на остальных местах стоят нули, т. е.

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \ldots & a_{1k} & 0 & \ldots & 0 \\ a_{21} & \ldots & a_{2k} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & 0 & \ldots & 0 \\ 0 & \ldots & 0 & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & a_{n,k+1} & \ldots & a_{n,n} \end{pmatrix}.$$

2. Собственные векторы и собственные значения. Особую роль в дальнейшем будут играть одномерные инвариантные подпространства.

Пусть R_1 — одномерное подпространство, порожденное вектором $x \neq 0$ (т. е. совокупность векторов вида αx). Ясно, что для того чтобы R_1 было инвариантным, необходимо и достаточно, чтобы вектор Ax лежал в R_1, т. е. был кратен вектору x:

$$Ax = \lambda x.$$

Определение 2. *Вектор $x \neq 0$, удовлетворяющий соотношению $Ax = \lambda x$, называется собственным вектором, а соответствующее число λ —*

собственным значением (*характеристическим числом*) *линейного преобразования* A.

Итак, если x — собственный вектор, то векторы αx образуют одномерное инвариантное подпространство.

Обратно, все отличные от нуля векторы одномерного инвариантного подпространства являются собственными.

Теорема 1. *В комплексном пространстве* *) R *всякое линейное преобразование* A *имеет хотя бы один собственный вектор.*

Доказательство. Выберем в R какой-либо базис $e_1, e_2, \ldots, e_n$. Линейному преобразованию A в этом базисе соответствует некоторая матрица $\|a_{ik}\|$.

Пусть

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

— произвольный вектор из R. Тогда координаты $\eta_1, \eta_2, \ldots, \eta_n$ вектора Ax выражаются следующими формулами (см. п. 2 § 9):

$$\begin{aligned}
\eta_1 &= a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n, \\
\eta_2 &= a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\eta_n &= a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n.
\end{aligned}$$

Условие того, что вектор собственный, т. е. равенство

$$Ax = \lambda x,$$

*) Доказательство теоремы пригодно для пространства над любым алгебраическим замкнутым полем, так как используется лишь существование решения у уравнения (2).

записывается в следующем виде:

$$\begin{aligned}
a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n &= \lambda\xi_1,\\
a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n &= \lambda\xi_2,\\
\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots\\
a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n &= \lambda\xi_n
\end{aligned}$$

или

$$\left.\begin{aligned}
(a_{11}-\lambda)\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n &= 0,\\
a_{21}\xi_1 + (a_{22}-\lambda)\xi_2 + \ldots + a_{2n}\xi_n &= 0,\\
\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots\\
a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + (a_{nn}-\lambda)\xi_n &= 0.
\end{aligned}\right\} \tag{1}$$

Для доказательства теоремы нужно доказать, таким образом, что существуют число λ и числа $\xi_1, \xi_2, \ldots, \xi_n$, не все равные нулю, удовлетворяющие системе (1).

Условием существования ненулевого решения однородной системы (1) является равенство нулю ее определителя

$$\begin{vmatrix}
a_{11}-\lambda & a_{12} & \ldots & a_{1n}\\
a_{21} & a_{22}-\lambda & \ldots & a_{2n}\\
\cdots & \cdots & \cdots & \cdots\\
a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda
\end{vmatrix} = 0. \tag{2}$$

Мы получили уравнение степени n относительно λ. Это уравнение имеет хотя бы один (вообще говоря, комплексный) корень λ_0.

Подставив в систему (1) вместо λ корень λ_0, мы получим однородную систему линейных уравнений, определитель которой равен нулю, и имеющую, следовательно, ненулевое решение $\xi_1^{(0)}, \xi_2^{(0)}, \ldots, \xi_n^{(0)}$. Тогда вектор

$$x^{(0)} = \xi_1^{(0)}e_1 + \xi_2^{(0)}e_2 + \ldots + \xi_n^{(0)}e_n$$

будет собственным вектором, а λ_0 — собственным значением, так как

$$Ax^{(0)} = \lambda_0 x^{(0)}.$$

Теорема доказана.

З а м е ч а н и е. Так как доказательство теоремы остается в силе, если преобразование A рассматривать не во всем пространстве, а в любом его инвариантном подпространстве, то *в любом инвариантном подпространстве существует хотя бы один собственный вектор преобразования* A.

Многочлен, стоящий в левой части уравнения (2), называется *характеристическим многочленом* матрицы преобразования A, а само уравнение (2) *характеристическим* или *вековым уравнением* этой матрицы. В процессе доказательства теоремы мы показали, что корни характеристического многочлена суть собственные значения преобразования A и, обратно, собственные значения преобразования A суть корни характеристического многочлена.

Так как собственные значения преобразования определены независимо от выбора базиса, то, следовательно, и корни характеристического многочлена также не зависят от выбора базиса. Мы покажем далее несколько больше *), а именно, что *сам характеристический многочлен не зависит от выбора базиса*, и поэтому мы в дальнейшем будем называть его *характеристическим многочленом преобразования* A (а не характеристическим многочленом матрицы преобразования A).

3. Среди линейных преобразований в известном смысле простейшими являются те, которые имеют n линейно независимых собственных векторов.

Пусть A — такое преобразование, а $e_1, e_2, \ldots, e_n$ — его линейно независимые собственные векторы, т. е.

$$Ae_i = \lambda_i e_i \quad (i = 1, 2, \ldots, n).$$

---

*) Из того, что корни характеристического многочлена одни и те же для разных базисов, еще не следует, что сам многочлен не зависит от выбора базиса; априори возможно, что в разных базисах кратности этих корней различны.

Примем $e_1, e_2, \ldots, e_n$ за базис в R. Равенства

$$\begin{aligned} &Ae_1 = \lambda_1 e_1, \\ &Ae_2 = \lambda_2 e_2, \\ &\ldots\ldots\ldots \\ &Ae_n = \lambda_n e_n \end{aligned}$$

означают, что матрица преобразования A в этом базисе имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

(является *диагональной матрицей*). Таким образом, имеет место

Теорема 2. *Если линейное преобразование A имеет n линейно независимых собственных векторов, то, выбрав их за базис, мы приведем матрицу преобразования A к диагональной форме. Обратно, если в некотором базисе матрица преобразования диагональна, то все векторы этого базиса являются собственными векторами.*

Замечание. Отметим один важный случай, когда линейное преобразование заведомо имеет n линейно независимых собственных векторов. Предварительно заметим следующее:

Если $e_1, e_2, \ldots, e_k$ — собственные векторы преобразования A и соответствующие им собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_k$, попарно различны, то $e_1, e_2, \ldots, e_k$ линейно независимы.

Для $k = 1$ утверждение очевидно. Пусть наше утверждение верно для $k-1$ векторов; докажем его для k векторов. Предположим противное, т. е. предположим, что

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k = 0, \tag{3}$$

причем хотя бы один из коэффициентов α_i, например α_1, отличен от нуля.

Применим к обеим частям равенства (3) преобразование A. Получим

$$A(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k) = 0,$$

т. е.

$$\alpha_1 \lambda_1 e_1 + \alpha_2 \lambda_2 e_2 + \ldots + \alpha_k \lambda_k e_k = 0.$$

Вычитая из последнего равенства равенство (3), умноженное на λ_k, мы получим выражение

$$\alpha_1(\lambda_1 - \lambda_k)e_1 + \alpha_2(\lambda_2 - \lambda_k)e_2 + \ldots \\ \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)e_{k-1} = 0,$$

где первый коэффициент по-прежнему отличен от нуля (так как по условию $\lambda_i \neq \lambda_k$ при $i \neq k$). Мы пришли к противоречию, так как по индуктивному предположению векторы $e_1, e_2, \ldots, e_{k-1}$ линейно независимы. Отсюда непосредственно следует, что:

Если характеристический многочлен преобразования A имеет n различных корней, то матрица преобразования A может быть приведена к диагональной форме.

Действительно, каждому корню λ_k характеристического уравнения отвечает хотя бы один собственный вектор. Так как соответствующие этим векторам собственные значения (корни характеристического уравнения) все различны, то, согласно доказанному выше, мы имеем n линейно независимых собственных векторов $e_1, e_2, \ldots, e_n$. Если векторы $e_1, e_2, \ldots, e_n$ принять за базис, то матрица преобразования A будет диагональной.

Если характеристический многочлен имеет кратные корни, то число линейно независимых собственных векторов может быть меньше, чем n. Например, преобразование A в пространстве многочленов степени не выше $n-1$, ставящее в соответствие каждому многочлену его производную, имеет лишь одно собственное значение $\lambda = 0$ и один (с точностью до пропорциональности) собственный вектор $P(t) = \text{const}$. В самом деле, для любого многочлена $P(t)$ степени $k > 0$ многочлен $P'(t)$ имеет степень $k - 1$, и потому равенство $P'(t) = \lambda P(t)$ возможно лишь, если $\lambda = 0$ и $P(t) = \text{const}$. Следовательно, для этого преобразования не существует базиса, в котором ему соответствовала бы диагональная матрица.

В главе III будет доказано, что если λ есть m-кратный корень характеристического уравнения, то ему отвечает не более чем m линейно независимых собственных векторов.

Ниже (в §§ 12 и 13) мы укажем некоторые классы линейных преобразований, приводимых к диагональной форме. Вопросу о том, к какому простейшему виду может быть приведено произвольное линейное преобразование, будет посвящена глава III.

4. Характеристический многочлен. В п. 2 мы уже определили характеристический многочлен преобразования A как определитель матрицы $\boldsymbol{A} - \lambda \boldsymbol{E}$, где $\boldsymbol{A}$ — матрица преобразования A, а $\boldsymbol{E}$ — единичная матрица. Докажем, что характеристический многочлен не зависит от выбора базиса. Действительно, при переходе к другому базису матрица $\boldsymbol{A}$ преобразования A принимает вид $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, где $\boldsymbol{C}$ есть матрица перехода к новому базису. Таким образом, в новом базисе характеристический многочлен есть определитель матрицы $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} - \lambda\boldsymbol{E}$. Но

$$|\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} - \lambda\boldsymbol{E}| = |\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} - \lambda\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{C}| = |\boldsymbol{C}^{-1}(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{C}|,$$

и так как определитель произведения равен произведению определителей, то

$$|\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} - \lambda\boldsymbol{E}| = |\boldsymbol{C}^{-1}||\boldsymbol{A} - \lambda\boldsymbol{E}||\boldsymbol{C}| = |\boldsymbol{A} - \lambda\boldsymbol{E}|,$$

и утверждение доказано. Таким образом, мы в дальнейшем можем говорить о характеристическом многочлене преобразования A (а не о характеристическом многочлене матрицы преобразования A).

Упражнения. 1. Найти характеристический многочлен матрицы

$$\begin{pmatrix} \lambda_0 & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda_0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda_0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & \lambda_0 \end{pmatrix}.$$

2. Найти характеристический многочлен матрицы

$$\begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_{n-1} & a_n \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Ответ. $(-1)^n(\lambda^n - a_1\lambda^{n-1} - a_2\lambda^{n-2} - \ldots - a_n)$.

Выразим характеристический многочлен явно через элементы матрицы $\boldsymbol{A}$ преобразования A. Вычислим сначала более общий определитель (который позже в § 12 нам тоже встретится): $|\boldsymbol{A} - \lambda\boldsymbol{B}|$, где $\boldsymbol{A}$ и $\boldsymbol{B}$ — две заданные матрицы. Нам нужно, следовательно, вычислить следующий многочлен относительно λ:

$$Q(\lambda) = \begin{vmatrix} a_{11} - \lambda b_{11} & a_{12} - \lambda b_{12} & \dots & a_{1n} - \lambda b_{1n} \\ a_{21} - \lambda b_{21} & a_{22} - \lambda b_{22} & \dots & a_{2n} - \lambda b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} - \lambda b_{n1} & a_{n2} - \lambda b_{n2} & \dots & a_{nn} - \lambda b_{nn} \end{vmatrix}.$$

Так как в этом определителе каждый столбец есть сумма двух столбцов, то определитель может быть разложен на сумму определителей. Свободный член в $Q(\lambda)$ есть

$$q_0 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{4}$$

Ясно, что *коэффициент при* $(-\lambda)^k$ *в* $Q(\lambda)$ *равен сумме определителей, каждый из которых получается заменой в* (4) *каких-либо* k *столбцов матрицы* $\|a_{ik}\|$ *соответствующими столбцами матрицы* $\|b_{ik}\|$.

Перейдем теперь к вычислению $|\boldsymbol{A}-\lambda\boldsymbol{E}|$. Для вычисления коэффициента при $(-\lambda)^k$ мы должны взять сумму определителей, каждый из которых получается заменой k столбцов матрицы $\|a_{ik}\|$ k столбцами единичной матрицы. Но каждый такой определитель есть главный минор $(n-k)$-го порядка матрицы $\|a_{ik}\|$. Таким образом, окончательно, *характеристический многочлен* $P(\lambda)$ *матрицы* $\boldsymbol{A}$ *имеет вид*

$$P(\lambda)=(-1)^n(\lambda^n-p_1\lambda^{n-1}+p_2\lambda^{n-2}-\ldots\pm p_n),$$

где p_1 *есть сумма диагональных элементов,* p_2 *— сумма главных миноров второго порядка и т. д.; наконец* p_n *есть определитель матрицы* $\boldsymbol{A}$.

Числа $p_1, p_2, \ldots, p_n$, построенные по матрице $\boldsymbol{A}$ преобразования A, зависят лишь от самого преобразования, поскольку этим свойством, как мы показали, обладает характеристический многочлен. Среди коэффициентов p_i наибольшую роль играют p_n — определитель матрицы $\boldsymbol{A}$ и p_1 — сумма диагональных элементов матрицы $\boldsymbol{A}$. *Сумма диагональных элементов матрицы называется следом матрицы* $\boldsymbol{A}$. След матрицы $\boldsymbol{A}$ обозначается $\operatorname{tr}\boldsymbol{A}$ (от английского слова trace — след). Ясно, что след матрицы равен сумме всех корней характеристического многочлена (собственных значений), причем каждый корень считается с той кратностью, с которой он входит в характеристический многочлен.

Упражнения. 1. Показать, что если $\boldsymbol{A}$ и $\boldsymbol{B}$ — матрицы n-го порядка, то

$$\operatorname{tr}\boldsymbol{AB}=\operatorname{tr}\boldsymbol{BA}.$$

2. Показать, что если $\boldsymbol{C}$ — невырожденная матрицы n-го порядка, то для любой матрицы $\boldsymbol{A}$ n-го порядка имеем:

$$\operatorname{tr}\boldsymbol{C}^{-1}\boldsymbol{AC}=\operatorname{tr}\boldsymbol{A}.$$

Вычисление собственных векторов линейного преобразования требует знания собственных значений и, следовательно, решения уравнения n-й степени — характеристического уравнения. В одном важном частном случае корни характеристического многочлена можно найти непосредственно.

Если матрица преобразования A треугольная, т. е. имеет вид

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}, \tag{5}$$

то собственными значениями будут числа, стоящие на диагонали, т. е. $a_{11}, a_{22}, \dots, a_{nn}$.

В самом деле, характеристический многочлен данной матрицы вычисляется непосредственно и есть

$$P(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda),$$

и следовательно, его корни — $a_{11}, a_{22}, \dots, a_{nn}$.

Упражнение. Найти собственные векторы, отвечающие собственным значениям a_{11}, a_{22}, a_{33} треугольной матрицы (5).

В заключение этого пункта укажем одно интересное свойство характеристического многочлена. Как мы уже указывали в п. 3 предыдущего параграфа, существует такой многочлен $P(t)$, что если в него подставить вместо t матрицу $\boldsymbol{A}$, то он обратится в нуль. Мы покажем сейчас, что одним из таких многочленов является характеристический многочлен. Докажем предварительно лемму:

Лемма. *Пусть многочлен*

$$P(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \dots + a_m$$

и матрица $\boldsymbol{A}$ связаны соотношением

$$P(\lambda)\boldsymbol{E} = (\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{C}(\lambda), \tag{6}$$

где $\boldsymbol{C}(\lambda)$ — многочлен от λ, коэффициенты которого являются матрицами, т. е.

$$\boldsymbol{C}(\lambda) = \boldsymbol{C}_0\lambda^{m-1} + \boldsymbol{C}_1\lambda^{m-2} + \dots + \boldsymbol{C}_{m-1}, \quad \boldsymbol{C}_i \text{ — матрицы.}$$

Тогда $P(\boldsymbol{A}) = 0$.

Заметим, что эта лемма является обобщением на многочлены с матричными коэффициентами теоремы Безу.

Доказательство. Мы имеем:

$$(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{C}(\lambda)=\boldsymbol{A}\boldsymbol{C}_{m-1}+(\boldsymbol{A}\boldsymbol{C}_{m-2}-\boldsymbol{C}_{m-1})\lambda+ \\ +(\boldsymbol{A}\boldsymbol{C}_{m-3}-\boldsymbol{C}_{m-2})\lambda^2+\ldots-\boldsymbol{C}_0\lambda^m. \quad (7)$$

Сравнивая между собой коэффициенты при одинаковых степенях λ в обеих частях равенства (6), мы получаем последовательность равенств:

$$\begin{aligned}
&\boldsymbol{A}\boldsymbol{C}_{m-1} &&= a_m\boldsymbol{E},\\
&\boldsymbol{A}\boldsymbol{C}_{m-2}-\boldsymbol{C}_{m-1} &&= a_{m-1}\boldsymbol{E},\\
&\boldsymbol{A}\boldsymbol{C}_{m-3}-\boldsymbol{C}_{m-2} &&= a_{m-2}\boldsymbol{E},\\
&\ldots\ldots\ldots\ldots\ldots\ldots &&\\
&\boldsymbol{A}\boldsymbol{C}_0\quad-\boldsymbol{C}_1 &&= a_1\boldsymbol{E},\\
&\qquad\quad\;\, -\boldsymbol{C}_0 &&= a_0\boldsymbol{E}.
\end{aligned} \quad (8)$$

Умножим теперь слева первое равенство на $\boldsymbol{E}$, второе на $\boldsymbol{A}$, третье на $\boldsymbol{A}^2$, ..., последнее на $\boldsymbol{A}^m$ и сложим их. Мы получим справа $P(\boldsymbol{A}) = a_m\boldsymbol{E}+a_{m-1}\boldsymbol{A}+\ldots+a_0\boldsymbol{A}^m$, а слева 0. Таким образом, $P(\boldsymbol{A})=0$ и лемма доказана *).

Теорема 3. *Если $P(\lambda)$ — характеристический многочлен матрицы $\boldsymbol{A}$, то $P(\boldsymbol{A})=0$.*

Доказательство. Рассмотрим матрицу, обратную матрице $\boldsymbol{A}-\lambda\boldsymbol{E}$. Мы имеем $(\boldsymbol{A}-\lambda\boldsymbol{E})(\boldsymbol{A}-\lambda\boldsymbol{E})^{-1}=\boldsymbol{E}$. Как известно, обратная матрица может быть записана в виде

$$(\boldsymbol{A}-\lambda\boldsymbol{E})^{-1}=\frac{1}{P(\lambda)}\boldsymbol{C}(\lambda),$$

где $\boldsymbol{C}(\lambda)$ — матрица из миноров $(n-1)$-го порядка матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$, а $P(\lambda)$ — определитель матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$, т. е. характеристический многочлен матрицы $\boldsymbol{A}$. Отсюда

$$(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{C}(\lambda)=P(\lambda)\boldsymbol{E}.$$

---

*) В алгебре теорема Безу доказывается тем, что в равенство (6) подставляется $\boldsymbol{A}$ вместо λ. Здесь, конечно, мы не вправе этого сделать непосредственно, так как λ есть число, а $\boldsymbol{A}$ — матрица. Однако, по существу, мы сделали то же самое. Действительно, k-е из равенств (8) получилось сравнением в (6) коэффициентов при λ^k. Умножая его на $\boldsymbol{A}^k$ и складывая затем все равенства, мы, по существу, подставляем $\boldsymbol{A}$ вместо λ.

Так как элементами матрицы $\boldsymbol{C}(\lambda)$ являются миноры матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$, т. е. многочлены степени не выше $n-1$ относительно λ, то согласно доказанной лемме

$$P(\boldsymbol{A})=0,$$

и теорема доказана.

Заметим, что если у характеристического многочлена матрицы $\boldsymbol{A}$ нет кратных корней, то не существует многочлена степени ниже n, обращающегося в нуль при подстановке в него матрицы $\boldsymbol{A}$ (см. следующее упражнение).

Упражнение. Пусть $\boldsymbol{A}$ — диагональная матрица вида

$$\boldsymbol{A}=\begin{pmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \lambda_n\end{pmatrix},$$

где все числа λ_i различны. Найти многочлен $P(t)$ возможно более низкой степени, для которого $P(\boldsymbol{A})=0$. (См. пример в § 9, п. 3.)

§ 11. Линейное преобразование, сопряженное к данному

1. Связь между линейными преобразованиями и билинейными формами в евклидовом пространстве. Мы рассматривали ранее в аффинном пространстве отдельно линейные преобразования и отдельно билинейные формы. В случае евклидова пространства между билинейными формами и линейными преобразованиями существует тесная связь *).

Всякому линейному преобразованию A отвечает в евклидовом пространстве билинейная форма $A(x;y)$,

*) Так как в данном базисе как линейные преобразования, так и билинейные формы задаются матрицами, то можно было бы попытаться в аффинном пространстве поставить друг другу в соответствие линейное преобразование и билинейную форму, задаваемые одной и той же матрицей. Однако это соответствие было бы случайным. Действительно, если в одном базисе матрицы билинейной формы и линейного преобразования совпадают, то в другом базисе они будут уже, вообще говоря, различны, так как

задаваемая формулой

$$A(x; y) \equiv (Ax, y).$$

Действительно, функция $A(x; y) \equiv (Ax, y)$ удовлетворяет условиям, определяющим билинейную форму. Имеем:

$$1^\circ \qquad (A(x_1+x_2), y) = (Ax_1+Ax_2, y) = (Ax_1, y)+(Ax_2, y),$$
$$(A\lambda x, y) = (\lambda Ax, y) = \lambda(Ax, y).$$

$$2^\circ \qquad (x, A(y_1+y_2)) = (x, Ay_1+Ay_2) = (x, Ay_1)+(x, Ay_2),$$
$$(x, A\mu y) = (x, \mu Ay) = \overline{\mu}(x, Ay).$$

Покажем, что преобразование A определяется соответствующей билинейной формой $A(x; y)$ однозначно.

Пусть

$$A(x; y) = (Ax, y)$$

и

$$A(x; y) = (Bx, y).$$

Тогда

$$(Ax, y) \equiv (Bx, y),$$

т. е.

$$(Ax - Bx, y) = 0$$

при переходе к другому базису матрица $\boldsymbol{A}$ билинейной формы переходит в $\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$ ($\boldsymbol{C}'$ — матрица, транспонированная к матрице $\boldsymbol{C}$) (см. § 4), а матрица линейного преобразования — в $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ (см. § 9).

Внимательный читатель сможет заметить, что устанавливаемое ниже соответствие между билинейными формами и линейными преобразованиями в евклидовом пространстве состоит в том, что сопоставляются друг другу линейные преобразования и билинейные формы, матрицы которых в нормированном ортогональном базисе получаются одна из другой транспонированием; это соответствие, как следует из дальнейшего, уже не зависит от выбора базиса.

для любого вектора y; но это значит, что $Ax - Bx = 0$. Таким образом, $Ax = Bx$ для любого x, т. е. $A = B$. Однозначность доказана.

Имеет место и обратное.

Пусть R — комплексное евклидово пространство и пусть $A(x; y)$ — билинейная форма в нем. Выберем в R какой-либо ортогональный нормированный базис e_1, $e_2, \ldots, e_n$. Если

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n \quad \text{и} \quad y = \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n,$$

то $A(x; y)$ можно записать в виде

$$\begin{aligned} A(x; y) = {} & a_{11}\xi_1\overline{\eta}_1 + a_{12}\xi_1\overline{\eta}_2 + \ldots + a_{1n}\xi_1\overline{\eta}_n + \\ & + a_{21}\xi_2\overline{\eta}_1 + a_{22}\xi_2\overline{\eta}_2 + \ldots + a_{2n}\xi_2\overline{\eta}_n + \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & + a_{n1}\xi_n\overline{\eta}_1 + a_{n2}\xi_n\overline{\eta}_2 + \ldots + a_{nn}\xi_n\overline{\eta}_n. \end{aligned} \tag{1}$$

Постараемся представить это выражение в виде некоторого скалярного произведения. Для этого перепишем его следующим образом:

$$\begin{aligned} A(x; y) = {} & (a_{11}\xi_1 + a_{21}\xi_2 + \ldots + a_{n1}\xi_n)\overline{\eta}_1 + \\ & + (a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{n2}\xi_n)\overline{\eta}_2 + \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & + (a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{nn}\xi_n)\overline{\eta}_n. \end{aligned}$$

Введем в рассмотрение вектор z с координатами

$$\begin{aligned} \zeta_1 &= a_{11}\xi_1 + a_{21}\xi_2 + \ldots + a_{n1}\xi_n, \\ \zeta_2 &= a_{12}\xi_1 + a_{22}\xi_2 + \ldots + a_{n2}\xi_n, \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \zeta_n &= a_{1n}\xi_1 + a_{2n}\xi_2 + \ldots + a_{nn}\xi_n. \end{aligned}$$

Вектор z получается из вектора x линейным преобразованием с матрицей, транспонированной к матрице $\|a_{ik}\|$ билинейной формы $A(x; y)$. Это преобразование мы обозначим буквой A, т. е. положим $z = Ax$. Мы получаем, следовательно, что

$$A(x;y) = \zeta_1\overline{\eta}_1 + \zeta_2\overline{\eta}_2 + \ldots + \zeta_n\overline{\eta}_n = (z,y) = (Ax,y).$$

Итак, *всякой билинейной форме $A(x;y)$ в евклидовом пространстве отвечает такое линейное преобразование A, что*

$$A(x;y) \equiv (Ax,y).$$

Таким образом, мы доказали следующую теорему.

Т е о р е м а 1. *Формула*

$$A(x;y) = (Ax,y) \tag{2}$$

устанавливает в евклидовом пространстве взаимно однозначное соответствие между билинейными формами и линейными преобразованиями.

Из однозначности соответствия, устанавливаемого формулой (2), следует, что оно не зависит от выбора базиса.

Связь между билинейными формами и линейными преобразованиями можно установить и другим способом. А именно, каждую билинейную форму можно представить также в виде

$$A(x;y) = (x, A^*y).$$

Для этого в формуле (1)

$$\begin{aligned} A(x;y) = {} & a_{11}\xi_1\overline{\eta}_1 + a_{12}\xi_1\overline{\eta}_2 + \ldots + a_{1n}\xi_1\overline{\eta}_n + \\ & + a_{21}\xi_2\overline{\eta}_1 + a_{22}\xi_2\overline{\eta}_2 + \ldots + a_{2n}\xi_2\overline{\eta}_n + \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & + a_{n1}\xi_n\overline{\eta}_1 + a_{n2}\xi_n\overline{\eta}_2 + \ldots + a_{nn}\xi_n\overline{\eta}_n \end{aligned}$$

мы будем выносить за скобки координаты $\xi_1, \xi_2, \ldots, \xi_n$ вектора x. Повторяя снова прежние рассуждения, мы получаем:

$$\begin{aligned}
A(x;y) &= \xi_1(a_{11}\overline{\eta}_1 + a_{12}\overline{\eta}_2 + \ldots + a_{1n}\overline{\eta}_n) + \\
&+ \xi_2(a_{21}\overline{\eta}_1 + a_{22}\overline{\eta}_2 + \ldots + a_{2n}\overline{\eta}_n) + \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ \xi_n(a_{n1}\overline{\eta}_1 + a_{n2}\overline{\eta}_2 + \ldots + a_{nn}\overline{\eta}_n) = \\
&= \xi_1\overline{(\overline{a}_{11}\eta_1 + \overline{a}_{12}\eta_2 + \ldots + \overline{a}_{1n}\eta_n)} + \\
&+ \xi_2\overline{(\overline{a}_{21}\eta_1 + \overline{a}_{22}\eta_2 + \ldots + \overline{a}_{2n}\eta_n)} + \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ \xi_n\overline{(\overline{a}_{n1}\eta_1 + \overline{a}_{n2}\eta_2 + \ldots + \overline{a}_{nn}\eta_n)} = (x, A^*y).
\end{aligned}$$

При этом матрица преобразования A^* получается из матрицы преобразования A в любом *ортогональном* базисе переходом к транспонированной и заменой ее элементов комплексно сопряженными.

Заметим, что в неортогональном базисе связь между матрицами преобразований A и A^* более сложна.

2. Операция перехода от преобразования A к сопряженному (операция *).

Определение 1. *Пусть A — линейное преобразование комплексного евклидова пространства. Преобразование A^*, определенное условием*

$$(Ax, y) = (x, A^*y),$$

называется сопряженным к A.

Теорема 2. *В евклидовом пространстве каждому линейному преобразованию отвечает сопряженное преобразование и притом только одно.*

Доказательство. Линейному преобразованию A однозначно соответствует согласно теореме 1 этого параграфа билинейная форма $A(x;y) = (Ax, y)$. Эту билинейную форму согласно сказанному в конце п. 1 можно представить, и притом однозначно, в виде

(x, A^*y). Окончательно мы имеем:

$$(Ax, y) = A(x; y) = (x, A^*y).$$

Матрица сопряженного преобразования A^* *получается из матрицы преобразования* A *в ортогональном базисе переходом к транспонированной и комплексно сопряженной матрице,* как это доказано в п. 1 этого параграфа.

Переход от A к A^* можно выразить в виде правила: если в выражении (Ax, y) мы желаем A перебросить на второе место, то к нему нужно приписать *.

Операция перехода от преобразования A к сопряженному преобразованию A^* («операция *») связана с определенными выше (§ 9) операциями сложения и умножения линейных преобразований следующими соотношениями:

1° $(AB)^* = B^*A^*$.
2° $(A^*)^* = A$.
3° $(A + B)^* = A^* + B^*$.
4° $(\lambda A)^* = \overline{\lambda} A^*$.
5° $E^* = E$.

Докажем, например, два первых из этих свойств.

1° $$(ABx, y) = (Bx, A^*y) = (x, B^*A^*y).$$

Но, с другой стороны, по определению $(AB)^*$ имеем:

$$(ABx, y) = (x, (AB)^*y).$$

Сравнивая правые части этих двух равенств и вспомнив, что линейное преобразование однозначно определяется соответствующей билинейной формой, получаем:

$$(AB)^* = B^*A^*.$$

2° По определению A^* имеем:

$$(Ax, y) = (x, A^*y).$$

Обозначим временно A^* через C. Тогда

$$(Ax, y) = (x, Cy),$$

откуда

$$(y, Ax) = (Cy, x).$$

Заменив y через x, а x через y и поменяв местами правую и левую части этого равенства, получим:

$$(Cx, y) = (x, Ay).$$

Но это равенство и означает, что $C^* = A$, и так как $C = A^*$, то

$$(A^*)^* = A.$$

Упражнения. 1. Доказать таким же способом свойства $3°$–$5°$.

2. Доказать свойства $1°$–$5°$, пользуясь тем, что матрица преобразования A^* получается из матрицы преобразования A в ортогональном базисе транспонированием и заменой всех элементов комплексно сопряженными.

3. Самосопряженные, унитарные и нормальные линейные преобразования. Операция * в известной мере аналогична операции перехода от данного комплексного числа α к сопряженному $\overline{\alpha}$. Эта аналогия не случайна. Действительно, для матриц первого порядка над комплексным полем, т. е. для комплексных чисел, операция * как раз и состоит в замене данного числа комплексно сопряженным.

Среди всех комплексных чисел действительные числа характеризуются тем свойством, что $\overline{\alpha} = \alpha$. Для линейных преобразований аналогичное понятие является весьма существенным.

Определение 2. *Линейное преобразование A называется самосопряженным (или эрмитовым), если $A^* = A$.*

Покажем, что *для того, чтобы линейное преобразование A было самосопряженным, необходимо и достаточно, чтобы билинейная форма (Ax, y) была эрмитовой.*

В самом деле, эрмитовость формы (Ax, y) означает, что

$$(Ax, y) = \overline{(Ay, x)}. \qquad \text{(а)}$$

Самосопряженность преобразования A означает, что

$$(Ax, y) = (x, Ay). \qquad \text{(б)}$$

Легко видеть, что равенства (а) и (б) эквивалентны.

Всякое комплексное число ζ представимо в виде $\zeta = \alpha + i\beta$, где α и β — действительные числа. Аналогично:

Всякое линейное преобразование A может быть записано в виде

$$A = A_1 + iA_2 \qquad (3),$$

где A_1 и A_2 — самосопряженные преобразования.

Действительно,

$$A = \frac{A + A^*}{2} + i\frac{A - A^*}{2i}.$$

Введем обозначения

$$\frac{A + A^*}{2} = A_1, \quad \frac{A - A^*}{2i} = A_2.$$

Тогда

$$A_1^* = \left(\frac{A + A^*}{2}\right)^* = \frac{1}{2}(A + A^*)^* = \frac{1}{2}(A^* + A^{**}) = \\ = \frac{1}{2}(A^* + A) = A_1$$

и

$$A_2^* = \left(\frac{A - A^*}{2i}\right)^* = -\frac{1}{2i}(A - A^*)^* = -\frac{1}{2i}(A^* - A^{**}) = \\ = -\frac{1}{2i}(A^* - A) = A_2,$$

т.е. A_1 и A_2 — самосопряженные преобразования.

Таким образом, самосопряженные преобразования играют среди всех линейных преобразований роль, аналогичную роли действительных чисел среди комплексных.

Упражнения. 1. Доказать единственность представления преобразования A в виде (3).

2. Доказать, что линейная комбинация с действительными коэффициентами самосопряженных преобразований есть снова самосопряженное преобразование.

3. Доказать, что если A — произвольное линейное преобразование, то преобразования AA^* и A^*A — самосопряженные.

Примечание. В отличие от комплексных чисел, AA^*, вообще говоря, не равно A^*A.

Произведение двух самосопряженных линейных преобразований не есть, вообще говоря, самосопряженное преобразование. Имеет место следующая

Теорема 3. *Пусть A и B — самосопряженные линейные преобразования. Для того чтобы преобразование AB было также самосопряженным, необходимо и достаточно, чтобы $AB = BA$, т. е. чтобы преобразования A и B были перестановочны.*

Доказательство. Нам дано, что

$$A^* = A \quad \text{и} \quad B^* = B.$$

Мы ищем необходимое и достаточное условие того, чтобы выполнялось равенство

$$(AB)^* = AB. \tag{4}$$

Но

$$(AB)^* = B^*A^* = BA.$$

Следовательно, равенство (4) имеет место тогда и только тогда, когда

$$AB = BA.$$

Теорема доказана.

Упражнение. Доказать, что если A и B — самосопряженные преобразования, то самосопряженными будут и преобразования $AB + BA$ и $i(AB - BA)$.

Аналогом комплексных чисел, равных по модулю единице, т.е. таких, что $z\overline{z} = 1$, являются унитарные преобразования.

Определение 3. *Линейное преобразование U называется унитарным, если $UU^* = U^*U = E$* *). *Другими словами, для унитарного преобразования $U^* = U^{-1}$.*

В § 13 мы познакомимся с весьма простой геометрической интерпретацией унитарных преобразований.

Упражнения. 1. Доказать, что произведение двух унитарных преобразований есть снова унитарное преобразование.

2. Показать, что если U — унитарное преобразование, а A — самосопряженное преобразование, то $U^{-1}AU$ — также самосопряженное.

Ниже (в § 15) мы докажем, что всякое линейное преобразование можно представить как произведение самосопряженного на унитарное. Эту теорему можно рассматривать как обобщение записи комплексного числа в тригонометрической форме.

Введем еще одно определение.

---

*) В n-мерном пространстве условия $U^*U = E$ и $UU^* = E$ эквивалентны. В бесконечномерном пространстве это — два различных условия.

Определение 4. *Линейное преобразование A называется нормальным, если $AA^* = A^*A$.*

Для комплексных чисел нет надобности в аналогичном понятии, так как умножение комплексных чисел коммутативно и, значит, $\overline{\alpha}\alpha$ всегда равно $\alpha\overline{\alpha}$.

Нетрудно убедиться, что как самосопряженные, так и унитарные преобразования являются частными случаями нормальных преобразований.

Более детальному изучению отдельных классов линейных преобразований в евклидовом пространстве будут посвящены дальнейшие параграфы этой главы. При этом мы получим для различных типов преобразований весьма простую геометрическую характеристику.

§ 12. Самосопряженные (эрмитовы) преобразования. Одновременное приведение пары квадратичных форм к сумме квадратов

1. Самосопряженные преобразования. В этом параграфе мы более подробно изучим класс самосопряженных преобразований n-мерного евклидова пространства. Эти преобразования часто встречаются в различных приложениях. (Существенную роль самосопряженные преобразования, правда в бесконечномерном пространстве, играют в квантовой механике).

Лемма 1. *Собственные значения самосопряженного преобразования вещественны.*

Доказательство. Пусть x — собственный вектор самосопряженного преобразования A и λ — соответствующее собственное значение, т. е.

$$Ax = \lambda x; \quad x \neq 0.$$

Так как $A^* = A$, то

$$(Ax, x) = (x, Ax),$$

т. е.

$$(\lambda x, x) = (x, \lambda x).$$

Вынося λ за скобки, получим:

$$\lambda(x, x) = \overline{\lambda}(x, x),$$

и так как $(x, x) \neq 0$, то $\lambda = \overline{\lambda}$, что и требовалось доказать.

Лемма 2. *Пусть A — самосопряженное линейное преобразование в n-мерном пространстве R и e — его собственный вектор. Совокупность R_1 векторов x, ортогональных к e, есть $(n-1)$-мерное подпространство, инвариантное относительно преобразования A.*

Доказательство. Совокупность R_1 векторов x, ортогональных к e, образует $(n-1)$-мерное подпространство.

Покажем, что R_1 инвариантно относительно A. Пусть $x \in R_1$. Это значит, что $(x, e) = 0$. Тогда и $(Ax, e) = 0$, т. е. $Ax \in R_1$. Действительно,

$$(Ax, e) = (x, A^*e) = (x, Ae) = (x, \lambda e) = \lambda(x, e) = 0.$$

Мы доказали, что преобразование A не выводит векторы, принадлежащие R_1, из R_1, т. е. доказали, что подпространство R_1 инвариантно относительно A.

Теорема 1. *Пусть A — самосопряженное преобразование в n-мерном евклидовом пространстве R. Тогда существует n попарно ортогональных собственных векторов преобразования A. Соответствующие им собственные значения вещественны.*

Доказательство. Согласно теореме 1 § 10 в R существует хотя бы один собственный вектор e_1 преобразования A. В силу леммы 2 совокупность векторов, ортогональных к e_1, образует $(n-1)$-мерное инвариантное подпространство R_1. Будем далее рассматривать наше преобразование A лишь в R_1. В R_1 существует собственный вектор e_2 (см. замечание к теореме 1

§ 10). Совокупность векторов из R_1, ортогональных к e_2, образует $(n-2)$-мерное инвариантное подпространство R_2. В нем существует собственный вектор e_3 и т. д.

Мы получаем, таким образом, n попарно ортогональных собственных векторов $e_1, e_2, \ldots, e_n$. Согласно лемме 1 соответствующие им собственные значения вещественны. Теорема доказана.

Так как произведение собственного вектора на любое отличное от нуля число есть снова собственный вектор, то векторы e_i можно выбрать так, чтобы их длины равнялись единице.

Теорема 2. *Пусть A — самосопряженное преобразование в n-мерном пространстве. Тогда существует ортогональный базис, в котором матрица преобразования диагональна и вещественна. Верно также и обратное.*

Доказательство. Выберем в качестве базиса построенные в теореме 1 попарно ортогональные собственные векторы $e_1, e_2, \ldots, e_n$.

Тогда

$$\begin{aligned} Ae_1 &= \lambda_1 e_1, \\ Ae_2 &= \lambda_2 e_2, \\ &\ldots\ldots\ldots \\ Ae_n &= \lambda_n e_n, \end{aligned}$$

т. е. в этом базисе матрица преобразования A имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \tag{1}$$

где все λ_i вещественны.

Обратно, пусть матрица преобразования A в ортогональном базисе имеет вид (1). В ортогональном нормированном базисе матрица сопряженного преобразования A^* получается из матрицы преобразования A транспонированием и заменой каждого элемента комплексно сопряженным (см. § 11). Проделав эти операции над матрицей вида (1) (где все λ_i вещественны), мы получим ту же самую матрицу. Следовательно, преобразованиям A и A^* соответствует одна и та же матрица, т. е. $A = A^*$. Теорема полностью доказана.

Отметим еще следующее свойство собственных векторов самосопряженного преобразования: *собственные векторы, соответствующие различным собственным значениям, взаимно ортогональны.*

Действительно, пусть

$$Ae_1 = \lambda_1 e_1, \quad Ae_2 = \lambda_2 e_2, \quad \lambda_1 \neq \lambda_2.$$

Имеем:

$$(Ae_1, e_2) = (e_1, A^* e_2) = (e_1, Ae_2),$$

т. е.

$$\lambda_1(e_1, e_2) = \lambda_2(e_1, e_2)$$

или

$$(\lambda_1 - \lambda_2)(e_1, e_2) = 0.$$

Так как $\lambda_1 \neq \lambda_2$, то

$$(e_1, e_2) = 0.$$

З а м е ч а н и е. Из доказанной теоремы следует, что наглядно-геометрический смысл произвольного самосопряженного преобразования таков; в пространстве выделяется n попарно ортогональных направлений (собственных направлений). Каждому из этих направлений ставится в соответствие действительное число (собственное значение). По каждому из этих направлений производится растяжение (сжатие) пространства в $|\lambda_i|$ раз и, кроме того, зеркальное отражение в плоскости, ортогональной к данному направлению, если соответствующее λ_i отрицательно.

Параллельно с понятием самосопряженного преобразования вводится понятие эрмитовой матрицы.

Матрица $\|a_{ik}\|$ называется *эрмитовой,* если $a_{ik} = \overline{a}_{ki}$.

Ясно, что для того чтобы преобразование A было самосопряженным, необходимо и достаточно, чтобы его матрица в каком-нибудь ортогональном базисе была эрмитовой.

Упражнение. Возвести матрицу

$$\begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 1 \end{pmatrix}$$

в 28-ю степень. Указание. Привести эту матрицу к диагональной форме, затем возвести ее в указанную степень и, наконец, вернуться к прежнему базису.

2. Приведение к главным осям. Одновременное приведение пары квадратичных форм к сумме квадратов. Применим полученные в п. 1 результаты к квадратичным формам.

Мы знаем, что всякой эрмитовой билинейной форме соответствует самосопряженное линейное преобразование. Из теоремы 2 этого параграфа вытекает важная

Теорема 3. *Пусть R — евклидово n-мерное пространство и пусть $A(x; y)$ — эрмитова билинейная форма в R. Тогда в R существует ортогональный нормированный базис, в котором соответствующая $A(x; y)$ квадратичная форма записывается в виде суммы квадратов:*

$$A(x; x) = \sum \lambda_i |\xi_i|^2,$$

где λ_i вещественны, а ξ_i — координаты вектора x *).

*) В § 8 мы доказали, что в *аффинном* пространстве можно всякую квадратичную (или, что то же самое, всякую эрмитову билинейную) форму привести к сумме квадратов. Здесь мы для *евклидова* пространства доказываем *более сильное* утверждение, именно существование *нормированного ортогонального базиса,* в котором данная эрмитова форма приводится к сумме квадратов.

Доказательство. Если $A(x; y)$ — эрмитова билинейная форма, т. е.

$$A(x; y) = \overline{A(y; x)},$$

то (см. § 11) существует такое самосопряженное линейное преобразование A, что

$$A(x; y) \equiv (Ax, y).$$

Выберем в R в качестве векторов ортогонального нормированного базиса систему попарно ортогональных собственных векторов самосопряженного преобразования A (это возможно в силу теоремы 1). Тогда

$$Ae_1 = \lambda_1 e_1,\ Ae_2 = \lambda_2 e_2,\ \ldots,\ Ae_n = \lambda_n e_n.$$

Пусть

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

Так как

$$(e_i, e_k) = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k, \end{cases}$$

то

$$\begin{aligned} A(x; y) &\equiv (Ax, y) = \\ &= (\xi_1 Ae_1 + \xi_2 Ae_2 + \ldots + \xi_n Ae_n,\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n) = \\ &= (\lambda_1 \xi_1 e_1 + \lambda_2 \xi_2 e_2 + \ldots + \lambda_n \xi_n e_n,\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n) = \\ &= \lambda_1 \xi_1 \overline{\eta}_1 + \lambda_2 \xi_2 \overline{\eta}_2 + \ldots + \lambda_n \xi_n \overline{\eta}_n. \end{aligned}$$

В частности,

$$A(x; x) = (Ax, x) = \lambda_1 |\xi_1|^2 + \lambda_2 |\xi_2|^2 + \ldots + \lambda_n |\xi_n|^2.$$

Теорема доказана.

Нахождение в евклидовом пространстве ортогонального нормированного базиса, в котором данная квадратичная форма приводится к сумме квадратов, называется приведением этой формы к главным осям.

Т е о р е м а 4. *Пусть R — аффинное n-мерное пространство и $A(x;x)$ и $B(x;x)$ — две эрмитовы квадратичные формы, причем форма $B(x;x)$ — положительно определенная. Тогда существует базис, в котором обе эти формы записываются в виде суммы квадратов.*

Д о к а з а т е л ь с т в о. Введем в R скалярное произведение, положив $(x,y) \equiv B(x;y)$, где $B(x;y)$ — отвечающая $B(x;x)$ билинейная форма. Это является законным, так как аксиомы скалярного произведения означают, что (x,y) есть эрмитова билинейная форма, соответствующая положительно определенной квадратичной форме (§ 8). Пространство R станет, таким образом, евклидовым. Согласно теореме 3 в R существует ортогональный *) нормированный базис $e_1, e_2, \ldots, e_n$, в котором форма $A(x;x)$ приводится к сумме квадратов, т. е. к виду

$$A(x;x) = \lambda_1|\xi_1|^2 + \lambda_2|\xi_2|^2 + \ldots + \lambda_n|\xi_n|^2. \qquad (2)$$

В нормированном ортогональном базисе скалярное произведение имеет вид

$$(x,x) = |\xi_1|^2 + |\xi_2|^2 + \ldots + |\xi_n|^2,$$

т. е.

$$B(x;x) = |\xi_1|^2 + |\xi_2|^2 + \ldots + |\xi_n|^2. \qquad (3)$$

Мы нашли, таким образом, базис, в котором обе квадратичные формы $A(x;x)$ и $B(x;x)$ *одновременно* приводятся к сумме квадратов, что и требовалось.

В теореме 4 показано, что в R существует базис, в котором эрмитовы квадратичные формы A и B имеют вид (2) и (3). Покажем, как найти числа $\lambda_1, \lambda_2, \ldots, \lambda_n$.

---

*) Относительно введенного нами скалярного произведения $(x,y) = B(x;y)$.

В каноническом виде матрицы квадратичных форм A и B имеют вид

$$\boldsymbol{A} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Следовательно,

$$\operatorname{Det}(\boldsymbol{A} - \lambda \boldsymbol{B}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\dots(\lambda_n - \lambda). \qquad (4)$$

При переходе к другому базису матрицы эрмитовых квадратичных форм A и B переходят в $\boldsymbol{A}_1 = \boldsymbol{C}^*\boldsymbol{A}\boldsymbol{C}$ и $\boldsymbol{B}_1 = \boldsymbol{C}^*\boldsymbol{A}\boldsymbol{C}$. Поэтому, если $e_1, e_2, \dots, e_n$ — произвольный базис, то в этом базисе

$$\operatorname{Det}(\boldsymbol{A}_1 - \lambda \boldsymbol{B}_1) = \operatorname{Det} \boldsymbol{C}^* \cdot \operatorname{Det}(\boldsymbol{A} - \lambda \boldsymbol{B}) \cdot \operatorname{Det} \boldsymbol{C},$$

т. е. отличается лишь постоянным множителем от выражения (4). Отсюда следует, что *числа* $\lambda_1, \lambda_2, \dots, \lambda_n$ *являются корнями следующего уравнения*:

$$\begin{vmatrix} a_{11} - \lambda b_{11} & a_{12} - \lambda b_{12} & \dots & a_{1n} - \lambda b_{1n} \\ a_{21} - \lambda b_{21} & a_{22} - \lambda b_{22} & \dots & a_{2n} - \lambda b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} - \lambda b_{n1} & a_{n2} - \lambda b_{n2} & \dots & a_{nn} - \lambda b_{nn} \end{vmatrix} = 0,$$

где $\|a_{ik}\|$ и $\|b_{ik}\|$ — матрицы форм $A(x;x)$ и $B(x;x)$ в каком-нибудь базисе $e_1, e_2, \dots, e_n$.

Замечание. Требование положительной определенности одной из форм является существенным, о чем свидетельствует следующий пример: две квадратичные формы

$$A(x;x) = |\xi_1|^2 - |\xi_2|^2, \quad B(x;x) = \xi_1\overline{\xi}_2 + \xi_2\overline{\xi}_1,$$

из которых ни одна не является положительно определенной, не могут быть одновременно приведены к сумме квадратов. В самом деле, первой форме соответствует матрица

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

а второй — матрица
$$\boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
Рассмотрим матрицу $\boldsymbol{A} - \lambda \boldsymbol{B}$, где λ — вещественный параметр. Ее детерминант равен $-(\lambda^2 + 1)$. Так как он не имеет вещественных корней, то, согласно сказанному выше, обе формы не могут быть приведены одновременно к сумме квадратов.

§ 13. Унитарные преобразования

Мы определили в § 11 унитарные преобразования равенством
$$UU^* = U^*U = E. \tag{1}$$
Это определение имеет простой геометрический смысл. А именно:

Всякое унитарное преобразование U в евклидовом n-мерном пространстве R сохраняет скалярное произведение, т. е.
$$(Ux, Uy) = (x, y)$$
для всех $x, y \in R$. Обратно, всякое линейное преобразование U, сохраняющее скалярное произведение, унитарно [*т. е. удовлетворяет условию* (1)].

В самом деле, если дано, что $U^*U = E$, то
$$(Ux, Uy) = (x, U^*Uy) = (x, y).$$

Обратно, если для любых векторов x и y
$$(Ux, Uy) = (x, y),$$
то
$$(U^*Ux, y) = (x, y),$$
т. е.
$$(U^*Ux, y) = (Ex, y).$$

Так как из равенства билинейных форм следует равенство соответствующих преобразований, то $U^*U = E$, т. е. U унитарно.

В частности, при $x = y$ имеем:

$$(Ux, Ux) = (x, x),$$

т. е. *унитарное преобразование U не меняет длин векторов.*

Упражнение. Доказать, что если линейное преобразование сохраняет длины всех векторов, то оно унитарно.

Запишем условия унитарности линейного преобразования в матричной форме. Для этого выберем какой-либо ортогональный нормированный базис $e_1, e_2, \ldots \ldots, e_n$. Пусть в этом базисе преобразованию U соответствует матрица

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}. \tag{2}$$

Тогда сопряженному преобразованию U^* соответствует матрица

$$\begin{pmatrix} \overline{a}_{11} & \overline{a}_{21} & \ldots & \overline{a}_{n1} \\ \overline{a}_{12} & \overline{a}_{22} & \ldots & \overline{a}_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ \overline{a}_{1n} & \overline{a}_{2n} & \ldots & \overline{a}_{nn} \end{pmatrix}. \tag{3}$$

Условие унитарности $UU^* = E$ означает, что произведение матриц (2) и (3) есть единичная матрица. Если перемножить их и приравнять элементы произведения соответственным элементам единичной матрицы, то получим:

$$\sum_{\alpha=1}^{n} a_{i\alpha}\overline{a}_{i\alpha} = 1, \quad \sum_{\alpha=1}^{n} a_{i\alpha}\overline{a}_{k\alpha} = 0 \quad (i \neq k). \tag{4}$$

Итак, *в ортогональном нормированном базисе условие $UU^* = E$ означает, что сумма произведений элементов какой-либо строки матрицы преобразования U на* *элементы, сопряженные к элементам другой строки, равна нулю, а сумма квадратов модулей элементов любой строки равна единице.*

Так как $U^*U = E$ также есть условие унитарности, то мы имеем также:

$$\sum_{\alpha=1}^{n} a_{\alpha i}\overline{a}_{\alpha i} = 1; \quad \sum_{\alpha=1}^{n} a_{\alpha i}\overline{a}_{\alpha k} = 0 \quad (i \neq k). \tag{5}$$

Это условие аналогично предыдущему, но вместо строк в нем участвуют столбцы матрицы.

Условие (5) имеет простой геометрический смысл. Действительно, скалярное произведение векторов

$$Ue_i = a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n$$

и

$$Ue_k = a_{1k}e_1 + a_{2k}e_2 + \ldots + a_{nk}e_n$$

равно $\sum a_{\alpha i}\overline{a}_{\alpha k}$ (так как $e_1, e_2, \ldots, e_n$ — это ортогональный нормированный базис); поэтому

$$(Ue_i, Ue_k) = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k. \end{cases} \tag{6}$$

Следовательно, *для того чтобы линейное преобразование было унитарным, необходимо и достаточно, чтобы оно переводило какой-либо ортогональный нормированный базис $e_1, e_2, \ldots, e_n$ снова в ортогональный и нормированный базис $Ue_1, Ue_2, \ldots, Ue_n$.*

Матрица $\|a_{ik}\|$, элементы которой удовлетворяют условиям (4), либо, что то же самое, условиям (5), называется *унитарной матрицей.* Унитарные матрицы являются, как мы видели, матрицами унитарных преобразований в *ортогональном нормированном* базисе. Так как переход от одного ортогонального нормированного к другому задается унитарным преобразованием, то матрица перехода от одного ортогонального нормированного базиса к другому такому же является унитарной.

Посмотрим, к какому простейшему виду можно привести матрицу унитарного преобразования при соответствующем выборе базиса.

Лемма 1. *Собственные значения унитарного преобразования по модулю равны* 1.

Доказательство. Пусть x — собственный вектор унитарного преобразования U и λ — соответствующее собственное значение, т. е.

$$Ux = \lambda x, \quad x \neq 0.$$

Тогда

$$(x, x) = (Ux, Ux) = (\lambda x, \lambda x) = \lambda\overline{\lambda}(x, x),$$

т. е. $\lambda\overline{\lambda} = 1$, значит $|\lambda| = 1$, что и требовалось доказать.

Лемма 2. *Пусть U — унитарное линейное преобразование в n-мерном пространстве R и e — его собственный вектор, т. е.*

$$Ue = \lambda e, \quad e \neq 0.$$

Тогда $(n-1)$-мерное подпространство R_1, состоящее из векторов x, ортогональных к e, инвариантно относительно U.

Доказательство. Пусть $x \in R_1$, т. е. $(x, e) = 0$. Покажем, что $Ux \in R_1$, т. е. что $(Ux, e) = 0$. В самом деле,

$$(Ux, Ue) = (U^*Ux, e) = (x, e) = 0.$$

А так как $Ue = \lambda e$, то $\overline{\lambda}(Ux, e) = 0$. Но в силу леммы 1 $\lambda \neq 0$, поэтому $(Ux, e) = 0$, т. е. $Ux \in R_1$. Следовательно, подпространство R_1 инвариантно относительно U.

Теорема 1. *Пусть U — унитарное преобразование в n-мерном евклидовом пространстве. Тогда* *существует n попарно ортогональных собственных векторов преобразования U. Соответствующие им собственные значения по модулю равны единице.*

Доказательство. В силу теоремы 1 § 10 преобразование U, как и всякое линейное преобразование, имеет в R хотя бы один собственный вектор. Обозначим его e_1. Согласно лемме 2, $(n-1)$-мерное подпространство R_1, состоящее из всех векторов пространства R, ортогональных к e_1, инвариантно относительно U. Следовательно, в R_1 имеется хотя бы один собственный вектор e_2 преобразования U. Через R_2 обозначим инвариантное подпространство, состоящее из всех векторов, принадлежащих R_1 и ортогональных к e_2. В R_2 содержится некоторый собственный вектор e_3 преобразования U и т. д.; продолжая этот процесс, мы построим n попарно ортогональных собственных векторов $e_1, e_2, \ldots, e_n$ преобразования U. Согласно лемме 1 собственные значения, соответствующие собственным векторам $e_1, e_2, \ldots, e_n$, по модулю равны 1.

Теорема 2. *Для каждого унитарного преобразования U в n-мерном пространстве R существует нормированный ортогональный базис, в котором матрица преобразования U диагональна, т. е. имеет вид:*

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \tag{7}$$

причем $\lambda_1, \lambda_2, \ldots, \lambda_n$ — числа, по модулю равные единице.

Доказательство. Пусть U — унитарное преобразование. Тогда n попарно ортогональных нормированных собственных векторов, построенных в предыдущей теореме, образуют искомый базис. Действительно,

$$\begin{aligned} Ue_1 &= \lambda_1 e_1, \\ Ue_2 &= \lambda_2 e_2, \\ &\ldots\ldots\ldots \\ Ue_n &= \lambda_n e_n \end{aligned}$$

и, следовательно, матрица преобразования U в базисе $e_1, e_2, \ldots, e_n$ имеет вид (7). Числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ по модулю равны 1 в силу леммы 1. Теорема доказана.

Упражнения. 1. Доказать, что верно и обратное, т. е. если в некотором ортогональном базисе матрица преобразования U имеет вид (7), то U унитарно.

2. Доказать, что если A — самосопряженное преобразование, то преобразование $(A-iE)^{-1}(A+iE)$ существует и является унитарным.

3. Пусть U — унитарное преобразование. Доказать, что если преобразование $U - E$ обратимо, то преобразование

$$A = i(U-E)^{-1}(U+E)$$

самосопряженное.

Так как матрица перехода от одного ортогонального нормированного базиса к другому задается унитарной матрицей, то полученный в этом параграфе результат мы можем в матричных терминах сформулировать следующим образом:

Пусть $\boldsymbol{U}$ — заданная унитарная матрица. Тогда существует такая унитарная матрица $\boldsymbol{V}$, что $\boldsymbol{U}$ представима в виде

$$\boldsymbol{U} = \boldsymbol{V}^{-1}\boldsymbol{D}\boldsymbol{V},$$

где $\boldsymbol{D}$ — диагональная матрица, у которой по диагонали стоят числа, по модулю равные 1.

Аналогично, основной результат в п. 1 § 12 в матричных терминах формулируется так:

Пусть $\boldsymbol{A}$ — заданная эрмитова матрица. Тогда $\boldsymbol{A}$ может быть представлена в виде

$$\boldsymbol{A} = \boldsymbol{V}^{-1}\boldsymbol{D}\boldsymbol{V},$$

где $\boldsymbol{V}$ — унитарная матрица, а $\boldsymbol{D}$ — диагональная матрица, у которой по диагонали стоят вещественные числа.

§ 14. Перестановочные линейные преобразования. Нормальные преобразования

1. Перестановочные преобразования. Мы видели (§ 12), что для всякого самосопряженного линейного преобразования есть свой ортогональный нормированный базис, в котором его матрица диагональна. Может оказаться что для нескольких самосопряженных преобразований существует один общий базис, в котором матрицы всех этих преобразований диагональны. Мы выясним здесь, при каких условиях это возможно. Разберем в первую очередь случай двух преобразований.

Лемма 1. *Пусть A и B — два перестановочных линейных преобразования, т. е.*

$$AB = BA.$$

Тогда совокупность всех собственных векторов преобразования A, отвечающих данному собственному значению λ, образует (вместе с нулевым вектором) подпространство R_λ, инвариантное относительно преобразования B.

Доказательство. Нам нужно показать, что если

$$x \in R_\lambda, \quad \text{т. е.} \quad Ax = \lambda x,$$

то и

$$Bx \in R_\lambda, \quad \text{т. е.} \quad ABx = \lambda Bx.$$

Но так как $AB = BA$, то

$$ABx = BAx = B\lambda x = \lambda Bx,$$

и лемма доказана.

Лемма 2. *Любые два перестановочных преобразования имеют общий собственный вектор.*

Доказательство. Пусть $AB = BA$ и R_λ — подпространство, состоящее из всех таких векторов x, что $Ax = \lambda x$, где λ — собственное значение преобразования A. Согласно лемме 1, R_λ инвариантно относительно B. Поэтому в нем существует вектор x_0, собственный для B. Этот вектор является собственным и для A, так как все векторы из R_λ являются собственными для A.

Замечание. Если $AB = BA$, то, вообще говоря, не всякий вектор, собственный для A, является собственным и для B. Например, если A есть единичное преобразование E, то для него любой вектор x является собственным. Однако x вовсе не будет собственным вектором для любого перестановочного с E преобразования, так как с E перестановочны все линейные преобразования.

Теорема 1. *Пусть A и B — два самосопряженных линейных преобразования в комплексном n-мерном пространстве R. Для того чтобы в R существовал ортогональный базис, в котором преобразования A и B одновременно приводятся к диагональной форме, необходимо и достаточно, чтобы они были перестановочны* (т. е. $AB = BA$).

Доказательство. Достаточность. Пусть $AB = BA$. Тогда, в силу леммы 2, существует вектор e_1, собственный и для A, и для B, т. е. такой, что

$$Ae_1 = \lambda_1 e_1, \quad Be_1 = \mu_1 e_1.$$

$(n-1)$-мерное подпространство R_1, ортогональное к e_1, инвариантно как для A, так и для B (см. лемму 2 § 12). Будем рассматривать преобразования A и B лишь в R_1. Согласно лемме 2 в R_1 существует вектор e_2, собственный и для A, и для B:

$$Ae_2 = \lambda_2 e_2, \quad Be_2 = \mu_2 e_2.$$

Совокупность векторов из R_1, ортогональных к e_2, образует $(n-2)$-мерное пространство, инвариантное как относительно A, так и относительно B, и т. д. Продолжая этот процесс, мы получим n попарно ортогональных векторов $e_1, e_2, \ldots, e_n$, собственных как для A, так и для B:

$$Ae_i = \lambda_i e_i, \quad Be_i = \mu_i e_i \quad (i = 1, \ldots, n).$$

Примем векторы $e_1, e_2, \ldots, e_n$ за базис в R. Тогда оба преобразования A и B запишутся в диагональной форме. Достаточность условия $AB = BA$ доказана.

Необходимость. Пусть в некотором ортогональном базисе матрицы преобразований A и B диагональны. Любые диагональные матрицы, как это легко проверить, перестановочны между собой. Но если матрицы преобразований в некотором базисе перестановочны, то перестановочны и сами преобразования.

Упражнение. Пусть U_1 и U_2 — перестановочные унитарные преобразования. Доказать, что существует базис, в котором они одновременно записываются в диагональной форме.

Замечание. Теорема 1 переносится на любое множество попарно перестановочных самосопряженных преобразований. Доказательство повторяется дословно, только вместо леммы 2 используется следующая

Лемма 2′. *У любого множества попарно перестановочный линейных преобразований есть общий собственный вектор.*

Доказательство будем вести по индукции. В одномерном пространстве $(n = 1)$ лемма очевидна. Предположим, что для пространств размерности $< n$ лемма доказана и докажем ее для n-мерного пространства.

Если каждый вектор из R является собственным для каждого из рассматриваемых преобразований *) $A, B, C, \ldots$, то все доказано. Предположим поэтому, что хотя бы один вектор не является собственным для какого-либо из наших преобразований, например для A.

---

*) Это означает, что каждое из преобразований $A, B, C, \ldots$ кратно единичному преобразованию.

Обозначим через R_1 совокупность всех собственных векторов преобразования A, отвечающих какому-нибудь собственному значению λ. Согласно лемме 1, R_1 инвариантно относительно $B, C, \ldots$ (и, само собой разумеется, инвариантно относительно A). При этом R_1 есть подпространство, отличное от нулевого и от всего R и имеющее, следовательно, размерность $\leqslant n-1$. Так как по предположению для пространств размерности, меньшей чем n, теорема доказана, то в R_1 преобразования $A, B, C, \ldots$ имеют общий собственный вектор, и лемма доказана.

2. Нормальные преобразования. В §§ 12 и 13 мы ознакомились с двумя классами линейных преобразований, приводимых в некотором нормированном ортогональном базисе к диагональной форме. Сейчас мы выясним, каков общий вид всех таких преобразований.

Т е о р е м а 2. *Для того чтобы существовал ортогональный базис, в котором линейное преобразование A приводится к диагональной форме, необходимо и достаточно, чтобы*

$$AA^* = A^*A.$$

(*Такие преобразования мы назвали в § 11 нормальными.*)

Д о к а з а т е л ь с т в о. Н е о б х о д и м о с т ь. Пусть в некотором ортогональном нормированном базисе матрица преобразования A диагональна, т. е. имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Так как базис ортогональный и нормированный, то матрица преобразования A^* имеет вид

$$\begin{pmatrix} \overline{\lambda}_1 & 0 & \ldots & 0 \\ 0 & \overline{\lambda}_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \overline{\lambda}_n \end{pmatrix}.$$

Матрицы преобразований A и A^* диагональны и, значит, перестановочны между собой. Следовательно, перестановочны и сами преобразования A и A^*.

Достаточность. Пусть A и A^* перестановочны. Тогда, согласно лемме 2 этого параграфа, у A и A^* существует общий собственный вектор e_1, т.е.

$$Ae_1 = \lambda_1 e_1, \quad A^* e_1 = \mu_1 e_1\ ^{*)}.$$

$(n-1)$-мерное подпространство R_1, состоящее из векторов, ортогональных к e_1, инвариантно как относительно A, так и относительно A^*. Действительно, пусть $x \in R_1$, т.е. $(x, e_1) = 0$. Тогда

$$(Ax, e_1) = (x, A^* e_1) = (x, \mu_1 e_1) = \overline{\mu}_1 (x, e_1) = 0,$$

т.е. $x \in R_1$. Инвариантность R_1 относительно A доказана. Аналогично доказывается инвариантность R_1 относительно A^*.

Применяя к R_1 ту же лемму 2, получим, что в R_1 существует вектор e_2, собственный одновременно и для A, и для A^*. Через R_2 обозначим $(n-2)$-мерное подпространство, состоящее из векторов подпространства R_1, ортогональных к e_2, и т.д. Продолжая таким образом, мы построим n попарно ортогональных векторов $e_1, e_2, \ldots, e_n$, каждый из которых является собственным как для A, так и для A^*. Векторы $e_1, e_2, \ldots, e_n$ образуют ортогональный базис, в котором как A, так и A^* приводятся к диагональной форме.

Другое доказательство достаточности. Положим

$$A_1 = \frac{A + A^*}{2}, \quad A_2 = \frac{A - A^*}{2i}.$$

Преобразования A_1 и A_2 — самосопряженные. Если A и A^* перестановочны, то A_1 и A_2 также перестановоч-

---

*) Упражнение: доказать, что $\mu_1 = \overline{\lambda}_1$.

ны. В силу теоремы 1 настоящего параграфа преобразования A_1 и A_2 могут быть одновременно приведены к диагональной форме. Но тогда и $A = A_1 + iA_2$ также записывается в диагональной форме.

Если A — самосопряженное преобразование, то

$$AA^* = A^*A = A^2,$$

т. е. A нормально. Нормальным является также всякое унитарное преобразование, так как в этом случае $UU^* = U^*U = E$. Поэтому теорема 2 этого параграфа содержит как частный случай результаты § 12 (п. 1) и § 13.

Упражнения. 1. Доказать, что любое множество попарно перестановочных нормальных преобразований приводится одновременно к диагональной форме.

2. Доказать, что всякое нормальное преобразование A может быть записано в виде

$$A = HU = UH,$$

где H — самосопряженное преобразование, а U — унитарное, причем H и U перестановочны.

Указание. Выбрать базис, в котором A и A^* приводятся к диагональной форме.

3. Доказать, что если $A = HU$, где H и U перестановочны, H — эрмитово, U — унитарно, то A — нормальное преобразование.

§ 15. Разложение линейного преобразования в произведение унитарного и эрмитова

Всякое комплексное число можно разложить в произведение положительного числа и числа, по модулю равного единице (так называемая тригонометрическая форма комплексного числа). Мы хотим получить для линейных преобразований аналог такого разложения.

Аналогом чисел, по модулю равных единице, являются унитарные линейные преобразования. Аналогом положительных действительных чисел являются так называемые положительно определенные линейные преобразования.

Определение 1. *Линейное преобразование H называется положительно определенным, если H самосопряженно и $(Hx, x) \geqslant 0$ для любого x.*

Теорема 1. *Всякое невырожденное линейное преобразование A может быть представлено в виде*

$$A = HU \quad (\text{либо } A = U_1 H_1),$$

где H (соотв. H_1) — невырожденное положительно определенное, а U (соотв. U_1) — унитарное преобразование.

Само доказательство мы проведем несколько позже: сейчас мы выясним, как по A найти соответствующие H и U, если указанное в теореме 1 разложение возможно; это подскажет нам путь для доказательства теоремы.

Пусть $A = HU$, где H — положительно определенное невырожденное преобразование, а U — унитарное преобразование. H легко выразить через A; в самом деле,

$$A^* = U^* H^* = U^{-1} H,$$

откуда

$$AA^* = H^2.$$

Следовательно, для того чтобы найти H, нужно «извлечь квадратный корень» из AA^*. Зная A и H, легко получить и U, полагая $U = H^{-1}A$.

Доказательству теоремы 1 предпошлем три леммы.

Лемма 1. *Каково бы ни было линейное преобразование A, преобразование AA^* — положительно определенное. Если A не вырождено, то AA^* также не вырождено.*

Доказательство. Преобразование AA^* положительно определенное. Действительно:

$$(AA^*x, x) = (A^*x, A^*x) \geqslant 0$$

для любого x. Кроме того,

$$(AA^*)^* = A^{**}A^* = AA^*,$$

т.е. AA^* — самосопряженное преобразование.

Если преобразование A не вырождено, то детерминант матрицы $\|a_{ik}\|$ преобразования A в любом ортогональном базисе не равен нулю. Детерминант матрицы $\|\overline{a}_{ik}\|$ преобразования A^* в том же базисе является комплексно сопряженным к детерминанту матрицы $\|a_{ik}\|$ и, следовательно, также не равен нулю. Поэтому в данном случае детерминант матрицы, соответствующий преобразованию AA^*, не равен нулю, а это означает, что преобразование AA^* — невырожденное.

Лемма 2. *Если B — положительно определенное преобразование, то его собственные значения неотрицательны. Обратно, если все собственные значения самосопряженного преобразования B неотрицательны, то B — положительно определенное преобразование.*

Доказательство. Пусть B положительно определено и $Be = \lambda e$. Тогда

$$(Be, e) = \lambda(e, e)$$

и так как $(Be, e) \geqslant 0$ и $(e, e) > 0$, то $\lambda \geqslant 0$.

Обратно, пусть все собственные значения самосопряженного преобразования B неотрицательны, $e_1, e_2, \ldots, e_n$ — нормированный ортогональный базис, состоящий из собственных векторов преобразования B. Пусть

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

— произвольный вектор из R. Тогда

$$\begin{aligned}(Bx, x) &= \\ &= (\xi_1 Be_1 + \xi_2 Be_2 + \ldots + \xi_n Be_n,\ \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) = \\ &= (\xi_1 \lambda_1 e_1 + \xi_2 \lambda_2 e_2 + \ldots + \xi_n \lambda_n e_n,\ \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) = \\ &= \lambda_1 |\xi_1|^2 + \lambda_2 |\xi_2|^2 + \ldots + \lambda_n |\xi_n|^2 \quad (1)\end{aligned}$$

и так как все λ_i неотрицательны, то $(Bx, x) \geqslant 0$.

Замечание. Из равенства (1) непосредственно видно, что если все λ_i положительны, то преобразование B не вырождено и обратно.

Лемма 3. *Если B — положительно определенное преобразование, то существует такое положительно определенное преобразование H, что $H^2 = B$ (мы запишем это так: $H = \sqrt{B} = B^{\frac{1}{2}}$). При этом, если B не вырождено, то и H не вырождено.*

Доказательство. Выберем в R ортогональный базис, в котором B записывается в диагональной форме:

$$B = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения B. Согласно лемме 2 все $\lambda_i \geqslant 0$. Положим

$$H = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \ldots & 0 \\ 0 & \sqrt{\lambda_2} & \ldots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \ldots & \sqrt{\lambda_n} \end{pmatrix},$$

где числа $\sqrt{\lambda_i}$ выбираются неотрицательными. В силу той же леммы 2 преобразование H положительно определено. При этом, если B не вырождено, то (см. замечание к лемме 2) $\lambda_i > 0$, значит, и $\sqrt{\lambda_i} > 0$ и, следовательно, H не вырождено.

Перейдем теперь к доказательству теоремы 1. Пусть A — любое невырожденное линейное преобразование. Положим

$$H = \sqrt{AA^*}.$$

В силу доказанных лемм 1 и 3, H есть невырожденное положительно определенное преобразование. Положим, далее,

$$U = H^{-1}A. \tag{2}$$

Преобразование U унитарно. В самом деле,

$$UU^* = H^{-1}A(H^{-1}A)^* = H^{-1}AA^*H^{-1} = H^{-1}H^2H^{-1} = E.$$

Из (2) следует,что $A = HU$. Теорема доказана.

Операцию извлечения квадратного корня, примененную в этом параграфе, можно использовать для доказательства следующего предложения:

Пусть A и B — два самосопряженных преобразования, причем A — невырожденное положительно определенное преобразование. Тогда собственные значения преобразования AB вещественны.

Доказательство. Мы знаем, что характеристические многочлены (а значит, и собственные значения) преобразований

$$X = AB \quad \text{и} \quad C^{-1}XC$$

совпадают. Положим $C = A^{\frac{1}{2}}$. Тогда

$$C^{-1}XC = A^{-\frac{1}{2}}ABA^{\frac{1}{2}} = A^{\frac{1}{2}}BA^{\frac{1}{2}}.$$

Легко видеть, что полученное преобразование является самосопряженным:

$$\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^* = \left(A^{\frac{1}{2}}\right)^*B^*\left(A^{\frac{1}{2}}\right)^* = A^{\frac{1}{2}}BA^{\frac{1}{2}}.$$

Утверждение доказано.

Упражнение. Доказать, что если A и B — положительно определенные преобразования, из которых одно не вырождено, то преобразование AB имеет неотрицательные собственные значения.

§ 16. Линейные преобразования в вещественном евклидовом пространстве

В этом параграфе мы будем заниматься линейными преобразованиями в вещественном пространстве. Из материала данной главы для этого достаточно знать содержание §§ 9–11.

1. Все определения § 10, а именно определения инвариантного подпространства, собственного вектора, собственного значения, были введены для линейного пространства над произвольным полем и поэтому имеют смысл также и для вещественного линейного пространства.

Существенную роль во всей теории играла доказанная в § 10 теорема о том, что в комплексном пространстве всякое линейное преобразование имеет собственный вектор (одномерное инвариантное подпространство). В случае вещественного пространства эта теорема неверна. Например, поворот плоскости вокруг начала координат на угол, отличный от $k\pi$, представляет собой линейное преобразование, не имеющее ни одного одномерного инвариантного подпространства. Однако имеет место следующая теорема:

Теорема 1. *У всякого линейного преобразования A в вещественном линейном пространстве R существует одномерное или двумерное инвариантное подпространство.*

Доказательство. Выберем в R базис $e_1, e_2, \ldots, e_n$. Преобразованию A отвечает в этом базисе матрица $\|a_{ik}\|$.

Рассмотрим систему уравнений

$$\left.\begin{array}{l} a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n = \lambda\xi_1, \\ a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n = \lambda\xi_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n = \lambda\xi_n \end{array}\right\} \qquad (1)$$

и будем искать для нее ненулевое решение $\xi_1^0, \xi_2^0, \ldots, \xi_n^0$. Такое решение существует тогда и только тогда, когда определитель

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix}$$

равен нулю. Приравняв его нулю, мы получим уравнение n-й степени относительно λ с вещественными коэффициентами. Пусть λ_0 есть корень этого уравнения. Возможны два случая:

a) λ_0 есть вещественный корень этого уравнения. Тогда можно найти вещественные не все равные нулю числа $\xi_1^0, \xi_2^0, \ldots, \xi_n^0$, являющиеся решением системы (1). Считая их координатами некоторого вектора x в базисе $e_1, e_2, \ldots, e_n$, мы можем систему (1) переписать в виде

$$Ax = \lambda_0 x,$$

т. е. x порождает одномерное инвариантное подпространство.

b) $\lambda_0 = \alpha + i\beta$, т. е. λ_0 комплексно. Пусть

$$\xi_1 + i\eta_1,\ \xi_2 + i\eta_2,\ \ldots,\ \xi_n + i\eta_n$$

есть решение системы (1); подставляя эти числа вместо $\xi_1, \xi_2, \ldots, \xi_n$ в (1) и отделяя вещественную часть от мнимой, мы получаем:

$$\left.\begin{aligned} a_{11}\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n &= \alpha\xi_1 - \beta\eta_1, \\ a_{21}\xi_1 + a_{22}\xi_2 + \ldots + a_{2n}\xi_n &= \alpha\xi_2 - \beta\eta_2, \\ \ldots\ldots\ldots\ldots\ldots&\ldots\ldots\ldots \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + a_{nn}\xi_n &= \alpha\xi_n - \beta\eta_n \end{aligned}\right\} \qquad (2)$$

и соответственно

$$\left.\begin{array}{r} a_{11}\eta_1 + a_{12}\eta_2 + \ldots + a_{1n}\eta_n = \alpha\eta_1 + \beta\xi_1, \\ a_{21}\eta_1 + a_{22}\eta_2 + \ldots + a_{2n}\eta_n = \alpha\eta_2 + \beta\xi_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\eta_1 + a_{n2}\eta_2 + \ldots + a_{nn}\eta_n = \alpha\eta_n + \beta\xi_n. \end{array}\right\} \qquad (2')$$

Будем теперь $\xi_1, \xi_2, \ldots, \xi_n$ (соотв. $\eta_1, \eta_2, \ldots, \eta_n$) считать координатами некоторого вектора x (соотв. y) в R; тогда соотношения (2) и (2′) можно записать следующим образом:

$$Ax = \alpha x - \beta y; \quad Ay = \alpha y + \beta x. \qquad (3)$$

Равенства (3) означают, что двумерное подпространство, порожденное векторами x и y, инвариантно относительно A.

В дальнейшем мы будем пользоваться тем, что в двумерном инвариантном подпространстве, отвечающем корню $\lambda = \alpha + i\beta$, преобразование имеет вид (3).

Упражнение. Доказать, что в вещественном пространстве нечетного числа измерений (в частности в трехмерном) у каждого линейного преобразования есть одномерное инвариантное подпространство.

2. Самосопряженные преобразования.

Определение 1. *Линейное преобразование A в вещественном евклидовом пространстве R называется самосопряженным, если для любых векторов x и y*

$$(Ax, y) = (x, Ay). \qquad (4)$$

Пусть $e_1, e_2, \ldots, e_n$ — ортогональный нормированный базис в R и

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

Пусть, далее, ζ_i — координаты вектора $z = Ax$, т. е.

$$\zeta_i = \sum_{k=1}^{n} a_{ik}\xi_k,$$

где $\|a_{ik}\|$ — матрица преобразования A в базисе e_1, $e_2, \ldots, e_n$. Следовательно,

$$(Ax, y) = (z, y) = \sum_{i=1}^{n} \zeta_i \eta_i = \sum_{i,k=1}^{n} a_{ik}\xi_k\eta_i.$$

Аналогично,

$$(x, Ay) = \sum_{i,k=1}^{n} a_{ik}\xi_i\eta_k. \tag{5}$$

Таким образом, условие (4) означает, что

$$a_{ik} = a_{ki}.$$

Итак, *для того чтобы преобразование было самосопряженным, необходимо и достаточно, чтобы в ортогональном нормированном базисе его матрица была симметрична.*

Всякая симметрическая билинейная форма $A(x; y)$ в произвольном базисе имеет вид

$$A(x; y) = \sum_{i,k=1}^{n} a_{ik}\xi_i\eta_k, \tag{6}$$

где $a_{ik} = a_{ki}$. Сравнивая (5) и (6), получаем следующий результат, который мы используем для доказательства теоремы 3 этого параграфа:

Для всякой симметрической билинейной формы $A(x; y)$ существует такое самосопряженное преобразование A, что

$$A(x; y) = (Ax, y).$$

Покажем, что для каждого самосопряженного преобразования существует ортогональный базис, в котором матрица этого преобразования диагональна.

Доказательство будет основано на содержании п. 1. Другое доказательство, не зависящее от п. 1 (и, следовательно, теоремы существования корня алгебраического уравнения), см. в § 17.

Предварительно докажем следующие леммы.

Лемма 1. *У всякого самосопряженного преобразования существует одномерное инвариантное подпространство.*

Доказательство. Согласно теореме 1 этого параграфа каждому корню λ характеристического уравнения отвечает одномерное инвариантное подпространство, если λ вещественно, и двумерное — если λ комплексно. Поэтому, для доказательства леммы достаточно показать, что все λ вещественны.

Предположим, что λ комплексно. При доказательстве теоремы 1 мы для такого $\lambda = \alpha + i\beta$ построили два таких вектора x и y, что

$$Ax = \alpha x - \beta y,$$
$$Ay = \beta x + \alpha y.$$

Но тогда

$$(Ax, y) = \alpha(x, y) - \beta(y, y),$$
$$(x, Ay) = \beta(x, x) + \alpha(x, y).$$

Так как $(Ax, y) = (x, Ay)$ то, вычитая из второго равенства первое, имеем:

$$0 = \beta[(x, x) + (y, y)]$$

и так как $(x, x) + (y, y) \neq 0$, то $\beta = 0$, т. е. λ вещественно.

Лемма 2. *Пусть A — самосопряженное преобразование, а e — его собственный вектор. Тогда совокупность R' векторов, ортогональных e, образует $(n-1)$-мерное инвариантное подпространство.*

Доказательство. Ясно, что совокупность R' векторов $x \in R$, ортогональных вектору e, есть $(n-1)$-мерное подпространство. Покажем, что R' инвариантно относительно преобразования A.

Пусть $x \in R'$, т.е. $(x, e) = 0$. Тогда

$$(Ax, e) = (x, Ae) = (x, \lambda e) = \lambda(x, e) = 0,$$

т.е. и $Ax \in R'$.

Теорема 2. *Существует ортогональный нормированный базис, в котором матрица самосопряженного преобразования A диагональна.*

Доказательство. Согласно лемме 1 преобразование A имеет хотя бы один собственный вектор e_1.

Обозначим через R' подпространство, состоящее из векторов, ортогональных e_1. Так как R' инвариантно, то в нем, согласно той же лемме 1, также существует собственный вектор; обозначим его e_2. Продолжая это построение, мы получим n собственных векторов, из которых каждый следующий по построению ортогонален к предыдущим, т.е. получим n попарно ортогональных собственных векторов $e_1, e_2, \ldots, e_n$.

Выберем их за базис в R. Так как

$$Ae_i = \lambda_i e_i \quad (i = 1, 2, \ldots, n),$$

то матрица преобразования A в этом базисе имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

т.е. является диагональной матрицей.

3. Приведение квадратичной формы в ортогональном базисе к сумме квадратов. (Приведение к главным осям.) Пусть в n-мерном евклидовом пространстве задана симметрическая билинейная форма $A(x; y)$. Как было показано выше, каждой симметрической билинейной форме $A(x;y)$ соответствует такое линейное самосопряженное преобразование A, что $A(x;y) = (Ax, y)$.

Согласно теореме 2 этого параграфа, существует ортогональный нормированный базис $e_1, e_2, \ldots, e_n$, состоящий из собственных векторов преобразования A (т. е. такой, что $Ae_i = \lambda_i e_i$). В этом базисе мы имеем:

$$\begin{aligned} A(x;y) &= (Ax, y) = \\ &= (A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n),\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n) = \\ &= (\lambda_1 \xi_1 e_1 + \lambda_2 \xi_2 e_2 + \ldots + \lambda_n \xi_n e_n,\ \eta_1 e_1 + \eta_2 e_2 + \ldots + \eta_n e_n) = \\ &= \lambda_1 \xi_1 \eta_1 + \lambda_2 \xi_2 \eta_2 + \ldots + \lambda_n \xi_n \eta_n. \end{aligned}$$

Полагая $y = x$, получаем следующую теорему:

Теорема 3. *Пусть $A(x;x)$ — квадратичная форма в n-мерном евклидовом пространстве. Тогда существует ортогональный нормированный базис, в котором эта квадратичная форма имеет вид:*

$$A(x;x) = \sum \lambda_i \xi_i^2.$$

Так как $\lambda_1, \lambda_2, \ldots, \lambda_n$ являются собственными значениями A, то они могут быть найдены из характеристического уравнения матрицы $\|a_{ik}\|$.

Для случая трехмерного пространства доказанная здесь теорема рассматривается в аналитической геометрии. Действительно, в этом случае уравнение

$$A(x;x) = 1$$

есть уравнение центральной поверхности второго порядка. Ортогональный нормированный базис, о котором идет речь в теореме 3, есть в этом случае система координат, в которой поверхность имеет канонический вид, а векторы e_1, e_2, e_3 являются направлениями главных осей поверхности второго порядка.

4. Одновременное приведение пары квадратичных форм к сумме квадратов.

Теорема 4. *Пусть в n-мерном пространстве R заданы две квадратичные формы $A(x;x)$ и $B(x;x)$, причем форма $B(x;x)$ положительно определенная. Тогда* *в R существует базис, в котором обе эти квадратичные формы записываются в виде суммы квадратов.*

Доказательство. Пусть $B(x;y)$ — билинейная форма, соответствующая квадратичной форме $B(x;x)$. Определим в R скалярное произведение формулой

$$(x,y) = B(x;y).$$

Согласно предыдущей теореме в R существует нормированный ортогональный *) базис $e_1, e_2, \ldots, e_n$, в котором форма приводится к сумме квадратов, т. е.

$$A(x;x) = \sum_{i=1}^{n} \lambda_i \xi_i^2. \tag{7}$$

Скалярный квадрат в нормированном ортогональном базисе имеет вид:

$$(x,x) = B(x;x) = \sum_{i=1}^{n} \xi_i^2. \tag{8}$$

Итак, в базисе $e_1, e_2, \ldots, e_n$ обе квадратичные формы записываются в виде суммы квадратов. Теорема доказана.

5. Ортогональные преобразования.

Определение 2. *Линейное преобразование A вещественного n-мерного евклидова пространства называется ортогональным преобразованием, если оно сохраняет скалярное произведение векторов, т. е.*

$$(Ax, Ay) = (x, y) \tag{9}$$

для всех $x, y \in R$.

Полагая в равенстве (9) $x = y$, получаем

$$|Ax|^2 = |x|^2, \tag{10}$$

*) В смысле определенного нами в R скалярного произведения.

т. е. *ортогональное преобразование сохраняет длины векторов.*

Упражнение. Доказать, что условие (10) является достаточным условием ортогональности линейного преобразования.

Так как

$$\cos\varphi = \frac{(x, y)}{|x||y|}$$

и так как и числитель, и знаменатель в этом выражении не меняются при ортогональном преобразовании, то ортогональное преобразование сохраняет углы между векторами.

Пусть $e_1, e_2, \ldots, e_n$ — ортогональный нормированный базис. Так как преобразование A сохраняет углы между векторами и их длины, то векторы $Ae_1, Ae_2, \ldots$ $\ldots, Ae_n$ также образуют ортогональный нормированный базис, т. е.

$$(Ae_i, Ae_k) = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k. \end{cases} \tag{11}$$

Пусть теперь $\|a_{ik}\|$ — матрица преобразования в ортогональном нормированном базисе $e_1, e_2, \ldots, e_n$. Так как столбцы этой матрицы являются координатами векторов Ae_i, то условие (11) записывается следующим образом:

$$\sum_{\alpha=1}^{n} a_{\alpha i} a_{\alpha k} = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k. \end{cases} \tag{12}$$

Упражнение. Показать, что условия (11), а следовательно, и условия (12), являются достаточными условиями ортогональности преобразования.

Условия (12) можно записать в матричной форме. Действительно, $\sum\limits_{\alpha=1}^{n} a_{\alpha i} a_{\alpha k}$ суть элементы произведения матрицы на ее транспонированную. Поэтому условия

(12) означают, что *произведение матрицы на ее транспонированную есть единичная матрица.* Так как определитель произведения матриц равен произведению их определителей, то мы получаем также, что квадрат определителя матрицы ортогонального преобразования равен единице, т. е. что *определитель матрицы ортогонального преобразования равен* ± 1.

Ортогональные преобразования, определитель которых равен $+1$, называются *собственными,* а те, определитель которых равен -1, называются *несобственными.*

Докажем следующее существенное свойство ортогональных преобразований.

Лемма 3. *Если* R_1 — *подпространство пространства* R, *инвариантное относительно ортогонального линейного преобразования* A, *то его ортогональное дополнение* R_2, *т. е. совокупность всех векторов, ортогональных к каждому* $x \in R_1$ (см. упражнение на стр. 52), *также есть инвариантное подпространство.*

Доказательство. Пусть $y \in R_2$, т. е. $(x, y) = 0$ для всякого $x \in R_1$.

Покажем, что при этом $(x, Ay) = 0$ для всякого $x \in R_1$. Так как преобразование A ортогонально, то оно невырождено и его образ на любом инвариантном подпространстве совпадает с этим подпространством. Поэтому всякое $x \in R_1$ представимо в виде

$$x = Az, \quad \text{где} \quad z \in R_1.$$

Отсюда $(x, Ay) = (Az, Ay) = (z, y) = 0$, т. е. $Ay \in R_2$. Лемма доказана.

Упражнение. Доказать, что произведение двух собственных или двух несобственных ортогональных преобразований есть собственное ортогональное преобразование, а произведение собственного на несобственное есть несобственное ортогональное преобразование.

З а м е ч а н и е. Разделение ортогональных преобразований на собственные и несобственные связано с тем, что ортогональное преобразование, которое можно получить непрерывным переходом от единичного преобразования, обязательно собственно. Действительно, пусть A_t — ортогональное преобразование, непрерывно зависящее от параметра t (это значит, что элементы матрицы этого преобразования в каком-либо базисе непрерывно зависят от t), причем $A_0 = E$. Тогда определитель этого преобразования есть также непрерывная функция от t. Так как непрерывная функция, принимающая лишь значения ± 1, должна быть постоянной, а при $t = 0$ определитель преобразования A_t равен 1, то и при $t \neq 0$ определитель преобразования A_t равен 1. Используя теорему 5 этого параграфа, можно показать и обратное, а именно, что всякое собственное ортогональное преобразование может быть получено непрерывным изменением из единичного.

Изучим ортогональные преобразования в одномерном и двумерном пространствах. Ниже мы увидим, что изучение ортогональных преобразований в пространстве любого числа измерений сводится к этим простейшим случаям.

Пусть e — вектор, порождающий одномерное пространство, в котором задано ортогональное преобразование A. Тогда $Ae = \lambda e$ и так как в силу ортогональности $(Ae, Ae) = (e, e)$, то $\lambda^2(e, e) = (e, e)$, т. е. $\lambda = \pm 1$.

Таким образом, в одномерном пространстве есть лишь два ортогональных преобразования: преобразование $Ax \equiv x$ и преобразование $Ax \equiv -x$; первое из них собственное, а второе — несобственное.

Рассмотрим ортогональное преобразование в двумерном пространстве R. Пусть e_1, e_2 — ортогональный нормированный базис в R. Пусть далее преобразование A в этом базисе задается матрицей

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{13}$$

Разберем сначала случай собственного ортогонального преобразования, т. е. положим $\alpha\delta - \beta\gamma = 1$.

Условие ортогональности преобразования означает, что произведение матрицы (13) на ее транспонированную есть единичная матрица, т. е. что

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^{-1} = \begin{pmatrix}\alpha & \gamma\\ \beta & \delta\end{pmatrix}. \tag{14}$$

Так как определитель матрицы (13) равен единице, то

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^{-1} = \begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}. \tag{15}$$

Из (14) и (15) следует, что матрица преобразования в этом случае имеет вид

$$\begin{pmatrix}\alpha & -\beta\\ \beta & \alpha\end{pmatrix},$$

где $\alpha^2+\beta^2=1$. Полагая $\alpha=\cos\varphi$, $\beta=\sin\varphi$, получаем что *всякое собственное ортогональное преобразование в двумерном пространстве имеет в ортогональном нормированном базисе матрицу вида*

$$\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

(поворот в плоскости на угол φ).

Пусть теперь A — несобственное преобразование, т. е. $\alpha\delta-\beta\gamma=-1$. В этом случае характеристическое уравнение матрицы имеет вид $\lambda^2-(\alpha+\delta)\lambda-1$ и, следовательно, имеет вещественные корни. Значит, у преобразования A имеется собственный вектор e, $Ae=\lambda e$. Из ортогональности преобразования следует, что $Ae=\pm e$. Но ортогональное преобразование не меняет углов между векторами и их длин. Значит, ортогональный e вектор e_1 переходит после преобразования в вектор, ортогональный $Ae=\pm e$, т. е. в $\pm e_1$. Итак, в базисе e, e_1 матрица преобразования A имеет вид

$$\begin{pmatrix}\pm1 & 0\\ 0 & \pm1\end{pmatrix}.$$

Так как определитель матрицы несобственного преобразования равен -1, то возможны лишь следующие канонические виды матрицы несобственного ортогонального преобразования в двумерном пространстве

$$\begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & +1 \end{pmatrix}.$$

(Зеркальные отражения относительно одной из осей.)

Найдем теперь простейший вид матрицы ортогонального преобразования для случая пространства любого числа измерений.

Т е о р е м а 5. *Пусть A — ортогональное преобразование в n-мерном евклидовом пространстве R. В R существует ортогональный нормированный базис $e_1, e_2, \ldots, e_n$, в котором матрица преобразования A имеет вид*

$$\begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& -1 &&&&&&& \\ &&&& \ddots &&&&&& \\ &&&&& -1 &&&&& \\ &&&&&& \cos\varphi_1 & -\sin\varphi_1 &&& \\ &&&&&& \sin\varphi_1 & \cos\varphi_1 &&& \\ &&&&&&&& \ddots && \\ &&&&&&&&& \cos\varphi_k & -\sin\varphi_k \\ &&&&&&&&& \sin\varphi_k & \cos\varphi_k \end{pmatrix}.$$

Все элементы, кроме выписанных, суть нули.

Д о к а з а т е л ь с т в о. Согласно теореме 1 этого параграфа в R можно выбрать либо одномерное, либо двумерное инвариантное подпространство $R^{(1)}$. Если существует одномерное инвариантное подпространство $R^{(1)}$, то обозначим через e_1 содержащийся в нем вектор единичной длины. Если же одномерного инвариантного подпространства нет, то возьмем двумерное и обозначим через e_1, e_2 его ортогональный нормированный базис.

В случае, если $R^{(1)}$ одномерно, преобразование A имеет в нем вид: $Ax = \pm x$.

В случае, если подпространство $R^{(1)}$ двумерно, наше преобразование является в нем собственным ортогональным преобразованием (так как в противном случаев в $R^{(1)}$ существовало бы одномерное инвариантное подпространство), и, следовательно, A имеет в $R^{(1)}$ матрицу

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Совокупность $\tilde{R}$ векторов, ортогональных ко всем векторам из $R^{(1)}$, согласно доказанной выше лемме, есть снова инвариантное подпространство.

В инвариантном подпространстве $\tilde{R}$ снова находим одномерное или двумерное инвариантное подпространство, выбираем в нем базис и так далее.

Мы получим таким образом n попарно ортогональных векторов длины 1. Примем их за базис в R. Тогда матрица преобразования в этом базисе будет иметь вид

$$\begin{pmatrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& -1 &&&&&&& \\
&&&& \ddots &&&&&& \\
&&&&& -1 &&&&& \\
&&&&&& \cos\varphi_1 & -\sin\varphi_1 &&& \\
&&&&&& \sin\varphi_1 & \cos\varphi_1 &&& \\
&&&&&&&& \ddots && \\
&&&&&&&&& \cos\varphi_k & -\sin\varphi_k \\
&&&&&&&&& \sin\varphi_k & \cos\varphi_k
\end{pmatrix},$$

где стоящие по диагонали ± 1 отвечают одномерным инвариантным подпространствам, а «клетки»

$$\begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix}$$

— двумерным инвариантным подпространствам. Теорема доказана.

Замечание. Назовем *простым вращением* собственное ортогональное преобразование, представляющее собой поворот в некоторой двумерной плоскости и оставляющее неизменным $(n-2)$-мерное подпространство, ортогональное к этой плоскости. Таким образом, простое вращение есть преобразование, матрица которого может быть приведена к виду

$$\begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \cos\varphi & -\sin\varphi &&& \\ &&& \sin\varphi & \cos\varphi &&& \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{pmatrix}.$$

Простым отражением мы назовем несобственное линейное преобразование, меняющее направление всех векторов, принадлежащих некоторому одномерному подпространству, на противоположное и оставляющее неизменными векторы его $(n-1)$-мерного ортогонального дополнения. Таким образом, простое отражение есть преобразование, матрица которого в некотором базисе имеет вид

$$\begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& -1 &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}.$$

Пользуясь результатом теоремы 5, нетрудно показать, что всякое ортогональное преобразование может быть представлено как произведение некоторого числа простых вращений и простых отражений. Доказательство предоставляется читателю.

§ 17. Экстремальные свойства собственных значений

Рассмотрим самосопряженное линейное преобразование A в n-мерном евклидовом пространстве. Мы покажем, что его собственные значения можно получить, рассматривая некоторую задачу на минимум, связанную с соответствующей A квадратичной формой (Ax, x). Это, в частности, позволит доказать существование собственных векторов, не пользуясь теоремой о существовании корня уравнения n-й степени. Эти экстремальные свойства полезны также при вычислении собственных значений. Мы рассмотрим сначала вещественное пространство, а затем перенесем полученные результаты на случай комплексного пространства.

Докажем сначала следующую лемму.

Лемма 1. *Пусть B — самосопряженное линейное преобразование в вещественном пространстве такое, что квадратичная форма (Bx, x) неотрицательна, т. е.*

$$(Bx, x) \geqslant 0 \quad \text{для любого} \quad x.$$

Тогда, если для некоторого $x = e$

$$(Be, e) = 0,$$

то и $Be = 0$.

Доказательство. Покажем, что для любого вектора h имеем $(Be, h) = 0$. Для этого положим $x = e + th$, где t — произвольное число, а h — вектор. Тогда имеем

$$(B(e + th), e + th) = \\ = (Be, e) + t(Be, h) + t(Bh, e) + t^2(Bh, h) \geqslant 0,$$

т. е. так как $(Bh, e) = (h, Be) = (Be, h)$ и $(Be, e) = 0$, то $2t(Be, h) + t^2(Bh, h) \geqslant 0$ для любых t. Отсюда следует, что $(Be, h) = 0$.

Действительно, функция $at + bt^2$ при $a \neq 0$ меняет знак в точке $t = 0$, мы же получили, что выражение

$$2t(Be, h) + t^2(Bh, h)$$

для любых t неотрицательно, следовательно,

$$(Be, h) = 0.$$

Так как h произвольно, то $Be = 0$, что и требовалось доказать.

Пусть A — некоторое самосопряженное линейное преобразование в n-мерном вещественном евклидовом пространстве. Соответствующую A квадратичную форму (Ax, x) будем рассматривать на *единичной сфере*, т. е. на множестве векторов x, для которых

$$(x, x) = 1.$$

Имеет место следующая теорема.

Теорема 1. *Пусть A — самосопряженное линейное преобразование. Тогда соответствующая A квадратичная форма (Ax, x) достигает на единичной сфере минимума λ_1. Вектор e_1, на котором этот минимум достигается, есть собственный вектор преобразования A, а значение минимума λ_1 — соответствующее собственное значение этого преобразования.*

Доказательство. Единичная сфера есть ограниченное замкнутое множество в n-мерном пространстве. Поэтому (Ax, x), как непрерывная на нем функция, достигает минимума в некоторой точке e_1. Обозначим этот минимум через λ_1. Тогда имеем

$$(Ax, x) \geqslant \lambda_1, \quad \text{если} \quad (x, x) = 1, \tag{1}$$

причем

$$(Ae_1, e_1) = \lambda_1, \quad \text{где} \quad (e_1, e_1) = 1.$$

Запишем неравенство (1) в виде

$$(Ax, x) \geqslant \lambda_1(x, x), \quad \text{где} \quad (x, x) = 1. \tag{2}$$

Оно справедливо для векторов длины единица. Так как при умножении x на некоторое число α как правая, так и левая части неравенства умножаются на α^2, то оно справедливо для векторов любой длины (поскольку любой вектор можно получить из вектора длины единица умножением его на некоторое число α).

Полученное неравенство можно переписать так:

$$(Ax - \lambda_1 x,\ x) \geqslant 0 \quad \text{для любых} \quad x,$$

причем для $x = e_1$ имеет место

$$(Ae_1 - \lambda_1 e_1,\ e_1) = 0.$$

Это значит, что преобразование $B = A - \lambda_1 E$ удовлетворяет условиям леммы 1. Отсюда, применяя эту лемму, получаем:

$$(A - \lambda_1 E)e_1 = 0, \quad \text{т. е.} \quad Ae_1 = \lambda_1 e_1.$$

Таким образом, e_1 является собственным вектором преобразования A, соответствующим собственному значению λ_1. Теорема доказана.

Для нахождения следующего собственного значения рассмотрим все векторы из R, ортогональные к собственному вектору e_1. Как было показано в п. 2 § 16 (лемма 2), эти векторы образуют $(n-1)$-мерное подпространство R_1, инвариантное относительно преобразования A.

Отыскивая минимум квадратичной формы (Ax, x), при условии $(x, x) = 1$, в этом подпространстве мы придем к новому собственному вектору e_2 и собственному значению λ_2.

Очевидно, что $\lambda_2 \geqslant \lambda_1$, так как минимум (Ax, x) во всем пространстве не может быть больше, чем минимум той же функции в подпространстве.

Следующий собственный вектор мы получим, решая ту же задачу в $(n-2)$-мерном инвариантном подпространстве, состоящем из векторов, ортогональных и e_1 и e_2. Значение минимума (Ax, x) в этом подпространстве будет третьим собственным значением.

Продолжая этот процесс, мы исчерпаем все n собственных значений и соответствующих им собственных векторов нашего преобразования.

Иногда бывает полезно определить второй, третий и т. д. собственные векторы преобразования из задачи на максимум или минимум непосредственно, не считая при этом известными предыдущие собственные векторы.

Пусть A — самосопряженное линейное преобразование. Обозначим через

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$$

его собственные значения, расположенные в возрастающем порядке, а через $e_1, e_2, \ldots, e_n$ — соответствующие им нормированные и ортогональные собственные векторы.

Покажем, что *если мы возьмем первые k собственных векторов*

$$e_1, e_2, \ldots, e_k$$

и порожденное ими подпространство S, то для любого вектора x из S имеет место неравенство

$$\lambda_1(x, x) \leqslant (Ax, x) \leqslant \lambda_k(x, x).$$

Действительно, пусть

$$x = \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_k e_k.$$

Так как $Ae_k = \lambda_k e_k$, $(e_k, e_k) = 1$, $(e_k, e_i) = 0$ при $i \neq k$, то

$$\begin{aligned}(Ax, x) &= (A(\xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_k e_k),\ \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_k e_k) = \\ &= (\lambda_1 \xi_1 e_1 + \lambda_2 \xi_2 e_2 + \ldots + \lambda_k \xi_k e_k,\ \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_k e_k) = \\ &= \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \ldots + \lambda_k \xi_k^2.\end{aligned}$$

Кроме того, так как векторы $e_1, e_2, \ldots, e_k$ ортогональны и нормированы, то

$$(x, x) = \xi_1^2 + \xi_2^2 + \ldots + \xi_k^2$$

и, следовательно,

$$(Ax, x) = \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \ldots + \lambda_k \xi_k^2 \geqslant \lambda_1(\xi_1^2 + \xi_2^2 + \ldots + \xi_k^2) = \lambda_1(x, x).$$

Аналогично
$$(Ax, x) \leqslant \lambda_k(x, x)$$
и, следовательно,
$$\lambda_1(x, x) \leqslant (Ax, x) \leqslant \lambda_k(x, x).$$

Пусть теперь R_k — произвольное подпространство $n - k + 1$ измерений. В § 7 (лемма п. 1) мы доказали, что если сумма размерностей двух подпространств n-мерного пространства превышает n, то существует отличный от нуля вектор, принадлежащий обоим подпространствам. Следовательно, так как $(n-k+1)+k > n$, то существует вектор x_0, принадлежащий как R_k, так и подпространству S, порожденному векторами $e_1, e_2, \ldots, e_k$. Мы можем считать, что длина его равна 1, т. е. что $(x_0, x_0) = 1$. Так как всюду в S, как мы уже доказали, $(Ax, x) \leqslant \lambda_k(x, x)$, то $(Ax_0, x_0) \leqslant \lambda_k$.

Итак, мы доказали, что в R_k существует вектор x_0 длины 1 такой, что
$$(Ax_0, x_0) \leqslant \lambda_k.$$
Но тогда и подавно, минимум (Ax, x), где x пробегает все векторы длины 1 из R_k, также меньше или равен λ_k.

Таким образом, *для любого $(n - k + 1)$-мерного подпространства R_k*
$$\min(Ax, x) \leqslant \lambda_k,$$
где $(x, x) = 1$ и $x \in R_k$.

Заметим, что среди подпространств R_k размерности $n - k + 1$ есть такое, для которого $\min(Ax, x)$, $x \in R_k$, $(x, x) = 1$, в точности равен λ_k. Таким подпространством является подпространство, состоящее из векторов, ортогональных первым $k - 1$ собственным векторам $e_1, e_2, \ldots, e_{k-1}$. Действительно, в этом параграфе мы доказали, что $\min\limits_{(x,x)=1}(Ax, x)$, распространенный по всем векторам, ортогональным первым $k - 1$ собственным векторам, равен λ_k.

Итак, мы доказали следующее утверждение:

Пусть R_k — некоторое $(n - k + 1)$-мерное подпространство пространства R. Тогда минимум (Ax, x) для всех x из R_k таких, что $(x, x) = 1$, меньше или равен λ_k. Подпространство R_k можно выбрать так, чтобы этот минимум равнялся λ_k.

Это утверждение можно записать следующей формулой:
$$\max_{R_k} \min_{\substack{(x,x)=1 \\ x \in R_k}} (Ax, x) = \lambda_k. \tag{3}$$

В этой формуле min берется по указанным векторам, а max по всевозможным подпространствам R_k размерности $n - k + 1$.

Из доказанной теоремы следует:

Пусть A — самосопряженное линейное преобразование, а B — положительно определенное линейное преобразование. Пусть $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ — собственные значения A, а $\mu_1 \leqslant \mu_2 \leqslant \ldots \ldots \leqslant \mu_n$ — собственные значения $A+B$; тогда $\lambda_k \leqslant \mu_k$.

Действительно, всюду

$$(Ax, x) \leqslant ((A+B)x, x).$$

Следовательно, в любом $(n-k+1)$-мерном подпространстве R_k имеет место неравенство:

$$\min_{\substack{(x,x)=1 \\ x \in R_k}} (Ax, x) \leqslant \min_{\substack{(x,x)=1 \\ x \in R_k}} ((A+B)x, x).$$

Значит, максимум левой части по всевозможным подпространствам R_k не превосходит максимума правой части. Так как в силу формулы максимум левой части равен λ_k, а максимум правой равен μ_k, то $\lambda_k \leqslant \mu_k$, что и требовалось доказать.

Перенесем полученные результаты на случай комплексного пространства.

Для этого нам придется заменить лишь лемму 1 следующей леммой.

Л е м м а 2. *Пусть B — самосопряженное преобразование в комплексном пространстве, и соответствующая ему эрмитова форма (Bx, x) не отрицательна, т. е.*

$$(Bx, x) \geqslant 0 \quad \text{для любых} \quad x.$$

Тогда, если для некоторого e $(Be, e) = 0$, то и $Be = 0$.

Д о к а з а т е л ь с т в о. Пусть t — произвольное вещественное число, а h — вектор. Тогда

$$(B(e+th), e+th) \geqslant 0,$$

или, так как $(Be, e) = 0$, то

$$t[(Be, h) + (Bh, e)] + t^2(Bh, h) \geqslant 0$$

для любого t. Отсюда следует, что

$$(Be, h) + (Bh, e) = 0. \tag{4}$$

Так как h произвольно, то, заменяя h на ih, получаем $(Be, ih) + (iBh, e) = 0$, т. е.

$$-i(Be, h) + i(Bh, e) = 0. \tag{5}$$

Из (4) и (5) получаем, что

$$(Be, h) = 0,$$

и так как h произвольно, то $Be = 0$. Лемма доказана.

Все остальные теоремы этого параграфа и их доказательства переносятся на случай комплексного пространства без всяких изменений.

ГЛАВА III

КАНОНИЧЕСКИЙ ВИД ПРОИЗВОЛЬНЫХ ЛИНЕЙНЫХ ПРЕОБРАЗОВАНИЙ

§ 18. Нормальная форма линейного преобразования

В главе II мы познакомились с различными классами линейных преобразований n-мерного пространства, имеющих n линейно независимых собственных векторов. Мы знаем, что в базисе, состоящем из собственных векторов такого преобразования, его матрица имеет особенно простой вид, так называемую диагональную форму.

Однако число линейно независимых собственных векторов у линейного преобразования может быть меньше, чем n *). Пример линейного преобразования с недостаточным числом собственных векторов мы приведем несколько позже (см. также § 10, п. 1, пример 3). Такое преобразование заведомо не может быть приведено к диагональной форме, так как базис, в котором матрица преобразования диагональна, состоит из соб-

*) Напомним, что если все корни характеристического многочлена различны, то преобразование имеет n линейно независимых собственных векторов. Поэтому, для того чтобы число собственных векторов было меньше чем n, необходимо наличие кратных корней у характеристического многочлена. Таким образом, этот случай является в некотором смысле исключительным.

ственных векторов. Возникает вопрос: каков простейший вид матрицы такого линейного преобразования?

В этой главе мы для произвольного преобразования укажем базис, в котором его матрица имеет сравнительно простой вид (так называемая *жорданова нормальная форма*). В случае, когда число линейно независимых собственных векторов преобразования равно размерности пространства, эта нормальная форма совпадает с диагональной. Мы сформулируем сейчас окончательный результат, который докажем в § 19.

Пусть задано произвольное линейное преобразование A в комплексном пространстве n измерений. Предположим, что у A имеется k $(k \leqslant n)$ линейно независимых собственных векторов

$$e_1, f_1, \ldots, h_1,$$

соответствующих собственным значениям $\lambda_1, \lambda_2, \ldots \ldots, \lambda_k$. Тогда существует базис, состоящий из k групп векторов *):

$$e_1, \ldots, e_p;\ f_1, \ldots, f_q;\ \ldots;\ h_1, \ldots, h_s, \tag{1}$$

в котором преобразование A имеет следующий вид:

$$\begin{aligned}
&Ae_1 = \lambda_1 e_1, \quad Ae_2 = e_1 + \lambda_1 e_2, \quad \ldots, \quad Ae_p = e_{p-1} + \lambda_1 e_p; \\
&Af_1 = \lambda_2 f_1, \quad Af_2 = f_1 + \lambda_2 f_2, \quad \ldots, \quad Af_q = f_{q-1} + \lambda_2 f_q; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&Ah_1 = \lambda_k h_1, \quad Ah_2 = h_1 + \lambda_k h_2, \quad \ldots, \quad Ah_s = h_{s-1} + \lambda_k h_s.
\end{aligned} \tag{2}$$

Мы видим, что базисные векторы каждой группы переходят при нашем преобразовании в линейную комбинацию векторов той же группы. Отсюда следует, что каждая группа базисных векторов порождает подпространство, инвариантное относительно

*) Ясно, что $p+q+\ldots+s = n$. Если же $k = n$, то каждая группа состоит из одного вектора, а именно собственного вектора.

преобразования A. Рассмотрим несколько подробнее преобразование, задаваемое формулами (2).

В подпространстве, порожденном каждой группой, есть собственный вектор; например, в подпространстве, порожденном векторами $e_1, e_2, \ldots, e_p$, таким собственным вектором является e_1.

Вектор e_2 называют иногда присоединенным собственным вектором первого порядка. Это значит, что Ae_2 пропорционально e_2 с точностью до собственного вектора, как это видно из равенства

$$Ae_2 = \lambda_1 e_2 + e_1.$$

Аналогично $e_3, e_4, \ldots$ называют присоединенными векторами второго, третьего и т. д. порядков.

Каждый из них является «как бы собственным», т. е. собственным с точностью до присоединенного вектора низшего порядка

$$Ae_k = \lambda_1 e_k + e_{k-1}.$$

Таким образом, базис каждого инвариантного подпространства состоит из одного собственного вектора и такого же количества присоединенных, которое нужно добавить, чтобы получить базис данного подпространства.

Покажем, что в каждом из этих подпространств имеется, с точностью до множителя, *лишь один* собственный вектор.

Действительно, рассмотрим, например, подпространство, порожденное векторами $e_1, e_2, \ldots, e_p$. Допустим, что некоторый вектор из этого подпространства, т. е. некоторая линейная комбинация вида

$$c_1e_1 + c_2e_2 + \ldots + c_pe_p,$$

где не все c_k равны нулю, является собственным вектором, т. е.

$$A(c_1e_1 + c_2e_2 + \ldots + c_pe_p) = \lambda(c_1e_1 + c_2e_2 + \ldots + c_pe_p).$$

Подставляя вместо левой части ее выражение по формулам (2), получаем равенство

$$c_1\lambda_1 e_1 + c_2(e_1 + \lambda_1 e_2) + \ldots + c_p(e_{p-1} + \lambda_1 e_p) = \\ = \lambda c_1 e_1 + \lambda c_2 e_2 + \ldots + \lambda c_p e_p.$$

Отсюда, приравнивая коэффициенты при каждом из базисных векторов, имеем систему уравнений для нахождения величин $\lambda, c_1, c_2, \ldots, c_p$:

$$\begin{aligned} c_1\lambda_1 + c_2 &= \lambda c_1, \\ c_2\lambda_1 + c_3 &= \lambda c_2, \\ &\ldots\ldots\ldots\ldots \\ c_{p-1}\lambda_1 + c_p &= \lambda c_{p-1}, \\ c_p\lambda_1 \qquad &= \lambda c_p. \end{aligned}$$

Покажем прежде всего, что $\lambda = \lambda_1$. Действительно, если бы $\lambda \neq \lambda_1$, то из последнего равенства мы имели бы $c_p = 0$ и затем из остальных равенств $c_{p-1} = c_{p-2} = c_2 = c_1 = 0$. Итак, $\lambda = \lambda_1$; тогда из первого уравнения имеем $c_2 = 0$, из второго $c_3 = 0$ и т. д. до $c_p = 0$. Значит, собственный вектор равен $c_1 e_1$ и, следовательно, с точностью до множителя совпадает с первым вектором соответствующей группы.

Выпишем матрицу преобразования (2). Так как векторы каждой группы преобразуются в линейные комбинации векторов той же группы, то в первых p столбцах матрицы преобразования могут быть отличны от нуля лишь элементы первых p строк, в следующих q столбцах могут быть отличны от нуля лишь элементы, стоящие в строках с теми же номерами, что и у этих столбцов, и т. д. Таким образом, в данном базисе матрица преобразования будет состоять из k клеток, расположенных по главной диагонали, а все элементы, не принадлежащие ни одной из этих клеток, будут равны нулю.

Для того чтобы понять, что стоит в каждой клетке матрицы преобразования A, достаточно еще раз написать, как преобразуются векторы одной группы. Мы имеем:

$$\begin{aligned}
Ae_1 &= \lambda_1 e_1, \\
Ae_2 &= \quad e_1 + \lambda_1 e_2, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
Ae_{p-1} &= \qquad\qquad\qquad e_{p-2} + \lambda_1 e_{p-1}, \\
Ae_p &= \qquad\qquad\qquad\qquad\qquad e_{p-1} + \lambda_1 e_p.
\end{aligned}$$

Вспоминая, как строится матрица, отвечающая данному преобразованию базиса, получаем, что клетка матрицы, соответствующая данной группе векторов, имеет вид

$$\boldsymbol{A}_1 = \begin{pmatrix} \lambda_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_1 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_1 \end{pmatrix}. \tag{3}$$

Вся же матрица оказывается составленной из таких клеток порядков $p, q, \dots, s$ соответственно, т. е. имеет вид

$$\begin{pmatrix}
\begin{matrix} \lambda_1 & 1 & 0 & \dots & 0 \\ 0 & \lambda_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_1 \end{matrix} & & & \\
& \begin{matrix} \lambda_2 & 1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_2 \end{matrix} & & \\
& & \ddots & \\
& & & \begin{matrix} \lambda_k & 1 & 0 & \dots & 0 \\ 0 & \lambda_k & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_k \end{matrix}
\end{pmatrix}, \tag{4}$$

где все элементы вне клеток — нули.

Заметим также, что не все λ_i обязаны быть различными.

Упражнение. Найти все инвариантные подпространства преобразования с матрицей (3).

Хотя приведенная здесь нормальная форма выглядит сложнее, чем, например, диагональная матрица, однако и с ней можно достаточно просто производить алгебраические операции. Мы покажем, например, как вычислить многочлен от матрицы (4). Матрица (4) имеет вид

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k \end{pmatrix},$$

где $\boldsymbol{A}_i$ — отдельные клетки, а все невыписанные элементы — нули. Тогда

$$\boldsymbol{A}^2 = \begin{pmatrix} \boldsymbol{A}_1^2 & & & \\ & \boldsymbol{A}_2^2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k^2 \end{pmatrix}, \ \ldots, \ \boldsymbol{A}^m = \begin{pmatrix} \boldsymbol{A}_1^m & & & \\ & \boldsymbol{A}_2^m & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k^m \end{pmatrix},$$

т. е. для того, чтобы возвести в некоторую степень матрицу $\boldsymbol{A}$, достаточно уметь возвести в эту степень каждую из клеток. Пусть теперь $P(t) = a_0 + a_1 t + \ldots + a_m t^m$ — произвольный многочлен. Тогда легко видеть, что

$$P(\boldsymbol{A}) = \begin{pmatrix} P(\boldsymbol{A}_1) & & & \\ & P(\boldsymbol{A}_2) & & \\ & & \ddots & \\ & & & P(\boldsymbol{A}_k) \end{pmatrix}.$$

Покажем теперь, как вычислить $P(\boldsymbol{A}_1)$, т. е. многочлен от одной клетки нормальной формы матрицы (3). Для этого запишем матрицу (3) в виде

$$\boldsymbol{A}_1 = \lambda_1 \boldsymbol{E} + \boldsymbol{I},$$

где $\boldsymbol{E}$ — единичная матрица порядка p, а матрица $\boldsymbol{I}$ имеет вид

$$\boldsymbol{I} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

Заметим, что матрицы $\boldsymbol{I}^2, \boldsymbol{I}^3, \ldots, \boldsymbol{I}^{p-1}$ имеют следующий вид *):

$$\boldsymbol{I}^2 = \begin{pmatrix} 0 & 0 & 1 & . & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & . & \ldots & 0 \\ 0 & 0 & 0 & . & \ldots & 0 \end{pmatrix}, \ \ldots, \ \boldsymbol{I}^{p-1} = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix},$$

а

$$\boldsymbol{I}^p = \boldsymbol{I}^{p+1} = \ldots = 0.$$

Теперь нетрудно вычислить произвольный многочлен от матрицы (3). Действительно, многочлен $P(t)$ можно по формуле Тейлора представить в виде

$$P(t) = P(\lambda_1) + (t-\lambda_1)P'(\lambda_1) + \frac{(t-\lambda_1)^2}{2!}P''(\lambda_1) + \ldots \\ \ldots + \frac{(t-\lambda_1)^n}{n!}P^{(n)}(\lambda_1),$$

где n — степень многочлена. Подставляя вместо t матрицу $\boldsymbol{A}_1$, имеем:

$$P(\boldsymbol{A}_1) = P(\lambda_1)\boldsymbol{E} + (\boldsymbol{A}_1 - \lambda_1\boldsymbol{E})P'(\lambda_1) + \frac{(\boldsymbol{A}_1 - \lambda_1\boldsymbol{E})^2}{2!}P''(\lambda_1) + \ldots \\ \ldots + \frac{(\boldsymbol{A}_1 - \lambda_1\boldsymbol{E})^n}{n!}P^{(n)}(\lambda_1).$$

Но $\boldsymbol{A}_1 - \lambda_1\boldsymbol{E} = \boldsymbol{I}$, следовательно,

$$P(\boldsymbol{A}_1) = P(\lambda_1)\boldsymbol{E} + P'(\lambda_1)\boldsymbol{I} + \frac{P''(\lambda_1)}{2!}\boldsymbol{I}^2 + \ldots + \frac{P^{(n)}(\lambda_1)}{n!}\boldsymbol{I}^n.$$

Подставляя вместо $\boldsymbol{I}, \boldsymbol{I}^2, \ldots, \boldsymbol{I}^{p-1}$ их выражения и учитывая, что $\boldsymbol{I}^p = \boldsymbol{I}^{p+1} = \ldots = 0$, получаем окончательный вид матрицы $P(\boldsymbol{A}_1)$:

$$P(\boldsymbol{A}_1) = \begin{pmatrix} P(\lambda_1) & \dfrac{P'(\lambda_1)}{1!} & \dfrac{P''(\lambda_1)}{2!} & \ldots & \dfrac{P^{(p-1)}(\lambda_1)}{(p-1)!} \\ 0 & P(\lambda_1) & \dfrac{P'(\lambda_1)}{1!} & \ldots & \dfrac{P^{(p-2)}(\lambda_1)}{(p-2)!} \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & P(\lambda_1) \end{pmatrix}.$$

*) Проще всего это сосчитать так. Мы имеем $Ie_1 = 0$, $Ie_2 = e_1, \ \ldots, \ Ie_p = e_{p-1}$. Следовательно, $I^2e_1 = 0$, $I^2e_2 = 0$, $I^2e_3 = e_1, \ \ldots, \ I^2e_p = e_{p-2}$. Аналогично, $I^3e_1 = I^3e_2 = I^3e_3 = 0$, $I^3e_4 = e_1, \ \ldots, \ I^3e_p = e_{p-3}$.

Мы видим, таким образом, что, для того чтобы вычислить многочлен от одной клетки нормальной формы порядка p, достаточно знать значение этого многочлена и его производных до порядка $p-1$ в точке λ_1, где λ_1 — собственное значение, отвечающее клетке. Отсюда следует, что если матрица $\boldsymbol{A}$ имеет нормальную форму (4) с клетками порядков $p, q, \ldots, s$, то для вычисления матрицы $P(\boldsymbol{A}_1)$ достаточно знать значения $P(t)$ в точках $t=\lambda_1, \lambda_2, \ldots, \lambda_k$ с производными до порядков $p-1, q-1, \ldots, s-1$ соответственно.

Мы докажем следующую теорему.

Теорема. *Пусть в комплексном n-мерном пространстве задано линейное преобразование A. Тогда можно найти базис, в котором матрица линейного преобразования имеет нормальную форму. Другими словами, можно найти базис, в котором линейное преобразование имеет вид* (2).

Два независимых доказательства сформулированной теоремы будут даны в § 19 и 20. Кроме того, важная теория инвариантных множителей и λ-матриц дает нам третье независимое доказательство этого результата.

§ 19. Приведение произвольного преобразования к нормальной форме

Мы уже упоминали в § 18, что в случае, когда у преобразования A не хватает линейно независимых собственных векторов (т.е. когда их число меньше размерности пространства), базис приходится дополнять за счет так называемых присоединенных векторов (их точное определение будет дано чуть позже). В этом параграфе дается способ построения базиса, в котором матрица преобразования A имеет жорданову нормальную форму. Этот базис мы непосредственно наберем из собственных и присоединенных векторов, и такой способ выбора является, в некотором смысле, наиболее естественным *).

Перед этим параграфом мы рекомендуем читателю перечитать п. 4 § 9 и разобрать приведенные там примеры.

1. Собственные и присоединенные векторы линейного преобразования. Пусть λ_0 — некоторое собственное значение преобразования A. Мы уже имели раньше такое определение.

Определение 1. *Вектор $x \neq 0$ называется собственным вектором преобразования A, отвечающим собственному значению λ_0, если*

$$Ax = \lambda_0 x, \quad \textit{т. е.} \quad (A - \lambda_0 E)x = 0. \tag{1}$$

Рассмотрим совокупность всех векторов, удовлетворяющих условию (1) при фиксированном λ_0. Ясно, что совокупность этих векторов является подпространством пространства R.

Мы обозначим его $N_{\lambda_0}^{(1)}$. Легко видеть, что $N_{\lambda_0}^{(1)}$ инвариантно относительно преобразования A (проверьте!).

Заметим, что подпространство $N_{\lambda_0}^{(1)}$ состоит из всех собственных векторов преобразования A, отвечающих собственному значению λ_0, к которым добавлен еще *нулевой вектор.*

Определение 2. *Вектор x называется присоединенным вектором 1-го порядка преобразования A, отвечающим собственному значению λ_0, если вектор*

$$y = (A - \lambda_0 E)x$$

является собственным вектором преобразования A.

Пусть λ_0 — собственное значение преобразования A.

*) См. также И. В. Проскуряков, Сборник задач по линейной алгебре, где имеется аналогичное доказательство.

Рассмотрим подпространство, состоящее из всех векторов x, для которых выполнено условие

$$(A - \lambda_0 E)^2 x = 0, \tag{2}$$

т. е. ядро преобразования $(A - \lambda_0 E)^2$. Обозначим это подпространство $N_{\lambda_0}^{(2)}$; $N_{\lambda_0}^{(2)}$ является инвариантным подпространством пространства R. В самом деле, пусть, $x \in N_{\lambda_0}^{(2)}$, т. е. $(A - \lambda_0 E)^2 x = 0$. Нам надо доказать, что и вектор $Ax \in N_{\lambda_0}^{(2)}$, т. е. что $(A - \lambda_0 E)^2 Ax = 0$. Но преобразование A перестановочно с $(A - \lambda_0 E)^2$, т. е.

$$(A - \lambda_0 E)^2 Ax = A(A - \lambda_0 E)^2 x = 0.$$

Рассмотрим несколько более подробно структуру пространства $N_{\lambda_0}^{(2)}$. В нем есть векторы двух типов.

Если $x \in N_{\lambda_0}^{(1)}$, т. е. $(A - \lambda_0 E)x = 0$, то подавно и $(A - \lambda_0 E)^2 x = 0$, т. е. $x \in N_{\lambda_0}^{(2)}$. Таким образом, $N_{\lambda_0}^{(1)}$ целиком содержится в $N_{\lambda_0}^{(2)}$. Если $x \in N_{\lambda_0}^{(2)}$, но $x \notin N_{\lambda_0}^{(1)}$, т. е.

$$\begin{gathered} (A - \lambda_0 E)x \neq 0, \\ (A - \lambda_0 E)^2 x = 0, \end{gathered}$$

то x — присоединенный вектор 1-го порядка. Действительно, в этом случае $y = (A - \lambda_0 E)x$ есть собственный вектор.

Таким образом, подпространство $N_{\lambda_0}^{(2)}$ получается, если к подпространству $N_{\lambda_0}^{(1)}$ добавить присоединенные векторы 1-го порядка.

Аналогично вводим подпространство $N_{\lambda_0}^{(k)}$, состоящее из всех векторов x, для которых

$$(A - \lambda_0 E)^k x = 0. \tag{3}$$

Это подпространство инвариантно относительно преобразования A. Ясно, что подпространство $N_{\lambda_0}^{(k)}$ содержит предыдущее подпространство $N_{\lambda_0}^{(k-1)}$.

О п р е д е л е н и е 3. *Вектор x называется присоединенным вектором k-го порядка, если вектор*

$$y = (A - \lambda_0 E)x$$

есть присоединенный вектор порядка $k-1$.

По индукции можно показать, что если x — присоединенный вектор k-го порядка, то

$$(A - \lambda_0 E)^k x \neq 0,$$
$$(A - \lambda_0 E)^{k+1} x = 0.$$

Другими словами, присоединенным вектором k-го порядка называется вектор, принадлежащий $N_{\lambda_0}^{(k+1)}$ и не принадлежащий $N_{\lambda_0}^{(k)}$.

П р и м е р. Пусть R — пространство многочленов степени $\leqslant n-1$ и преобразование A — дифференцирование:

$$AP(t) = \frac{d}{dt}P(t).$$

Легко видеть, что $\lambda = 0$ есть собственное значение. Соответствующий ему собственный вектор $P(t) = \text{const}$. Найдем для этого преобразования пространства $N_0^{(k)}$. По определению $N_0^{(k)}$состоит из всех многочленов $P(t)$, для которых $A^k P(t) = 0$, т. е.

$$\frac{d^k}{dt^k}P(t) = 0.$$

Это будут все многочлены, степень которых не превышает $k-1$. Присоединенными векторами k-го порядка будут многочлены, степень которых в точности равна $k-1$.

В этом примере размерность каждого из подпространств равна $N_0^{(k)}$ и она растет от 1 до n вместе с ростом k. Подпространство $N_0^{(n)}$ уже совпадает со всем пространством R, и если мы захотим определить $N_0^{(n+1)}, N_0^{(n+2)}$ и т. д., то все эти подпространства будут совпадать с $N_0^{(n)}$.

Легко видеть также, что в этом примере $AN_0^{(k+1)} = N_0^{(k)}$. Это следует из того, что каждый многочлен степени k есть производная от многочлена степени $k+1$.

Упражнение. Показать, что для любого линейного преобразования A имеет место включение

$$(A - \lambda_0 E)N_{\lambda_0}^{(k+1)} \subset N_{\lambda_0}^{(k)}.$$

Пусть A — линейное преобразование, а λ_0 — его собственное значение. Покажем, что подпространства $N_{\lambda_0}^{(1)}, N_{\lambda_0}^{(2)}, \ldots$ сначала строго возрастают с ростом индекса, а затем, начиная с некоторого номера $p \leqslant n$, этот рост прекращается, т. е.

$$N_{\lambda_0}^{(p)} = N_{\lambda_0}^{(p+1)} = \ldots$$

(см. приведенный в этом пункте пример).

Мы уже показали, что каждое подпространство $N_{\lambda_0}^{(k)}$ содержит $N_{\lambda_0}^{(k-1)}$, т. е. что с увеличением номера подпространства $N_{\lambda_0}^{(k)}$, а значит, и их размерности, могут только увеличиваться.

Так как наше пространство конечномерно, то для какого-то $p \leqslant n$ мы впервые получим, что $N_{\lambda_0}^{(p)} = N_{\lambda_0}^{(p+1)}$ (см. упражнение на стр. 25).

Докажем, что в этом случае $N_{\lambda_0}^{(p+1)} = N_{\lambda_0}^{(p+2)} = \ldots$, т. е. что дальнейшего возрастания подпространств происходить не будет.

Действительно, предположим противное, а именно, что $N_{\lambda_0}^{(p+1)} = N_{\lambda_0}^{(p)}$, но для некоторого $i > 0$ подпространство $N_{\lambda_0}^{(p+i+1)}$ строго больше, чем $N_{\lambda_0}^{(p+i)}$. Тогда существует вектор x такой, что

$$x \in N_{\lambda_0}^{(p+i+1)}, \; x \notin N_{\lambda_0}^{(p+i)}.$$

Это значит, что

$$(A - \lambda_0 E)^{p+i+1} x = 0, \quad \text{но} \quad (A - \lambda_0 E)^{p+1} x \neq 0. \tag{4}$$

Обозначим через y вектор $y = (A - \lambda_0 E)^i x$. Тогда первое из равенств (4) означает, что $y \in N_{\lambda_0}^{(p+1)}$, а второе, что $y \notin N_{\lambda_0}^{(p)}$, что невозможно, так как подпространства $N_{\lambda_0}^{(p+1)}$ и $N_{\lambda_0}^{(p)}$ по предположению совпадают.

Итак, пусть λ_0 — некоторое собственное значение преобразования A. Основным результатом этого пункта является построение инвариантного подпространства $N_{\lambda_0}^{(p)}$, состоящего из всех собственных и присоединенных векторов, отвечающих этому собственному значению.

Кроме того, в п. 3 нам понадобится более детальная структура $N_{\lambda_0}^{(p)}$. А именно, обозначая через $N_{\lambda_0}^{(k)}$ подпространство, состоящее из присоединенных векторов порядка $\leqslant k-1$, мы получили возрастающую цепочку инвариантных подпространств

$$0 \subset N_{\lambda_0}^{(1)} \subset N_{\lambda_0}^{(2)} \subset \ldots \subset N_{\lambda_0}^{(p)}. \tag{5}$$

Все члены этой цепочки различны. Подпространство $N_{\lambda_0}^{(k)}$ состоит при этом из всех векторов x, для которых

$$(A - \lambda_0 E)^k x = 0,$$

т. е. это есть ядро преобразования $(A - \lambda_0 E)^k$.

Преобразование $A - \lambda_0 E$ переводит каждое из подпространств цепочки (5) в предшествующее.

2. Выделение подпространства, в котором преобразование A имеет только одно собственное значение. Пусть λ_1 — некоторое собственное значение преобразования A. В этом пункте мы покажем, что пространство R можно разложить в прямую сумму двух инвариантных подпространств, в первом из которых преобразование A имеет лишь одно собственное значение λ_1, а во втором у преобразования A уже нет собственного значения λ_1.

Не ограничивая общности, можно считать, что $\lambda_1 = 0$.

Действительно, пусть $\lambda_1 \neq 0$. Рассмотрим преобразование $B = A - \lambda_1 E$; оно уже имеет собственное значение, равное нулю *). Очевидно также, что инвариантные подпространства преобразований A и B совпадают.

Итак, впредь мы будем считать, что преобразование имеет собственное значение $\lambda = 0$. Докажем наше утверждение сначала для частного случая, когда *в пространстве нет присоединенных векторов, отвечающих этому собственному значению, а есть только собственные векторы* **).

Нам нужно построить два инвариантных подпространства, прямая сумма которых равна R. В качестве первого из них, в котором $\lambda = 0$ есть единственное собственное значение, можно взять совокупность N_0 всех

---

*) В самом деле, если λ_1 — собственное значение преобразования A, т. е. $Af = \lambda_1 f$ $(f \neq 0)$, то $Bf = (A - \lambda_1 E)f = 0$, т. е. f — собственный вектор B, отвечающий собственному значению $\lambda = 0$.

**) Хотя потом наше утверждение будет независимо доказано для общего случая, рассмотрение этого частного случая полезно, так как, во-первых, на нем более выпукло видна основная идея доказательства и, во-вторых, становится очевидной необходимость введения подпространств, отличных от введенных здесь N_0 и M.

собственных векторов, отвечающих собственному значению $\lambda = 0$ или, другими словами, ядро преобразования A.

В качестве второго подпространства возьмем образ M пространства R при преобразовании A, т. е. совокупность векторов $y = Ax$, где x пробегает все пространство R. Легко видеть, что каждое из этих подпространств инвариантно (это доказано в п. 4 § 9).

Докажем, что они дают разложение пространства в прямую сумму. Так как сумма размерностей ядра и образа для любого преобразования A равна n (см. п. 4 § 9), то достаточно доказать, что пересечение этих подпространств равно нулю.

Предположим, что это не так, т. е. пусть существует вектор $y \neq 0$ такой, что $y \in M$ и $y \in N_0$. Так как $y \in M$, то он имеет вид

$$y = Ax, \tag{6}$$

где x — некоторый вектор из R. Так как $y \in N_0$, то

$$Ay = 0, \quad \text{где} \quad y \neq 0. \tag{7}$$

Равенство (7) означает, что y есть собственный вектор преобразования A, отвечающий собственному значению $\lambda = 0$, а равенство (6) при этом означает, что x есть присоединенный вектор первого порядка, отвечающий тому же собственному значению. Мы же предположили, что у преобразования A нет присоединенных векторов, отвечающих собственному значению $\lambda = 0$.

Таким образом, доказано, что подпространства M и N_0 не имеют общих векторов, кроме нулевого.

Вспоминая, что сумма размерностей образа и ядра равна n, мы получаем отсюда, что пространство R разложимо в прямую сумму инвариантных подпространств M и N_0:

$$R = M \oplus N_0.$$

З а м е ч а н и е. Из приведенного выше доказательства видно, что образ и ядро имеют пересечение, отличное от нуля в том и только том случае, когда преобразование A имеет присоединенные векторы, отвечающие собственному значению $\lambda = 0$.

Разобранный частный случай дает нам идею того, как проводить доказательство в общем случае, когда A имеет также и присоединенные векторы, отвечающие собственному значению $\lambda = 0$. Подпространство N_0 при этом оказывается слишком узким, и его естественно расширить за счет добавления всех присоединенных векторов, отвечающих собственному значению $\lambda = 0$. Второе же подпространство M оказывается при этом слишком большим *).

Итак, рассмотрим введенное в п. 1 инвариантное подпространство $N_0^{(p)}$, состоящее из всех собственных и присоединенных векторов преобразования A, отвечающих собственному значению $\lambda = 0$. Как мы помним, оно является ядром преобразования A^p, т. е. состоит из всех векторов x, для которых

$$A^p x = 0.$$

В качестве второго слагаемого прямой суммы мы возьмем подпространство $M^{(p)}$ — образ пространства R при том же преобразовании A^p.

Легко видеть, что $M^{(p)}$ также инвариантно относительно преобразования A. Действительно, если $y \in M^{(p)}$, т. е. $y = A^p x$, то

$$Ay = A^{p+1}x = A^p(Ax),$$

т. е. Ay также принадлежит $M^{(p)}$.

*) Что M «слишком велико», видно при этом не только из соображений размерности, но также и из того, что M пересекается даже с самим N_0, а не только с его расширением.

Т е о р е м а 1. *Пространство R можно разложить в прямую сумму инвариантных подпространств $N_0^{(p)}$ и $M^{(p)}$. При этом подпространство $N_0^{(p)}$ состоит только из собственных и присоединенных векторов, отвечающих собственному значению $\lambda = 0$, а в подпространстве $M^{(p)}$ преобразование A обратимо (т. е. $\lambda = 0$ не является собственным значением преобразования A в подпространстве $M^{(p)}$).*

Для доказательства первого утверждения нам, как и в рассмотренном выше частном случае, достаточно показать, что пересечение подпространств $N_0^{(p)}$ и $M^{(p)}$ равно нулю. Допустим противное, т. е. пусть существует вектор $y \neq 0$ такой, что $y \in M^{(p)}$ и $y \in N_0^{(p)}$. Так как $y \in M^{(p)}$, то

$$y = A^p x. \tag{8}$$

Далее, так как $y \in N_0^{(p)}$, то

$$A^p y = 0. \tag{9}$$

Но из равенств (8) и (9) следует, что существует такой вектор x, для которого

$$A^p x \neq 0$$

и в то же время

$$A^{2p} x = A^p y = 0.$$

Это значит, что x есть присоединенный вектор преобразования A с собственным значением $\lambda = 0$, не принадлежащий подпространству $N_0^{(p)}$, что невозможно, так как $N_0^{(p)}$ состоит из всех таких векторов.

Таким образом, мы доказали, что пересечение $N_0^{(p)}$ и $M^{(p)}$ равно нулю. Так как сумма размерностей этих подпространств равна n (это ядро и образ преобразования A^p), то отсюда следует, что пространство R

раскладывается в прямую сумму этих подпространств:

$$R = M^{(p)} \oplus N_0^{(p)}. \tag{10}$$

Докажем теперь второе утверждение теоремы, т. е. что в подпространстве $M^{(p)}$ преобразование A не имеет нулевого собственного значения. Действительно, если бы это было не так, то в $M^{(p)}$ существовал бы вектор $x \neq 0$ такой, что

$$A^p x = 0.$$

Но это равенство означает, что $x \in N_0^{(p)}$, т. е. является общим вектором $M^{(p)}$ и $N_0^{(p)}$, а мы доказали, что таким вектором может быть только нуль.

Теорема доказана полностью.

Теперь мы можем освободиться от предположения, что выделенное подпространство отвечает нулевому собственному значению, и считать установленным следующий факт.

Если λ_1 — *некоторое собственное значение преобразования* A, *то пространство* R *можно разложить в прямую сумму инвариантных подпространств* R_1 *и* $\tilde{R}$, *в первом из которых преобразование* A *имеет только собственное значение* λ_1, *а во втором все собственные значения* A *отличны от* λ_1.

Применяя полученный результат к преобразованию A в пространстве $\tilde{R}$ и к некоторому собственному значению λ_2 этого преобразования, мы «отщепим» инвариантное подпространство, отвечающее собственному значению λ_2. Продолжая этот процесс, пока не будут исчерпаны все собственные значения преобразования A, мы получим доказательство следующей теоремы:

Теорема 2. *Пусть преобразование* A *пространства* R *имеет* k *различных собственных значений* $\lambda_1, \ldots, \lambda_k$. *Тогда* R *можно разложить в прямую сумму*

k инвариантных подпространств $N_{\lambda_1}^{(p_1)}, \ldots, N_{\lambda_k}^{(p_k)}$:

$$R = N_{\lambda_1}^{(p_1)} \oplus \ldots \oplus N_{\lambda_k}^{(p_k)}. \tag{11}$$

Каждое из подпространств $N_{\lambda_i}^{(p_i)}$ состоит только из собственных и присоединенных векторов, отвечающих собственному значению λ_i.

Другими словами, для каждого i существует такое число p_i, что для всех $x \in N_{\lambda_i}^{(p_i)}$

$$(A - \lambda_i E)^{p_i} x = 0.$$

У нас осталась еще только одна, впрочем, не менее важная задача — выбрать в каждом из этих подпространств базис, в котором матрица преобразования имеет жорданову нормальную форму. Это будем сделано в следующем пункте.

3. Приведение к нормальной форме матрицы с одним собственным значением. В случае, если пространство состоит только из собственных векторов, базис в пространстве можно выбирать произвольно и матрица преобразования в этом базисе имеет диагональный вид.

В общем случае неосторожный выбор базиса может запутать картину.

Чтобы выбрать базис, в котором матрица преобразования имеет наиболее простой вид, мы будем тянуть цепочки собственных и присоединенных векторов, выбрав некоторый базис в подпространстве $N^{(p)}$ и последовательно применяя к векторам этого базиса преобразование A.

Введем предварительно некоторые понятия, удобные для дальнейшего.

Определение 4. *Векторы из пространства R называются относительно линейно независимыми над подпространством R_1, если никакая их линейная комбинация, отличная от нуля, не принадлежит R_1.*

Заметим, что всякие линейно зависимые векторы из R относительно линейно зависимы над любым подпространством.

О п р е д е л е н и е 5. *Базисом пространства R относительно подпространства R_1 называется такая система $e_1, \ldots, e_k$ линейно независимых векторов из R, которая после пополнения каким-нибудь базисом из R_1 образует базис во всем пространстве.*

Такой базис легко построить. Для этого достаточно выбрать какой-нибудь базис в R_1, дополнить его до базиса во всем пространстве и затем отбросить векторы исходного базиса из R_1. Число векторов в таком *относительном базисе* равно разности размерностей пространства и подпространства.

Всякую систему относительно линейно независимых векторов над R_1 можно дополнить до относительного базиса. Для этого нужно к выбранным векторам добавить какой-нибудь базис подпространства R_1. Получится некоторая система векторов из R, которые, как легко проверить, линейно независимы. Чтобы получить относительный базис, нужно дополнить эту систему до базиса во всем пространстве R, а затем отбросить базис подпространства.

Итак, пусть преобразование A в пространстве R имеет только одно собственное значение. Не ограничивая общности, можно предположить, что оно равно нулю.

Рассмотрим снова цепочку (5) подпространств, полученных в п. 1:

$$0 \subset N_0^{(1)} \subset \ldots \subset N_0^{(p)} = N_0^{(p+1)} = \ldots,$$

где подпространство $N_0^{(k)}$ есть ядро преобразования A^k. Так как преобразование A в пространстве R не имеет отличных от нуля собственных значений, то, очевидно, $N^{(p)}$ совпадает при этом со всем пространством R.

Выберем в максимальном из этих подпространств $N_0^{(p)}$ базис относительно содержащегося в нем подпространства $N_0^{(p-1)}$. Пусть векторы этого базиса будут

$$e_1, \ldots, e_q.$$

Очевидно, что это будут присоединенные векторы $(p-1)$-го порядка. Мы уже видели (см. упражнение на стр. 211), что $AN_0^{(p)} \subset N_0^{(p-1)}$. Поэтому векторы

$$Ae_1, \ldots, Ae_q$$

лежат в $N_0^{(p-1)}$. Покажем, что эти векторы линейно независимы в $N_0^{(p-1)}$ относительно лежащего в нем подпространства $N_0^{(p-2)}$. Действительно, пусть не все $\alpha_i = 0$ и

$$\alpha_1 Ae_1 + \ldots + \alpha_q Ae_q = A(\alpha_1 e_1 + \ldots + \alpha_q e_q) \in N_0^{(p-2)}.$$

Тогда вектор $x = \alpha_1 e_1 + \ldots + \alpha_q e_q \in N_0^{(p-1)}$, а это противоречит предположению, что векторы $e_1, \ldots, e_q$ линейно независимы над $N_0^{(p-1)}$.

Дополним векторы $Ae_1, \ldots, Ae_q$ до базиса в $N_0^{(p-1)}$ относительно $N_0^{(p-2)}$. Мы получим тогда $q+s$ векторов

$$Ae_1, \ldots, Ae_q, f_1, \ldots, f_s,$$

которые представляют собой максимальное число линейно независимых присоединенных векторов порядка $p-2$.

Снова применим к этим векторам преобразование A и полученную систему векторов из $N_0^{(p-2)}$ дополним, как и выше, до базиса в $N_0^{(p-2)}$ относительно $N_0^{(p-3)}$.

Продолжая этот процесс, мы дойдем до подпространства $N_0^{(1)}$ и выберем базис в этом пространстве, состоящий из максимального числа линейно независимых собственных векторов.

Расположим полученные векторы в следующую таблицу

$$
\begin{array}{cccccccc}
e_1 \;\ldots & e_q & & & & & & \\
Ae_1 \;\ldots & Ae_q & f_1 \;\ldots & f_s & & & & \\
A^2e_1 \;\ldots & A^2e_q & Af_1 \;\ldots & Af_s & & & & \\
\multicolumn{8}{c}{\dotfill} \\
\multicolumn{8}{c}{\dotfill} \\
A^{p-1}e_1 \;\ldots & A^{p-1}e_q & A^{p-2}f_1 \;\ldots & A^{p-2}f_s & \ldots & h_1 & \ldots & h_r
\end{array}
\qquad (12)
$$

Векторы нижней строчки образуют базис в подпространстве $N_0^{(1)}$. Векторы двух нижних строчек образуют базис в $N_0^{(2)}$, так как это есть базис $N_0^{(2)}$ относительно $N_0^{(1)}$в соединении с базисом $N_0^{(1)}$. Векторы трех нижних строчек образуют базис в $N_0^{(3)}$ и т. д. Наконец все векторы таблицы образуют базис в $N_0^{(p)}$, т. е. во всем пространстве R.

Покажем, что в этом базисе матрица преобразования A имеет жорданову нормальную форму. Действительно, рассмотрим произвольный столбец таблицы (12), например, для определенности первый.

Обозначим для удобства $A^{p-1}e_1$ через $\overline{e}_1$, $A^{p-2}e_1$ — через $\overline{e}_2$ и т. д. и рассмотрим действие преобразования A на каждый из этих векторов. Так как $\overline{e}_1$ — собственный вектор, отвечающий нулевому собственному значению, то

$$A\overline{e}_1 = 0.$$

Дальше, по определению,

$$A\overline{e}_2 = AA^{p-2}e_1 = A^{p-1}e_1 = \overline{e}_1$$

и аналогично

$$\begin{aligned} &A\overline{e}_3 = \overline{e}_2, \\ &\ldots\ldots\ldots \\ &A\overline{e}_p = \overline{e}_{p-1}. \end{aligned}$$

Таким образом, преобразование A переводит векторы первого столбца снова в себя, т. е. подпространство R_1, натянутое на эти векторы, инвариантно относительно A. Матрица преобразования A в подпространстве R_1 в базисе $\overline{e}_1, \ldots, \overline{e}_p$ имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}, \tag{13}$$

т. е. это есть жорданова клетка, отвечающая собственному значению $\lambda = 0$. Аналогичное инвариантное подпространство отвечает каждому из столбцов таблицы (12), и размерность каждого такого подпространства равна числу векторов в соответствующем столбце. Так как матрица преобразования A в базисе, состоящем из векторов какого-либо столбца таблицы (12), имеет вид (13), то матрица преобразования во всем пространстве R в базисе, состоящем из всех векторов таблицы (12), состоит из жордановых клеток, число которых равно числу столбцов в этой таблице, а размер каждой клетки равен числу векторов соответствующего столбца.

Если вместо преобразования A рассмотреть преобразование $A + \lambda_1 E$, то, так как матрица преобразования $\lambda_1 E$ диагональна, мы получим тот же результат для преобразования пространства R, имеющего только одно собственное значение, равное произвольному числу λ_1. Соответствующие жордановы клетки матрицы преобразования $A + \lambda_1 E$ будут иметь вид:

$$\begin{pmatrix} \lambda_1 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_1 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_1 \end{pmatrix}. \tag{14}$$

Вспоминая теперь, что для произвольного преобразования A мы можем разложить пространство R в сумму инвариантных подпространств, в каждом из которых преобразование A имеет только одно собственное значение (см. формулу (11)), мы получаем отсюда полное доказательство теоремы § 18.

§ 20. Другое доказательство теоремы о приведении к нормальной форме

Это доказательство мы будем вести по индукции, именно предположим, что для линейного преобразования в пространстве n измерений такой базис существует, и докажем, что мы можем найти нужный базис в пространстве $n + 1$ измерений. Для доказательства теоремы нам понадобится следующая

Лемма. *У всякого линейного преобразования A в n-мерном комплексном пространстве R существует хотя бы одно $(n-1)$-мерное инвариантное подпространство R'.*

Доказательство. Рассмотрим преобразование A^*; у него, как и у всякого преобразования, есть собственный вектор e

$$A^*e = \lambda e.$$

Покажем, что $(n-1)$-мерное подпространство R', состоящее из векторов x, ортогональных *) e, т. е. для которых $(x, e) = 0$, инвариантно относительно A. Действительно, пусть $x \in R'$, т. е. $(x, e) = 0$. Тогда

$$(Ax, e) = (x, A^*e) = (x, \lambda e) = 0$$

и, значит, Ax также принадлежит R'. Инвариантность подпространства R' относительно преобразования A доказана.

Докажем теперь сформулированную выше основную теорему этого параграфа.

Пусть A — произвольное линейное преобразование в $(n+1)$-мерном пространстве R. Согласно лемме, в R существует n-мерное подпространство R', инвариантное относительно A. Так как в n-мерном пространстве мы предполагаем теорему доказанной, то в R' существует базис, в котором линейное преобразование имеет нормальную форму. Обозначим этот базис в R' через

$$e_1, e_2, \ldots, e_p;\ f_1, f_2, \ldots, f_q;\ \ldots;\ h_1, h_2, \ldots, h_s,$$

где $p + q + \ldots + s = n$. В этом базисе линейное преобра-

---

*) Мы здесь используем наличие в R скалярного произведения, т. е. считаем пространство R евклидовым. Незначительным изменением доказательства можно освободиться от этого предположения и на протяжении всей главы считать R аффинным пространством.

зование в n-мерном подпространстве R' имеет вид

$$\begin{array}{llll}
\begin{aligned}
Ae_1 &= \lambda_1 e_1,\\
Ae_2 &= e_1 + \lambda_1 e_2,\\
&\ldots\ldots\ldots\\
Ae_p &= e_{p-1} + \lambda_1 e_p,
\end{aligned} & & & \\
& \begin{aligned}
Af_1 &= \lambda_2 f_1,\\
Af_2 &= f_1 + \lambda_2 f_2,\\
&\ldots\ldots\ldots\\
Af_q &= f_{q-1} + \lambda_2 f_q,
\end{aligned} & & \\
& & \ddots & \\
& & & \begin{aligned}
Ah_1 &= \lambda_k h_1,\\
Ah_2 &= h_1 + \lambda_k h_2,\\
&\ldots\ldots\ldots\\
Ah_s &= h_{s-1} + \lambda_k h_s.
\end{aligned}
\end{array}$$

Дополним этот базис каким-нибудь вектором e, который вместе с $e_1, e_2, \ldots, e_p;\ f_1, f_2, \ldots, f_q;\ \ldots;\ h_1, h_2, \ldots \ldots, h_s$ составляет базис в R.

Применим к e преобразование A и разложим полученный вектор Ae по векторам базиса:

$$Ae = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots + \beta_q f_q + \ldots \\ \ldots + \delta_1 h_1 + \ldots + \delta_s h_s + \tau e \ ^{*}).$$

Мы можем считать, что $\tau = 0$. В самом деле, если в некотором базисе A имеет нормальную форму, то $A - \tau E$

---

*) Линейное преобразование A имеет в $(n+1)$-мерном пространстве R собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_k$ и τ. Действительно, напишем матрицу преобразования A в базисе $e_1, e_2, \ldots \ldots, e_p; f_1, f_2, \ldots, f_q; \ldots; h_1, h_2, \ldots, h_s; e$. Она будет треугольной матрицей, в которой по диагонали стоят числа $\lambda_1, \lambda_2, \ldots, \lambda_k$ и τ.

Так как собственными значениями треугольной матрицы являются числа, стоящие по диагонали (см., например, § 10, п. 4), то $\lambda_1, \lambda_2, \ldots, \lambda_k$ и τ будут собственными значениями A в $(n+1)$-мерном пространстве. Таким образом, при переходе от инвариантного n-мерного к $(n+1)$-мерному пространству добавилось одно собственное значение: τ.

также имеет в этом базисе нормальную форму. Поэтому, если $\tau \neq 0$, то можно вместо преобразования A рассмотреть преобразование $A-\tau E$, причем A и $A-\tau E$, согласно сделанному замечанию, имеют нормальную форму в одном и том же базисе.

Мы полагаем, следовательно, что

$$Ae = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots \\ \ldots + \beta_q f_q + \ldots + \delta_1 h_1 + \ldots + \delta_s h_s. \quad (1)$$

Теперь постараемся заменить вектор e вектором e' так, чтобы после этой замены вектор Ae' стал возможно проще. Будем искать вектор e' в виде

$$e' = e - \varkappa_1 e_1 - \ldots - \varkappa_p e_p - \mu_1 f_1 - \ldots \\ \ldots - \mu_q f_q - \ldots - \omega_1 h_1 - \ldots - \omega_s h_s. \quad (2)$$

Мы имеем

$$Ae' = Ae - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) - \\ - A(\mu_1 f_1 + \ldots + \mu_q f_q) - \ldots - A(\omega_1 h_1 + \ldots + \omega_s h_s),$$

или, пользуясь формулой (1),

$$Ae' = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots + \beta_q f_q + \ldots \\ \ldots + \delta_1 h_1 + \ldots + \delta_s h_s - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) - \\ - A(\mu_1 f_1 + \ldots + \mu_q f_q) - \ldots - A(\omega_1 h_1 + \ldots + \omega_s h_s). \quad (3)$$

Коэффициенты $\varkappa_1, \ldots, \varkappa_p$; $\mu_1, \ldots, \mu_q$; $\ldots$; $\omega_1, \ldots, \omega_s$ мы можем выбрать произвольно. Подберем их так, чтобы в правой части формулы (3) осталось как можно меньше слагаемых.

Мы знаем, что каждой группе базисных векторов n-мерного подпространства R', в котором преобразование A имеет нормальную форму, отвечает свое собственное значение λ_1, λ_2 и т. д. Мы рассмотрим отдельно два случая, а именно разберем сначала случай, когда ни одно из этих собственных значений не равно нулю, а затем случай, когда это не так.

Разберем первый случай, когда $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\ldots$, $\lambda_k \neq 0$. В этом случае мы покажем, что вектор e' можно выбрать так, чтобы $Ae' = 0$, т. е. подобрать $\varkappa_1, \ldots, \omega_s$ так, чтобы все слагаемые в правой части (3) сократились. Так как векторы каждой группы переходят при преобразовании в комбинацию векторов той же группы, то векторы различных групп можно уничтожать независимо друг от друга. Покажем, как подобрать коэффициенты $\varkappa_1, \varkappa_2, \ldots, \varkappa_p$, чтобы в правой части (3) сократились векторы $e_1, e_2, \ldots, e_p$. Члены, содержащие эти векторы, имеют вид

$$\begin{aligned}
&\alpha_1 e_1 + \ldots + \alpha_p e_p - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) = \\
&\qquad = \alpha_1 e_1 + \ldots + \alpha_p e_p - \varkappa_1 \lambda_1 e_1 - \\
&\qquad\qquad - \varkappa_2(e_1 + \lambda_1 e_2) - \ldots - \varkappa_p(e_{p-1} + \lambda_1 e_p) = \\
&\qquad = (\alpha_1 - \varkappa_1\lambda_1 - \varkappa_2)e_1 + (\alpha_2 - \varkappa_2\lambda_1 - \varkappa_3)e_2 + \ldots \\
&\qquad\qquad \ldots + (\alpha_{p-1} - \varkappa_{p-1}\lambda_1 - \varkappa_p)e_{p-1} + (\alpha_p - \varkappa_p\lambda_1)e_p.
\end{aligned}$$

Приравнивая нулю коэффициент при e_p, определяем $\varkappa_p$, что возможно, так как $\lambda_1 \neq 0$, затем, приравнивая нулю коэффициент при e_{p-1}, определяем $\varkappa_{p-1}$ и так далее до $\varkappa_1$. Таким образом, мы уничтожили в (3) члены с $e_1, e_2, \ldots, e_p$. Аналогично вычисляем другие группы коэффициентов.

Мы получили, таким образом, вектор e', для которого

$$Ae' = 0.$$

Добавляя этот вектор к имеющемуся базису, получаем базис $e'; e_1, e_2, \ldots, e_p; f_1, f_2, \ldots, f_q; \ldots; h_1, h_2, \ldots, h_s$ в $(n+1)$-мерном пространстве, в котором преобразование имеет канонический вид. Вектор e' образует при этом отдельную группу с собственным значением, равным нулю (следовательно, с собственным значением τ, если бы мы не рассматривали вместо A преобразование $A - \tau E$).

Рассмотрим теперь второй случай, а именно пусть некоторым группам векторов базиса в n-мерном пространстве R' соответствуют собственные значения преобразования A, равные нулю. Тогда в правой части формулы (3) у нас будут слагаемые двух сортов — соответствующие группам с отличными от нуля собственными значениями и группам, для которых собственное значение равно нулю. С группами, у которых собственные значения отличны от нуля, мы можем поступить так же, как и в первом случае, т. е. подбором коэффициентов уничтожить векторы в правой части (3). Допустим, что после этой операции у нас останутся, например, три группы слагаемых $e_1, e_2, \ldots, e_p; f_1, f_2, \ldots, f_q;$ $g_1, g_2, \ldots, g_r$ с собственным значением, равным нулю, т. е. что $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Тогда

$$\begin{aligned} Ae' = \alpha_1 e_1 + \ldots + \alpha_p e_p + \beta_1 f_1 + \ldots + \beta_q f_q + \ldots + \gamma_1 g_1 + \ldots \\ \ldots + \gamma_r g_r - A(\varkappa_1 e_1 + \ldots + \varkappa_p e_p) - \\ - A(\mu_1 f_1 + \ldots + \mu_q f_q) - A(\nu_1 g_1 + \ldots + \nu_r g_r). \quad (4) \end{aligned}$$

Так как $\lambda_1 = \lambda_2 = \lambda_3 = 0$, то

$$\begin{aligned} &Ae_1 = 0, \quad Ae_2 = e_1, \quad \ldots, \quad Ae_p = e_{p-1}, \\ &Af_1 = 0, \quad Af_2 = f_1, \quad \ldots, \quad Af_q = f_{q-1}, \\ &Ag_1 = 0, \quad Ag_2 = g_1, \quad \ldots, \quad Ag_r = g_{r-1}. \end{aligned}$$

Поэтому линейная комбинация векторов $e_1, e_2, \ldots, e_p$, входящая в правую часть равенства (4), будет иметь вид

$$\alpha_1 e_1 + \ldots + \alpha_p e_p - \varkappa_2 e_1 - \ldots - \varkappa_p e_{p-1}.$$

Полагая $\varkappa_2 = \alpha_1, \ldots, \varkappa_p = \alpha_{p-1}$, мы уничтожим здесь все слагаемые кроме одного, равного $\alpha_p e_p$. Проделав ту же операцию в группах $f_1, \ldots, f_q$ и $g_1, \ldots, g_r$, мы получим вектор e', для которого

$$Ae' = \alpha_p e_p + \beta_q f_q + \gamma_r g_r.$$

Случайно может оказаться, что $\alpha_p = \beta_q = \gamma_r = 0$; тогда мы приходим к вектору e', для которого

$$Ae' = 0,$$

и тогда, как и в первом случае, наше преобразование уже в базисе $e'; e_1, e_2, \ldots, e_p; f_1, f_2, \ldots, f_q; \ldots; h_1, h_2, \ldots \ldots, h_s$ имеет нормальную форму. Вектор e' в этом случае образует новую клетку с собственным значением, равным нулю.

Пусть теперь хотя бы один из коэффициентов $\alpha_p, \beta_q, \gamma_r$ отличен от нуля. В этом случае, в отличие от рассмотренных ранее, нам придется для приведения к нормальной форме также изменить некоторые из векторов базиса, уже имеющегося в R'. Расположим числа p, q, r по их величине. Пусть, например, $p > q > r$. Тогда строим новую группу, начиная с e', следующим образом. Полагаем $e'_{p+1} = e'$, $e'_p = Ae'_{p+1}$, $e'_{p-1} = Ae'_p$, ..., $e'_1 = Ae'_2$. Мы имеем, следовательно,

$$\begin{aligned}
e'_{p+1} &= e', \\
e'_p &= Ae'_{p+1} = \alpha_p e_p + \beta_q f_q + \gamma_r g_r, \\
e'_{p-1} &= Ae'_p = \alpha_p e_{p-1} + \beta_q f_{q-1} + \gamma_r g_{r-1}, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
e'_{p-r+1} &= Ae'_{p-r+2} = \alpha_p e_{p-r+1} + \beta_q f_{q-r+1} + \gamma_r g_1, \\
e'_{p-r} &= Ae'_{p-r+1} = \alpha_p e_{p-r} + \beta_q f_{q-r}, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
e'_1 &= Ae'_2 = \alpha_p e_1.
\end{aligned}$$

Заменим теперь векторы $e', e_1, e_2, \ldots, e_p$ базиса векторами

$$e'_1, e'_2, \ldots, e'_p, e'_{p+1},$$

а остальные оставим без изменения. Мы получим тогда нормальную форму преобразования, причем размеры первой клетки увеличились на единицу. Теорема полностью доказана.

Мы видим, что в процессе построения нормальной формы нужно было различать два случая.

1. Случай, когда добавленное собственное значение τ (мы его полагали равным нулю) не совпадает ни с одним из прежних собственных значений $\lambda_1, \ldots, \lambda_k$. В этом случае добавлялась отдельная клетка первого порядка.

2. Случай, когда добавленное собственное значение совпадало с одним из уже имевшихся. В этом случае, вообще говоря, размер одной из имевшихся клеток увеличивался на 1. Если же коэффициенты α, β, γ равны нулю, то, как и в первом случае, добавлялась новая клетка.

§ 21. Инвариантные множители

В этом параграфе мы укажем способ, дающий возможность находить жорданову нормальную форму линейного преобразования. Из результатов этого параграфа будем также вытекать до сих пор еще не доказанная единственность этой формы.

Определение. *Матрицы $\boldsymbol{A}$ и $\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, где $\boldsymbol{C}$ — произвольная невырожденная матрица, называются подобными.*

Если матрица $\boldsymbol{A}_1$ подобна матрице $\boldsymbol{A}_2$, то и обратно, $\boldsymbol{A}_2$ подобна $\boldsymbol{A}_1$. Действительно, пусть

$$\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{A}_2\boldsymbol{C}.$$

Тогда отсюда

$$\boldsymbol{A}_2 = \boldsymbol{C}\boldsymbol{A}_1\boldsymbol{C}^{-1},$$

т. е. если положить $\boldsymbol{C}^{-1} = \boldsymbol{C}_1$, имеем:

$$\boldsymbol{A}_2 = \boldsymbol{C}_1^{-1}\boldsymbol{A}_1\boldsymbol{C}_1$$

и, следовательно, $\boldsymbol{A}_2$ подобна $\boldsymbol{A}_1$.

Легко также показать, что если две матрицы $\boldsymbol{A}_1$ и $\boldsymbol{A}_2$ подобны одной и той же матрице $\boldsymbol{A}$, то они подобны между собой. Действительно, пусть

$$\boldsymbol{A} = \boldsymbol{C}_1^{-1}\boldsymbol{A}_1\boldsymbol{C}_1, \quad \boldsymbol{A} = \boldsymbol{C}_2^{-1}\boldsymbol{A}_2\boldsymbol{C}_2.$$

Тогда $\boldsymbol{C}_1^{-1}\boldsymbol{A}_1\boldsymbol{C}_1 = \boldsymbol{C}_2^{-1}\boldsymbol{A}_2\boldsymbol{C}_2$, т. е.

$$\boldsymbol{A}_1 = \boldsymbol{C}_1\boldsymbol{C}_2^{-1}\boldsymbol{A}_2\boldsymbol{C}_2\boldsymbol{C}_1^{-1},$$

и если положить $\boldsymbol{C}_2\boldsymbol{C}_1^{-1} = \boldsymbol{C}$, то получим:

$$\boldsymbol{A}_1 = \boldsymbol{C}^{-1}\boldsymbol{A}_2\boldsymbol{C},$$

т. е. $\boldsymbol{A}_1$ и $\boldsymbol{A}_2$ подобны.

Пусть $\boldsymbol{A}$ — матрица преобразования A в некотором базисе. При переходе к другому базису матрица $\boldsymbol{A}$ заменяется подобной ей матрицей $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, где $\boldsymbol{C}$ — матрица перехода от первого базиса ко второму (§ 9). Таким образом, подобные матрицы — это матрицы одного и того же линейного преобразования в различных базисах.

Наша задача — по матрице преобразования построить *инварианты* самого преобразования, т. е. выражения, зависящие лишь от самого преобразования A. Другими словами, нам нужно построить функции от элементов матрицы, совпадающие для подобных матриц.

Один такой инвариант установлен уже в § 10. Именно, там было доказано, что характеристический многочлен матрицы $\boldsymbol{A}$, т. е. определитель матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$:

$$D_n(\lambda) = |\boldsymbol{A} - \lambda\boldsymbol{E}|$$

не меняется при замене матрицы $\boldsymbol{A}$ подобной матрицей. Мы построим здесь ряд инвариантов, среди которых будет содержаться и характеристический многочлен; они будут *полной системой инвариантов,* в том смысле, что из их совпадения для двух матриц следует подобие этих матриц.

Пусть $\boldsymbol{A}$ — произвольная матрица n-го порядка. Миноры k-го порядка матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$ суть некоторые многочлены от λ. *Обозначим через* $D_k(\lambda)$ *их наибольший общий делитель* *). В частности, $D_n(\lambda)$ — определитель матрицы $\boldsymbol{A} - \lambda\boldsymbol{E}$, т. е. характеристический многочлен матрицы $\boldsymbol{A}$. В дальнейшем мы покажем, что все $D_k(\lambda)$ являются инвариантами.

Заметим, что $D_n(\lambda)$ делится на $D_{n-1}(\lambda)$. Действительно, по определению $D_{n-1}(\lambda)$ все миноры $(n-1)$-го порядка делятся на $D_{n-1}(\lambda)$. Разлагая определитель на сумму произведений элементов какой-нибудь строки на их алгебраические дополнения, мы получаем, что и определитель $D_n(\lambda)$ делится на $D_{n-1}(\lambda)$. Аналогично, $D_{n-1}(\lambda)$ делится на $D_{n-2}(\lambda)$ и т. д.

Упражнение. Найти $D_k(\lambda)$ $(k = 1, 2, 3)$ для матрицы

$$\begin{pmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{pmatrix}.$$

Ответ. $D_3(\lambda) = (\lambda - \lambda_0)^3$, $D_2(\lambda) = D_1(\lambda) = 1$.

Лемма 1. *Если* $\boldsymbol{C}$ *— произвольная невырожденная матрица, то общие наибольшие делители миноров k-го порядка матриц* $\boldsymbol{A} - \lambda\boldsymbol{E}$ *и* $\boldsymbol{C}(\boldsymbol{A} - \lambda\boldsymbol{E})$ *совпадают. Аналогичное утверждение имеет место и для* $(\boldsymbol{A} - \lambda\boldsymbol{E})\boldsymbol{C}$.

---

*) Наибольший общий делитель определен с точностью до числового множителя. Мы выбираем $D_k(\lambda)$ так, чтобы старший коэффициент был равен 1. В частности, если миноры k-го порядка взаимно просты, то $D_k(\lambda) = 1$.

Доказательство. Строки матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ являются линейными комбинациями строк матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ с коэффициентами, являющимися элементами матрицы $\boldsymbol{C}$, т. е. не зависящими от λ. Действительно, обозначим через a_{ik} элементы матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ и через a'_{ik} элементы матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$. Тогда, например,

$$a'_{1k}=\sum_{j=1}^{n} c_{1j}a_{jk},$$

т. е. элементы первой строки матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ являются линейными комбинациями строк матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ с коэффициентами c_{1j}. Аналогично показываем это и для других строк.

Поэтому минор матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ разлагается на сумму миноров матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ с некоторыми численными коэффициентами. Следовательно, всякий делитель миноров k-го порядка матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$ будет также делителем миноров того же порядка матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$. Так как от матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ мы можем перейти к матрице $\boldsymbol{A}-\lambda\boldsymbol{E}$ умножением на $\boldsymbol{C}^{-1}$, то и обратно, каждый делитель миноров k-го порядка матрицы $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ является делителем миноров k-го порядка матрицы $\boldsymbol{A}-\lambda\boldsymbol{E}$. Следовательно, у $\boldsymbol{A}-\lambda\boldsymbol{E}$ и $\boldsymbol{C}(\boldsymbol{A}-\lambda\boldsymbol{E})$ общие делители миноров k-го порядка совпадают.

Лемма 2. *У подобных матриц многочлены $D_k(\lambda)$ совпадают.*

Доказательство. Пусть $\boldsymbol{A}$ и $\boldsymbol{A}'=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ — две подобные матрицы. Согласно предыдущей лемме, общие наибольшие делители миноров k-го порядка у $\boldsymbol{A}-\lambda\boldsymbol{E}$ и $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{C}$ совпадают. По той же лемме совпадают между собой общие наибольшие делители миноров k-го порядка у $\boldsymbol{C}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{E})$ и $\boldsymbol{C}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{C}=$ $=\boldsymbol{A}'-\lambda\boldsymbol{E}$. Следовательно, $D_k(\lambda)$ для $\boldsymbol{A}$ и $\boldsymbol{A}'$ равны между собой.

Так как при переходе от одного базиса к другому матрица линейного преобразования заменяется подобной, то из леммы 2 вытекает следующая

Теорема 1. *Пусть A — линейное преобразование. Тогда наибольший общий делитель $D_k(\lambda)$ миноров k-го порядка матрицы $\boldsymbol{A} - \lambda \boldsymbol{E}$, где $\boldsymbol{A}$ — матрица преобразования A в некотором базисе, не зависит от выбора базиса.*

Перейдем к вычислению многочленов $D_k(\lambda)$ для данного линейного преобразования A. В силу теоремы 1, при их вычислении можно пользоваться матрицей линейного преобразования в любом базисе. Выберем базис, в котором матрица линейного преобразования имеет жорданову нормальную форму. Нам нужно, следовательно, вычислить многочлены $D_k(\lambda)$ для матрицы $\boldsymbol{A}$, имеющей нормальную форму.

Найдем сначала все $D_k(\lambda)$ для матрицы n-го порядка вида

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \ldots & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_0 \end{pmatrix}, \tag{1}$$

т. е. для одной «клетки» нормальной формы. Мы имеем $D_n(\lambda) = (\lambda - \lambda_0)^n$. Если в матрице (1) зачеркнуть первый столбец и последнюю строку, то получим матрицу $\boldsymbol{A}_1$, в которой по диагонали стоят единицы, а над диагональю нули. Поэтому $D_{n-1}(\lambda) = 1$. Вычеркивая далее в матрице $\boldsymbol{A}_1$ строки и столбцы с одинаковыми номерами, мы сможем доказать, что и $D_{n-2}(\lambda) = \ldots$ $\ldots = D_1(\lambda) = 1$. Окончательно мы имеем, что *для отдельной клетки* [*матрицы* (1)] *последовательность* $D_k(\lambda)$ *следующая*:

$$(\lambda - \lambda_0)^n,\ 1,\ 1,\ \ldots,\ 1.$$

Далее, заметим: *пусть матрица* $\boldsymbol{B}$ *имеет вид*

$$\begin{pmatrix} \boldsymbol{B}_1 & 0 \\ 0 & \boldsymbol{B}_2 \end{pmatrix},$$

где $\boldsymbol{B}_1$ *и* $\boldsymbol{B}_2$ — *какие-либо матрицы порядков* n_1 *и* n_2. *Тогда отличные от нуля миноры* m*-го порядка матрицы* $\boldsymbol{B}$ *имеют вид*

$$\Delta_m = \Delta_{m_1}^{(1)} \Delta_{m_2}^{(2)}, \quad m_1 + m_2 = m,$$

где $\Delta_{m_1}^{(1)}$ — *миноры* m_1*-го порядка матрицы* $\boldsymbol{B}_1$, *а* $\Delta_{m_2}^{(2)}$ — *миноры* m_2*-го порядка матрицы* $\boldsymbol{B}_2$ *). Действительно, если выделить те из первых n_1 строк, которые входят в состав данного минора, и разложить по ним минор (воспользовавшись теоремой Лапласа), то этот минор будет либо равен нулю, либо иметь вид $\Delta_{m_1}^{(1)} \Delta_{m-m_1}^{(2)}$.

Найдем теперь многочлены $D_k(\lambda)$ для произвольной матрицы $\boldsymbol{A}$, имеющей жорданову нормальною форму. Мы предположим, что в матрице $\boldsymbol{A}$ имеется p клеток, отвечающих собственному значению λ_1, q клеток, отвечающих собственному значению λ_2, и т. д. Обозначим порядки клеток, отвечающих собственному значению λ_1, через $n_1, n_2, \ldots, n_p$ $(n_1 \geqslant n_2 \geqslant n_3 \geqslant \ldots \geqslant n_p)$.

Матрица $\boldsymbol{B} = \boldsymbol{A} - \lambda \boldsymbol{E}$ распадается на отдельные клетки $\boldsymbol{B}_i$, из которых, например, $\boldsymbol{B}_1$ имеет вид

$$\boldsymbol{B_1} = \begin{pmatrix} \lambda_1 - \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda_1 - \lambda & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_1 - \lambda \end{pmatrix}.$$

*) Отличный от нуля минор Δ_k k-го порядка матрицы $\boldsymbol{B}$ может, конечно, иметь вид $\Delta_k^{(1)}$, т. е. быть составлен из элементов $\boldsymbol{B}_1$. В этом случае мы его запишем формально в виде $\Delta_k = \Delta_k^{(1)} \Delta_0^{(2)}$, где положено $\Delta_0^{(2)} = 1$.

Вычислим сначала $D_n(\lambda)$, т. е. определитель матрицы $\boldsymbol{B}$. Он равен произведению определителей матриц $\boldsymbol{B}_1$, т. е.

$$D_n(\lambda) = (\lambda - \lambda_1)^{n_1+n_2+\ldots+n_p}(\lambda - \lambda_2)^{m_1+m_2+\ldots+m_q}\ldots$$

Перейдем теперь к вычислению $D_{n-1}(\lambda)$. Так как $D_{n-1}(\lambda)$ есть делитель многочлена $D_n(\lambda)$, то $D_{n-1}(\lambda)$ состоит из множителей $\lambda - \lambda_1$, $\lambda - \lambda_2$, ... Вычислим, в какой степени в $D_{n-1}(\lambda)$ входит $\lambda - \lambda_1$. Для этого заметим, что произвольный отличный от нуля минор $(n-1)$-го порядка матрицы $\boldsymbol{B} = \boldsymbol{A} - \lambda\boldsymbol{E}$ имеет вид

$$\Delta_{n-1} = \Delta_{t_1}^{(1)}\Delta_{t_2}^{(2)}\ldots\Delta_{t_k}^{(k)},$$

где $t_1+t_2+\ldots+t_k = n-1$, $\Delta_{t_i}^{(i)}$ миноры порядка t_i матрицы $\boldsymbol{B}_i$. Так как сумма порядков миноров $\Delta_{t_i}^{(1)}, \ldots$ равна $n-1$, то один и только один из этих миноров имеет порядок на единицу ниже, чем порядок соответствующей матрицы $\boldsymbol{B}_i$, т. е. получается из соответствующей клетки матрицы $\boldsymbol{B}$ вычеркиванием одной строки и одного столбца. Мы видели, что в отдельной клетке мы можем вычеркиванием одной строки и одного столбца получить минор, равный единице (см. стр. 234). Поэтому мы можем подобрать Δ_{n-1} так, чтобы какой-нибудь один из миноров $\Delta_{t_i}^{(i)}$ стал равным единице, не меняя при этом остальных, равных определителям соответствующих клеток. Отсюда ясно, что, для того чтобы получить минор, содержащий $\lambda - \lambda_1$ в возможно более низкой степени, достаточно вычеркнуть строку и столбец в клетке, отвечающей λ_1 и имеющей наибольший порядок, а именно порядок n_1. Таким образом, наибольший общий делитель $D_{n-1}(\lambda)$ миноров $(n-1)$-го порядка содержит $\lambda - \lambda_1$ в степени $n_2 + n_3 + \ldots + n_p$.

Аналогично, среди миноров $(n-2)$-го порядка наинизшую степень $\lambda - \lambda_1$ содержит минор Δ_{n-2}, полученный вычеркиванием по строке и столбцу из клеток, соответствующих собственному значению λ_1 и имеющих порядки n_1 и n_2. Таким образом, $D_{n-2}(\lambda)$ содержит $\lambda - \lambda_1$ в степени $n_3 + n_4 + \ldots + n_p$ и т. д. Наконец, $D_{n-p}(\lambda), D_{n-p-1}(\lambda), \ldots, D_1(\lambda)$ вовсе не содержит $\lambda - \lambda_1$.

Совершенно так же мы выясняем, в каких степенях в $D_k(\lambda)$ входят множители $\lambda - \lambda_2$, $\lambda - \lambda_3$, ...

Итак, мы доказали следующее утверждение:

Пусть матрица преобразования A имеет жорданову нормальную форму, в которой имеется p клеток порядков $n_1, n_2, \ldots, n_p$ $(n_1 \geqslant n_2 \geqslant \ldots \geqslant n_p)$, отвечающих собственному значению λ_1, q клеток порядков $m_1, m_2, \ldots, m_q$ $(m_1 \geqslant m_2 \geqslant \ldots \geqslant m_q)$, отвечающих собственному значению λ_2, и т. д.; тогда

$$D_n(\lambda) = (\lambda - \lambda_1)^{n_1+n_2+n_3+\ldots+n_p} (\lambda - \lambda_2)^{m_1+m_2+m_3+\ldots+m_q} \ldots,$$
$$D_{n-1}(\lambda) = (\lambda - \lambda_1)^{n_2+n_3+\ldots+n_p} (\lambda - \lambda_2)^{m_2+m_3+\ldots+m_q} \ldots,$$
$$D_{n-2}(\lambda) = (\lambda - \lambda_1)^{n_3+\ldots+n_p} (\lambda - \lambda_2)^{m_3+\ldots+m_q} \ldots,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

При этом, начиная с $D_{n-p}(\lambda)$ множитель $(\lambda - \lambda_1)^{\cdots}$ заменяется единицей, начиная с $D_{n-q}(\lambda)$, множитель $(\lambda - \lambda_2)^{\cdots}$ заменяется единицей и т. д.

Рассмотрим важный пример. Пусть собственному значению λ_1 отвечает лишь одна клетка, порядок которой равен n_1, собственному значению λ_2 — только одна клетка порядка m_1, λ_3 — одна клетка порядка k_1 и т. д. (т. е. собственные значения, отвечающие различным клеткам, различны). Тогда $D_i(\lambda)$ имеет вид

$$D_n(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{m_1} (\lambda - \lambda_3)^{k_1} \ldots,$$
$$D_{n-1}(\lambda) = 1,$$
$$D_{n-2}(\lambda) = 1,$$
$$\ldots\ldots\ldots$$

Указанный выше общий вид для $D_k(\lambda)$ показывает, что вместо многочленов $D_k(\lambda)$ удобнее ввести их отношения

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}.$$

Многочлены $E_k(\lambda)$ называются *инвариантными множителями.* Таким образом, если *матрица* $\boldsymbol{A}$ *имеет жорданову нормальную форму, в которой имеется* p *«клеток» порядков* $n_1, n_2, \ldots, n_p$ $(n_1 \geqslant n_2 \geqslant \ldots \geqslant n_p)$, *отвечающих собственному значению* λ_1, q *«клеток» порядков* $m_1, m_2, \ldots, m_q$ $(m_1 \geqslant m_2 \geqslant \ldots \geqslant m_q)$, *отвечающих собственному значению* λ_2 *и т. д., то инвариантные множители* $E_k(\lambda)$ *имеют вид*

$$\begin{aligned} &E_n(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{m_1} \ldots, \\ &E_{n-1}(\lambda) = (\lambda - \lambda_1)^{n_2} (\lambda - \lambda_2)^{m_2} \ldots, \\ &E_{n-2}(\lambda) = (\lambda - \lambda_1)^{n_3} (\lambda - \lambda_2)^{m_3} \ldots, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

Мы видим, что задание последовательности инвариантных множителей $E_n(\lambda), E_{n-1}(\lambda), \ldots$ полностью определяет жорданову нормальную форму матрицы $\boldsymbol{A}$; собственные значения λ_i получаются как корни уравнения $E_n(\lambda) = 0$. Размеры же $n_1, n_2, \ldots, n_p$ клеток, отвечающих данному собственному значению λ_1, равны степеням, с которыми двучлен $\lambda - \lambda_1$ содержится соответственно в $E_n(\lambda), E_{n-1}(\lambda), \ldots$

Мы теперь в состоянии сформулировать необходимые и достаточные условия существования базиса, в котором матрица линейного преобразования диагональна.

Для того чтобы существовал базис, в котором матрица преобразования диагональна, необходимо и достаточно, чтобы инвариантные множители этой матрицы имели лишь простые корни.

Действительно, мы видели, что кратности корней $\lambda_1, \lambda_2, \ldots$ инвариантных множителей определяют порядки клеток в жордановой нормальной форме. Простота корней инвариантных множителей означает, таким образом, что эти клетки первого порядка, т. е. что жорданова нормальная форма матрицы сводится к диагональной.

Т е о р е м а 2. *Для того чтобы две матрицы были подобны, необходимо и достаточно, чтобы их инвариантные множители совпадали.*

Д о к а з а т е л ь с т в о. Мы доказали (лемма 2), что у подобных матриц совпадают многочлены $D_k(\lambda)$. Следовательно, совпадают и инвариантные множители $E_k(\lambda)$, являющиеся их отношениями.

Обратно, пусть инвариантные множители матриц $\boldsymbol{A}$ и $\boldsymbol{B}$ совпадают. Мы знаем, что каждая матрица подобна некоторой матрице, имеющей жорданову нормальную форму. Так как инвариантные множители у $\boldsymbol{A}$ и $\boldsymbol{B}$ совпадают, то их нормальные жордановы формы тоже совпадают. Таким образом, матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ подобны одной и той же матрице и, значит, $\boldsymbol{A}$ подобна $\boldsymbol{B}$.

Т е о р е м а 3. *Нормальная форма линейного преобразования однозначно определяется самим линейным преобразованием.*

Д о к а з а т е л ь с т в о. Матрицы преобразования A в различных базисах подобны. Так как подобные матрицы имеют одинаковые инвариантные множители, а инвариантными множителями однозначно определяется нормальная форма, то нормальная форма данного линейного преобразования определена однозначно, и теорема доказана.

Мы в состоянии теперь найти жорданову нормальную форму матрицы линейного преобразования. Для этого достаточно взять матрицу линейного преобразования в каком-нибудь базисе и найти инвариантные множители матрицы $\boldsymbol{A}$. Разложив инвариантные множители в произведение вида $(\lambda-\lambda_1)^n(\lambda-\lambda_2)^m\ldots$, мы будем знать как собственные значения, так и отвечающие им порядки клеток.

§ 22. λ-матрицы

1. λ-матрицей (полиномиальной матрицей) называется матрица, элементами которой являются многочлены относительно некоторой буквы λ. Степенью λ-матрицы называется наивысшая из степеней многочленов, входящих в состав матрицы. Ясно, что λ-матрица степени n может быть представлена в виде

$$A_0\lambda^n + A_1\lambda^{n-1} + \ldots + A_n,$$

где A_k — матрицы, уже не зависящие от λ *). Частный случай λ-матриц нам уже встречался неоднократно, а именно матрицы вида $A-\lambda E$. Результаты, которые мы получим в этом параграфе, для случая λ-матриц вида $A-\lambda E$ содержат как частный случай многие из результатов, полученных в предыдущих параграфах этой главы.

λ-матрицы встречаются во всех вопросах математики. Так, например, решение системы однородных линейных дифференциальных уравнений первого порядка с постоянными коэффициентами

$$\frac{dy_i}{dx} = \sum_{k=1}^{n} a_{ik}y_k \quad (i=1,2,\ldots,n) \tag{1}$$

ищется обычно в виде

$$y_k = c_k e^{\lambda x}, \tag{2}$$

где λ и c_k — некоторые постоянные. Для их определения подставим функции (2) в систему и сократим уравнение на $e^{\lambda x}$. Мы

*) В этом параграфе мы будем матрицы обозначать светлыми буквами.

получим систему линейных уравнений

$$\lambda c_i = \sum_{k=1}^{n} a_{ik} c_k,$$

матрица которой есть $A - \lambda E$, где A — матрица из коэффициентов системы (1). Таким образом, изучение системы дифференциальных уравнений (1) тесно связано с λ-матрицей первой степени относительно λ: $A - \lambda E$.

Аналогично, исследование системы уравнений порядка выше первого приводит к исследованию λ-матриц высших степеней. Например, исследование системы уравнений

$$\sum_{k=1}^{n} a_{ik} \frac{d^2 y_k}{dx^2} + \sum_{k=1}^{n} b_{ik} \frac{dy_k}{dx} + \sum_{k=1}^{n} c_{ik} y_k = 0$$

приводит к исследованию λ-матрицы $A\lambda^2 + B\lambda + C$, где $A = \|a_{ik}\|$, $B = \|b_{ik}\|$, $C = \|c_{ik}\|$.

Мы рассмотрим сейчас вопрос о каноническом виде λ-матриц относительно так называемых элементарных преобразований.

Элементарными преобразованиями λ-матриц называются преобразования следующих типов.

1° Перестановка между собой двух каких-либо строк или столбцов матрицы.

2° Прибавление к строке какой-либо другой строки, умноженной на некоторый многочлен $\varphi(\lambda)$, и, аналогично, прибавление к столбцу другого столбца, умноженного на некоторый многочлен.

3° Умножение строки или столбца на некоторое число, отличное от нуля.

Определение. *Две λ-матрицы называются эквивалентными, если одна может быть получена из другой некоторой последовательностью элементарных преобразований.*

Обратное к каждому элементарному преобразованию есть снова элементарное преобразование. Это легко проверяется для каждого из трех типов элементарных преобразований. Так, если λ-матрица $B(\lambda)$

получается из λ-матрицы $A(\lambda)$ перестановкой строк, то обратной перестановкой строк мы можем из $B(\lambda)$ получить $A(\lambda)$. Если $B(\lambda)$ получается из $A(\lambda)$ прибавлением к k-й строке i-й, умноженной на $\varphi(\lambda)$, то, обратно, $A(\lambda)$ можно получить из $B(\lambda)$ прибавлением к k-й строке i-й, умноженной на $-\varphi(\lambda)$.

Из сделанного замечания следует, что если λ-матрица $K(\lambda)$ эквивалентна $L(\lambda)$, то и обратно, $L(\lambda)$ эквивалентна $K(\lambda)$. В самом деле, пусть из $K(\lambda)$ применением некоторой последовательности элементарных преобразований получается $L(\lambda)$. Тогда, применяя к $L(\lambda)$ в обратном порядке обратные преобразования, мы придем к $K(\lambda)$.

Если две λ-матрицы $K_1(\lambda)$ и $K_2(\lambda)$ эквивалентны некоторой матрице $K(\lambda)$, то они эквивалентны между собой. Действительно, если сначала провести последовательность элементарных преобразований, переводящих $K_1(\lambda)$ в $K(\lambda)$, а затем элементарные преобразования, переводящие $K(\lambda)$ в $K_2(\lambda)$, то мы переведем $K_1(\lambda)$ в $K_2(\lambda)$, т.е. $K_1(\lambda)$ эквивалентна $K_2(\lambda)$.

Основной результат п. 1 этого параграфа состоит в доказательстве теоремы о том, что всякую λ-матрицу можно элементарными преобразованиями привести к диагональному виду. Доказательству этого предложения предпошлем лемму:

Л е м м а 1. *Если элемент $a_{11}(\lambda)$ в λ-матрице $A(\lambda)$ не равен нулю и если не все элементы $a_{ik}(\lambda)$ матрицы $A(\lambda)$ делятся на многочлен $a_{11}(\lambda)$, то можно подобрать эквивалентную $A(\lambda)$ λ-матрицу $B(\lambda)$, для которой элемент $b_{11}(\lambda)$ также не равен нулю и имеет степень более низкую, чем $a_{11}(\lambda)$.*

Д о к а з а т е л ь с т в о. Предположим сначала, что не делящийся на $a_{11}(\lambda)$ элемент матрицы $A(\lambda)$ находится в первой строке. Пусть, например, $a_{1k}(\lambda)$ не делится на $a_{11}(\lambda)$. Тогда $a_{1k}(\lambda)$ можно представить в виде

$$a_{1k}(\lambda) = a_{11}(\lambda)\varphi(\lambda) + b(\lambda),$$

где $\varphi(\lambda)$ — частное, $b(\lambda) \neq 0$ — остаток от деления $a_{1k}(\lambda)$ на $a_{11}(\lambda)$ и, следовательно, степень $b(\lambda)$ ниже, чем степень $a_{11}(\lambda)$. Вычтем из k-го столбца первый, умноженный на $\varphi(\lambda)$. Получим матрицу, где вместо $a_{1k}(\lambda)$ стоит теперь многочлен $b(\lambda)$, имеющий более низкую степень, чем $a_{11}(\lambda)$. Переставляя теперь k-й столбец с первым, мы переведем $b(\lambda)$ в левый верхний угол.

Рассмотрим теперь случай, когда все элементы первой строки и первого столбца делятся на $a_{11}(\lambda)$, а некоторый элемент $a_{ik}(\lambda)$ не делится на $a_{11}(\lambda)$. Этот случай мы сведем к предыдущему следующим образом: $a_{i1}(\lambda)$ делится на $a_{11}(\lambda)$, т. е. имеет вид $a_{i1}(\lambda) = \varphi(\lambda)a_{11}(\lambda)$. Вычтем из i-й строки первую, умноженную на $\varphi(\lambda)$. Тогда $a_{i1}(\lambda)$ заменится нулем, элемент $a_{ik}(\lambda)$ заменится элементом $a'_{ik}(\lambda) = a_{ik}(\lambda) - \varphi(\lambda)a_{1k}(\lambda)$, который по-прежнему не делится на $a_{11}(\lambda)$ (так как $a_{1k}(\lambda)$ по предположению делится на $a_{11}(\lambda)$). Прибавим теперь i-ю строку к первой. Так как на первом месте в i-й строке теперь стоит нуль, то $a_{11}(\lambda)$ не изменится, а на k-м месте в первой строке теперь будет стоять $a_{1k}(\lambda) + a'_{ik}(\lambda) = a_{1k}(\lambda)(1 - \varphi(\lambda)) + a_{ik}(\lambda)$ и, следовательно, в первой строке имеется элемент, не делящийся на $a_{11}(\lambda)$. Мы свели этот случай к рассмотренному выше и, следовательно, лемма доказана.

В дальнейшем мы будем пользоваться также следующим замечанием: если все элементы λ-матрицы $B(\lambda)$ делятся на некоторый многочлен $E(\lambda)$, то после элементарных преобразований над матрицей $B(\lambda)$ мы снова получим матрицу, элементы которой делятся на $E(\lambda)$.

Перейдем теперь к приведению λ-матрицы к диагональному виду.

Мы можем считать, что $a_{11}(\lambda) \neq 0$, так как, если в матрице есть хоть один элемент, отличный от нуля, то перестановками строк и столбцов его можно перевести на это место. Если не все элементы матрицы делятся на $a_{11}(\lambda)$, то мы можем способом, указанном в лемме, заменить матрицу эквивалентной, в которой элемент, стоящий в левом верхнем углу, имеет более низкую степень и по-прежнему отличен от нуля. Если не все элементы делятся на него, то мы можем опять понизить степень этого элемента и т. д. Процесс закончится, когда мы придем к матрице $B(\lambda)$, в которой все элементы делятся на $b_{11}(\lambda)$.

Так как элементы $b_{12}(\lambda), \ldots, b_{1n}(\lambda)$ первой строки делятся на $b_{11}(\lambda)$, то, вычитая из второго, третьего и т. д. столбца первый, умноженный на соответственно подобранные многочлены от λ, мы можем обратить в нуль 2-й, 3-й, ..., n-й элементы первой строки. Аналогично обратим в нуль все элементы, начиная со второго, в первом столбце. Так как в матрице $B(\lambda)$ все элементы делились на $b_{11}(\lambda)$, то в полученной матрице все элементы также делятся на $b_{11}(\lambda)$. Разделим все элементы первой строки на старший коэффициент многочлена $b_{11}(\lambda)$. На первом месте получится многочлен со старшим коэффициентом 1, который мы обозначим через $E_1(\lambda)$, а на остальных местах будут по-прежнему нули.

Мы пришли, таким образом, к матрице следующего вида:

$$\begin{pmatrix} E_1(\lambda) & 0 & 0 & \ldots & 0 \\ 0 & c_{22}(\lambda) & c_{23}(\lambda) & \ldots & c_{2n}(\lambda) \\ 0 & c_{32}(\lambda) & c_{33}(\lambda) & \ldots & c_{3n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & c_{n2}(\lambda) & c_{n3}(\lambda) & \ldots & c_{nn}(\lambda) \end{pmatrix}, \tag{3}$$

все элементы которой делятся на $E_1(\lambda)$.

Мы можем теперь повторить с матрицей $(n-1)$-го порядка $\|c_{ik}(\lambda)\|$ те же операции, что с матрицей n-го порядка. Заметим, что всякое элементарное преобразование матрицы $\|c_{ik}\|$ есть в то же время элементарное преобразование матрицы (3), так как в первой строке и столбце все элементы, кроме $E_1(\lambda)$, равны нулю.

Таким образом, мы обратим в нуль все элементы второй строки и второго столбца, кроме диагонального. Полученный диагональный элемент (старший коэффициент которого также считаем равным единице) обозначим $E_2(\lambda)$. Все элементы $c_{ik}(\lambda)$ делятся на $E_1(\lambda)$. Поэтому все дальнейшие элементарные преобразования всегда приводят нас к элементам, делящимся на $E_1(\lambda)$. В частности, $E_2(\lambda)$ делится на $E_1(\lambda)$.

Мы пришли, таким образом, к матрице, у которой в первых двух строках и столбцах все элементы, кроме диагональных, равны нулю, а по диагонали стоят $E_1(\lambda)$ и $E_2(\lambda)$, причем $E_2(\lambda)$ делится на $E_1(\lambda)$. Мы сможем продолжать этот процесс далее, пока не приведем всю матрицу к диагональному виду. Может, конечно, оказаться, что мы закончим процесс раньше, придя к матрице, состоящей сплошь из нулей.

Итак, доказана следующая

Т е о р е м а 1. *Всякая λ-матрица может быть элементарными преобразованиями приведена к виду*

$$\begin{pmatrix} E_1(\lambda) & 0 & 0 & \dots & 0 \\ 0 & E_2(\lambda) & 0 & \dots & 0 \\ 0 & 0 & E_3(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & E_n(\lambda) \end{pmatrix}, \tag{4}$$

где многочлены $E_k(\lambda)$, стоящие по диагонали, имеют старшие коэффициенты, равные единице, многочлен $E_2(\lambda)$ делится на $E_1(\lambda)$, $E_3(\lambda)$ делится на $E_2(\lambda)$,

$E_4(\lambda)$ на $E_3(\lambda)$ и т. д. Этот вид называется нормальной диагональной формой λ-матрицы.

Конечно, некоторое число последних многочленов $E_k(\lambda)$ в матрице (4) может оказаться равным нулю:

$$E_{r+1}(\lambda) = E_{r+2}(\lambda) = \ldots = 0.$$

З а м е ч а н и е. Мы привели матрицу $A(\lambda)$ к нормальному диагональному виду, в котором каждый диагональный элемент делится на предшествующий. Если поставить себе целью приведение матрицы к какому-нибудь диагональному виду, отбросив требование делимости, то задача решается проще.

Действительно, для того чтобы обратить в нуль все элементы первой строки и первого столбца кроме $a_{11}(\lambda)$, достаточно, чтобы эти элементы (а не все элементы матрицы) делились на $a_{11}(\lambda)$. Как видно из доказательства леммы, для того чтобы этого достигнуть, требуется гораздо меньшее число элементарных преобразований, чем для приведения к нормальной диагональной форме. Обратив в нуль все элементы первой строки и первого столбца, кроме диагонального, мы можем проделать то же самое с оставшейся матрицей $(n-1)$-го порядка и т. д., пока матрица не будет приведена к диагональному виду. Этим путем можно привести матрицу к различным диагональным формам, т. е. диагональная форма не определена однозначно. В следующем пункте этого параграфа мы покажем, что *нормальная* диагональная форма данной λ-матрицы определяется уже однозначно.

У п р а ж н е н и е. Привести λ-матрицу

$$\begin{pmatrix} \lambda - \lambda_1 & 0 \\ 0 & \lambda - \lambda_2 \end{pmatrix}, \quad \lambda_1 \neq \lambda_2,$$

к нормальной диагональной форме.

Ответ.

$$\begin{pmatrix} 1 & 0 \\ 0 & (\lambda-\lambda_1)(\lambda-\lambda_2) \end{pmatrix}.$$

2. В этом пункте мы докажем, что нормальная диагональная форма данной матрицы определена однозначно. Для этого мы построим систему многочленов, связанных с данной λ-матрицей, которые не меняются при элементарных преобразованиях и которыми, как мы увидим, нормальная диагональная форма λ-матрицы вполне определяется.

Пусть дана произвольная λ-матрица. Наибольший общий делитель всех миноров k-го порядка данной λ-матрицы обозначим через $D_k(\lambda)$. Так как наибольший общий делитель определен с точностью до постоянного множителя, то будем считать, что старший коэффициент у $D_k(\lambda)$ равен единице. В частности, если общий наибольший делитель миноров k-го порядка равен постоянной, то мы полагаем $D_k(\lambda) = 1$.

Докажем, что элементарные преобразования не меняют многочленов $D_k(\lambda)$, т.е. что у эквивалентных λ-матриц многочлены $D_k(\lambda)$ совпадают.

Для элементарных преобразований вида 1°, переставляющих строки или столбцы, это очевидно, так как при них каждый минор k-го порядка либо вовсе не меняется, либо меняет знак, либо заменяется другим минором k-го порядка, что, конечно, не меняет общего наибольшего делителя всех таких миноров. Аналогично, не меняют $D_k(\lambda)$ элементарные преобразования вида 3°, так как при этих преобразованиях миноры самое большее умножаются на постоянное. Рассмотрим теперь элементарное преобразование вида 2°, например, прибавим к i-му столбцу j-й, умноженный на $\varphi(\lambda)$. При этом минор k-го порядка вовсе не изменится, если он содержит и i-й и j-й столбцы либо если он не содержит ни одного из них.

В случае же, если минор содержит i-й столбец и не содержит j-го столбца, то его можно представить как комбинацию двух миноров, которые имелись у исходной матрицы. Таким образом, наибольший общий делитель миноров k-го порядка и в этом случае не изменится.

Если все миноры порядка k, а следовательно, и более высоких порядков, матрицы $A(\lambda)$ равны нулю, то мы будем считать $D_k(\lambda) \equiv D_{k+1}(\lambda) \equiv \ldots \equiv D_n(\lambda) = 0$. Заметим, что из совпадения у всех эквивалентных матриц многочленов $D_k(\lambda)$ следует, что эквивалентные матрицы имеют один и тот же ранг.

Найдем многочлены $D_k(\lambda)$ для матрицы, приведенной к нормальной диагональной форме:

$$\begin{pmatrix} E_1(\lambda) & 0 & \ldots & 0 \\ 0 & E_2(\lambda) & \ldots & 0 \\ \hdotsfor{4} \\ 0 & 0 & \ldots & E_n(\lambda) \end{pmatrix}. \tag{5}$$

Заметим, что у диагональной матрицы отличны от нуля только главные миноры, т. е. миноры, в которые входят строки и столбцы с одинаковыми номерами. Эти миноры имеют вид $E_{i_1}(\lambda)E_{i_2}(\lambda)\ldots E_{i_k}(\lambda)$.

Так как $E_2(\lambda)$ делится на $E_1(\lambda)$, $E_3(\lambda)$ делится на $E_2(\lambda)$ и т. д., то наибольший общий делитель миноров первого порядка $D_1(\lambda)$ равен $E_1(\lambda)$. Так как все многочлены $E_k(\lambda)$ делятся на $E_1(\lambda)$, а все многочлены, кроме $E_1(\lambda)$, делятся на $E_2(\lambda)$, то произведение $E_i(\lambda)E_j(\lambda)$ $(i < j)$ всегда делится на минор $E_1(\lambda)E_2(\lambda)$. Таким образом, $D_2(\lambda) = E_1(\lambda)E_2(\lambda)$. Так как, кроме того, все $E_k(\lambda)$, кроме $E_1(\lambda)$ и $E_2(\lambda)$, делятся на $E_3(\lambda)$, то $E_i(\lambda)E_j(\lambda)E_k(\lambda)$ $(i < j < k)$ делится на минор $E_1(\lambda)E_2(\lambda)E_3(\lambda)$ и, следовательно, $D_3(\lambda) = E_1(\lambda)E_2(\lambda)E_3(\lambda)$.

Таким же образом для матрицы (4)

$$D_k(\lambda) = E_1(\lambda)E_2(\lambda)\ldots E_k(\lambda) \quad (k = 1, 2, \ldots, n). \tag{6}$$

Очевидно, что если, начиная с некоторого r, $E_{r+1}(\lambda) = E_{r+2}(\lambda) = \ldots = E_n(\lambda) = 0$, то $D_{r+1}(\lambda) = D_{r+2}(\lambda) = \ldots = D_n(\lambda) = 0$.

Отсюда получается, что для λ-матрицы, имеющей нормальную диагональную форму (5), диагональные элементы $E_k(\lambda)$ вычисляются по формулам

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)}.$$

При этом, если $D_{r+1}(\lambda) = \ldots = D_n(\lambda) = 0$, то надо положить $E_{r+1}(\lambda) = \ldots = E_n(\lambda) = 0$.

Многочлены $E_k(\lambda)$ *называются инвариантными множителями.* В § 20 мы уже определили их для матриц вида $A - \lambda E$.

Т е о р е м а 2. *Нормальная диагональная форма данной λ-матрицы $A(\lambda)$ определяется по ней однозначно. Если $D_k(\lambda)$ $(k = 2, 3, \ldots, r)$ — наибольший общий делитель миноров k-го порядка матрицы $A(\lambda)$, а $D_{r+1}(\lambda) = \ldots = D_n(\lambda) = 0$, то элементы нормальной диагональной формы* (5) *определяются по формулам*

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)} \quad (k = 1, 2, \ldots, r),$$

а

$$E_{r+1}(\lambda) = E_{r+2}(\lambda) = \ldots = E_n(\lambda) = 0.$$

Д о к а з а т е л ь с т в о. Мы показали, что при элементарных преобразованиях многочлены $D_k(\lambda)$ не меняются. Поэтому, если матрица $A(\lambda)$ эквивалентна диагональной нормальной матрице (5), то $D_k(\lambda)$ у них совпадают. Так как для матрицы (5) мы получили, что

$$D_k(\lambda) = E_1(\lambda) \ldots E_k(\lambda) \quad (k = 1, 2, \ldots, r;\ r \leqslant n)$$

и что $D_{r+1}(\lambda) = D_{r+2}(\lambda) = \ldots = D_n(\lambda) = 0$, то теорема доказана.

Следствие. *Для того чтобы две λ-матрицы $A(\lambda)$ и $B(\lambda)$ были эквивалентны, необходимо и достаточно, чтобы для них совпадали многочлены $D_1(\lambda)$, $D_2(\lambda), \ldots, D_n(\lambda)$.*

Действительно, если многочлены $D_k(\lambda)$ у $A(\lambda)$ и $B(\lambda)$ совпадают, то эти матрицы эквивалентны одной и той же нормальной диагональной λ-матрице и, следовательно, эквивалентны между собой.

3. Назовем λ-матрицу $P(\lambda)$ обратимой, если матрица $[P(\lambda)]^{-1}$ также есть λ-матрица. Если $\operatorname{Det} P(\lambda)$ равен постоянной, отличной от нуля, то $P(\lambda)$ обратима. Действительно, элементы обратной матрицы равны минорам $(n-1)$-го порядка, деленным на $\operatorname{Det} P(\lambda)$, т.е. в нашем случае они будут многочленами от λ и, значит $[P(\lambda)]^{-1}$, будет λ-матрицей.

Обратно, если $P(\lambda)$ обратима, то $\operatorname{Det} P(\lambda) = \mathrm{const} \neq \neq 0$. В самом деле, пусть $[P(\lambda)]^{-1} = P_1(\lambda)$. Тогда $\operatorname{Det} P(\lambda) \operatorname{Det} P_1(\lambda) = 1$, а произведение двух многочленов может быть тождественно равно единице лишь в том случае, если многочлены суть отличные от нуля постоянные. Таким образом, мы показали, что *λ-матрица обратима тогда и только тогда, когда ее определитель есть постоянная, отличная от нуля.*

Все обратимые матрицы эквивалентны единичной матрице. В самом деле, определитель обратимой матрицы равен постоянной, отличной от нуля, и значит, $D_n(\lambda) = 1$. Так как $D_n(\lambda)$ делится на $D_k(\lambda)$, то и $D_k(\lambda) = 1$ $(k = 1, 2, \ldots, n)$. Поэтому все инвариантные множители $E_k(\lambda)$ обратимой матрицы равны 1, и нормальная диагональная форма для них будет совпадать с единичной матрицей.

Теорема 3. *Для того чтобы λ-матрицы $A(\lambda)$ и $B(\lambda)$ были эквивалентны между собой, необходимо* *и достаточно, чтобы существовали обратимые λ-матрицы $P(\lambda)$ и $Q(\lambda)$ такие, что*

$$A(\lambda) = P(\lambda)B(\lambda)Q(\lambda). \tag{7}$$

Доказательство. Докажем сначала, что если матрицы $A(\lambda)$ и $B(\lambda)$ эквивалентны, то можно подобрать обратимые матрицы $P(\lambda)$ и $Q(\lambda)$ так, чтобы выполнялось равенство (7). Для этого заметим, что каждое элементарное преобразование λ-матрицы $A(\lambda)$ можно осуществить, умножая $A(\lambda)$ слева или справа на некоторую обратимую λ-матрицу — матрицу этого элементарного преобразования.

Покажем это для всех трех типов элементарных преобразований. Пусть дана λ-матрица

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \dots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \dots & a_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ a_{n1}(\lambda) & a_{n2}(\lambda) & \dots & a_{nn}(\lambda) \end{pmatrix}.$$

Чтобы поменять местами, например, первый и второй столбцы (соответственно строки) этой матрицы, надо умножить $A(\lambda)$ справа (соответственно слева) на матрицу

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{8}$$

полученную из единичной перестановкой тех же столбцов (или, что все равно, строк).

Чтобы умножить второй столбец (соответственно строку) матрицы $A(\lambda)$ на число α, нужно умножить

$A(\lambda)$ справа (соответственно слева) на матрицу

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & \alpha & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{9}$$

полученную из единичной также умножением на α второго столбца (или, что все равно, второй строки).

Наконец, чтобы прибавить к первому столбцу $A(\lambda)$ второй, умноженный на $\varphi(\lambda)$, надо умножить $A(\lambda)$ справа на матрицу

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \varphi(\lambda) & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{10}$$

полученную с помощью той же операции из единичной, а чтобы прибавить к первой строке вторую, умноженную на $\varphi(\lambda)$, нужно умножить $A(\lambda)$ слева на матрицу

$$\begin{pmatrix} 1 & \varphi(\lambda) & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{11}$$

которая также получается из единичной с помощью соответствующего элементарного преобразования.

Мы видим, таким образом, что матрицы элементарных преобразований — это матрицы, полученные одним элементарным преобразованием из E, причем, чтобы произвести элементарное преобразование над столбцами, $A(\lambda)$ надо умножать на матрицу преобразования справа, а чтобы преобразовать строки, $A(\lambda)$ надо умножать на соответствующую матрицу слева.

Можно сосчитать определитель каждой из приведенных матриц (8)–(11) и, таким образом, проверить, что он равен отличной от нуля постоянной; следовательно, все эти матрицы обратимы. Так как определитель произведения матриц равен произведению определителей, то и произведение матриц элементарных преобразований есть обратимая матрица.

Так как мы предположили, что $A(\lambda)$ и $B(\lambda)$ эквивалентны, то $A(\lambda)$ можно получить, применяя к $B(\lambda)$ некоторую цепочку элементарных преобразований. Каждое элементарное преобразование можно осуществить, умножая $B(\lambda)$ на обратимую λ-матрицу; следовательно, весь переход от $B(\lambda)$ к $A(\lambda)$ можно получить, умножая $B(\lambda)$ последовательно на некоторую совокупность обратимых λ-матриц слева и аналогично на некоторую совокупность справа. Так как произведение обратимых матриц также есть обратимая матрица, то первая часть теоремы тем самым доказана.

Отсюда следует, что всякая обратимая матрица есть произведение матриц элементарных преобразований. Действительно, всякая обратимая матрица $Q(\lambda)$ эквивалентна единичной матрице и поэтому может быть представлена в виде

$$Q(\lambda) = P_1(\lambda) E P_2(\lambda),$$

где $P_1(\lambda)$ и $P_2(\lambda)$ — произведения матриц элементарных преобразований. Но это значит, что и сама $Q(\lambda) = P_1(\lambda)P_2(\lambda)$ есть произведение матриц элементарных преобразований.

Этим замечанием можно воспользоваться для доказательства второй половины теоремы. Действительно, пусть дано, что

$$A(\lambda) = P(\lambda)B(\lambda)Q(\lambda),$$

где $P(\lambda)$ и $Q(\lambda)$ обратимы. Но, согласно только что сделанному замечанию, умножение слева на $P(\lambda)$ и справа на $Q(\lambda)$ эквивалентно некоторой совокупности элементарных преобразований, произведенных над $B(\lambda)$. Таким образом, $A(\lambda)$ эквивалентна $B(\lambda)$, что и требовалось доказать.

4. *) В этом пункте мы будем заниматься λ-матрицами вида $A-\lambda E$, где A — постоянная матрица. Основной вопрос, который будет решен, это вопрос об эквивалентности λ-матриц первой степени $A-\lambda E$ и $B-\lambda E$ **).

Легко видеть, что если матрицы A и B подобны, т. е. существует такая невырожденная постоянная матрица C, что $B = C^{-1}AC$, то λ-матрицы $A-\lambda E$ и $B-\lambda E$ эквивалентны. Действительно, если

$$B = C^{-1}AC,$$

то

$$B - \lambda E = C^{-1}(A - \lambda E)C.$$

Так как постоянная невырожденная матрица есть частный случай обратимой λ-матрицы, то, по теореме 3, из этого равенства следует эквивалентность $A - \lambda E$ и $B - \lambda E$.

Мы покажем позднее и обратное, что из эквивалентности λ-матриц $A-\lambda E$ и $B-\lambda E$ следует подобие матриц A и B. Отсюда мы получим, в частности, новое доказательство того, что всякая матрица подобна матрице, имеющей нормальную жорданову форму.

---

*) Этот пункт можно пропустить, так как он содержит другое, независимое от §§ 19 и 20 доказательство того, что всякую матрицу можно привести к жордановой форме.

**) Произвольная λ-матрица первой степени $A_0 + \lambda A_1$, у которой $\operatorname{Det} A_1 \neq 0$, эквивалентна некоторой матрице вида $A - \lambda E$. Действительно, в этом случае $A_0 + \lambda A_1 = -A_1(-A_1^{-1}A_0 - \lambda E)$ и, обозначая $-A_1^{-1}A_0$ через A, имеем $A_0 + \lambda A_1 = -A_1(A - \lambda E)$, откуда по теореме 3 следует эквивалентность матриц $A - \lambda E$ и $A_0 + \lambda A_1$.

Доказательству предпошлем лемму:

Л е м м а 2. *Произвольную λ-матрицу*

$$P(\lambda) = P_0\lambda^n + P_1\lambda^{n-1} + \ldots + P_n$$

можно разделить слева на матрицу вида $A - \lambda E$ (где A — любая постоянная матрица), т. е. можно найти такие матрицы $S(\lambda)$ и R (R постоянна), что,

$$P(\lambda) = (A - \lambda E)S(\lambda) + R.$$

Процесс деления, с помощью которого доказывается лемма, отличается от обычного деления многочленов только тем, что при умножении нельзя изменять порядок сомножителей.

Пусть

$$P(\lambda) = P_0\lambda^n + P_1\lambda^{n-1} + \ldots + P_n,$$

где P_k — постоянные матрицы.

Легко видеть, что λ-матрица

$$P(\lambda) + (A - \lambda E)P_0\lambda^{n-1}$$

будет иметь степень не выше $n - 1$.

Если

$$P(\lambda) + (A - \lambda E)P_0\lambda^{n-1} = P_0'\lambda^{n-1} + P_1'\lambda^{n-2} + \ldots + P_{n-1}',$$

то аналогично многочлен

$$P(\lambda) + (A - \lambda E)P_0\lambda^{n-1} + (A - \lambda E)P_0'\lambda^{n-2}$$

есть многочлен степени не выше $n - 2$. Продолжая этот процесс, мы придем к многочлену

$$P(\lambda) + (A - \lambda E)(P_0\lambda^{n-1} + P_0'\lambda^{n-2} + \ldots)$$

степени не выше нулевой, т. е. не зависящему от λ. Обозначив полученную постоянную матрицу через R, мы получим

$$P(\lambda) = (A - \lambda E)[-P_0\lambda^{n-1} - P_0'\lambda^{n-2} - \ldots] + R.$$

Если теперь обозначить многочлен в квадратных скобках через $S(\lambda)$, то мы будем иметь

$$P(\lambda) = (A - \lambda E)S(\lambda) + R,$$

т. е. лемма доказана.

Аналогично доказывается возможность деления справа, т. е. существование матриц $S_1(\lambda)$ и R_1 таких, что

$$P(\lambda) = S_1(\lambda)(A - \lambda E) + R_1.$$

Заметим кстати, что здесь, как и в обычной теореме Безу, можно утверждать, что

$$R = R_1 = P(A).$$

Теорема 4. *Для того чтобы λ-матрицы $A-\lambda E$ и $B-\lambda E$ были эквивалентны, необходимо и достаточно, чтобы матрицы A и B были подобны.*

Доказательство. Достаточность была доказана в начале этого пункта. Докажем необходимость. Нам надо доказать, что если λ-матрицы $A-\lambda E$ и $B-\lambda E$ эквивалентны, то матрицы A и B подобны. По теореме 3 существуют такие обратимые λ-матрицы $P(\lambda)$ и $Q(\lambda)$, что

$$B - \lambda E = P(\lambda)(A - \lambda E)Q(\lambda). \tag{12}$$

Покажем сначала, что в равенстве (12) $P(\lambda)$ и $Q(\lambda)$ можно заменить постоянными матрицами.

С этой целью разделим $P(\lambda)$ на $B-\lambda E$ слева, $Q(\lambda)$ — справа. Мы получим равенства

$$\begin{aligned} P(\lambda) &= (B - \lambda E)P_1(\lambda) + P_0, \\ Q(\lambda) &= Q_1(\lambda)(B - \lambda E) + Q_0, \end{aligned} \tag{13}$$

где P_0 и Q_0 — постоянные матрицы.

Подставим в формулу (12) выражение для $P(\lambda)$ и произведем умножение. Мы получим:

$$B-\lambda E = (B-\lambda E)P_1(\lambda)(A-\lambda E)Q(\lambda)+P_0(A-\lambda E)Q(\lambda).$$

Во второе слагаемое подставим выражение для $Q(\lambda)$, произведем умножение и перенесем слагаемое $P_0(A - \lambda E)Q_0$ в левую часть равенства. Мы получим:

$$B - \lambda E - P_0(A - \lambda E)Q_0 = K(\lambda), \tag{14}$$

где

$$K(\lambda) = (B - \lambda E)P_1(\lambda)(A - \lambda E)Q(\lambda) + \\ + P_0(A - \lambda E)Q(\lambda)(B - \lambda E). \tag{15}$$

Из равенства (13) следует, что $P_0 = P(\lambda) - (B - \lambda E)P_1(\lambda)$. Заменив этим выражением P_0 во втором слагаемом, получим:

$$K(\lambda) = (B - \lambda E)P_1(\lambda)(A - \lambda E)Q(\lambda) + \\ + P(\lambda)(A - \lambda E)Q_1(\lambda)(B - \lambda E) - \\ - (B - \lambda E)P_1(\lambda)(A - \lambda E)Q_1(\lambda)(B - \lambda E). \tag{16}$$

Но из равенства (12) мы имеем

$$(A - \lambda E)Q(\lambda) = P^{-1}(\lambda)(B - \lambda E),$$
$$P(\lambda)(A - \lambda E) = (B - \lambda E)Q^{-1}(\lambda).$$

Пользуясь этими равенствами, мы можем ввести множитель $B - \lambda E$ в конец первого и начало второго слагаемого в выражении для $K(\lambda)$, после чего получим окончательно

$$K(\lambda) = (B - \lambda E)[P_1(\lambda)P^{-1}(\lambda) + Q^{-1}(\lambda)Q_1(\lambda) - \\ - P_1(\lambda)(A - \lambda E)Q_1(\lambda)](B - \lambda E).$$

Докажем теперь, что $K(\lambda) = 0$. Выражение в квадратных скобках, в силу обратимости $P(\lambda)$ и $Q(\lambda)$, есть многочлен относительно λ. Докажем, что он равен нулю. Предположим, что этот многочлен отличен от нуля и имеет степень m. Нетрудно убедиться тогда, что $K(\lambda)$ имеет степень $m + 2$ и так как $m \geqslant 0$, является многочленом не ниже второй степени. Но из равенства (14) следует, что $K(\lambda)$ не выше первой степени. Следовательно, выражение в квадратных скобках, а значит, и $K(\lambda) = 0$.

Мы получили таким образом, что

$$B - \lambda E = P_0(A - \lambda E)Q_0, \tag{17}$$

где P_0 и Q_0 — постоянные матрицы, т. е. в равенстве (12) можно матрицы $P(\lambda), Q(\lambda)$ заменить постоянными матрицами.

Сравнивая коэффициенты при первой степени λ в обеих частях равенства (17), мы получаем

$$P_0 Q_0 = E,$$

откуда следует невырожденность каждой из матриц P_0 и Q_0 и равенство

$$P_0 = Q_0^{-1}.$$

Сравнение свободных членов дает

$$B = P_0 A Q_0 = Q_0^{-1} A Q_0,$$

т. е. B и A подобны. Теорема доказана.

Так как условием эквивалентности $A - \lambda E$ и $B - \lambda E$ служит совпадение их инвариантных множителей, то из доказанной теоремы следует, что, *для того чтобы матрицы A и B были подобны, необходимо и достаточно, чтобы инвариантные множители у $A - \lambda E$ и $B - \lambda E$ совпадали между собой.* Покажем теперь, что *всякая матрица A подобна матрице, имеющей жорданову нормальную форму.*

Для этого рассмотрим матрицу $A - \lambda E$ и найдем ее инвариантные множители. По этим инвариантным множителям построим, как было указано в § 21, матрицу B, имеющую жорданову нормальную форму. Тогда $B - \lambda E$ имеет те же инвариантные множители, что и $A - \lambda E$, и, значит, B подобна A.

Как было указано на стр. 254 (сноска), изложенное в п. 4 является другим, заменяющим §§ 19 и 20, доказательством того, что всякая матрица подобна матрице, имеющей жорданову нормальную форму. С другой стороны, конечно, содержание п. 4 может быть непосредственно выведено из содержания §§ 19 или 20 и 21.

ГЛАВА IV

ПОНЯТИЕ О ТЕНЗОРАХ

§ 23. Сопряженное (двойственное) пространство

1. Определение сопряженного пространства. Пусть R — линейное пространство. Одновременно с R часто рассматривают другое, тесно связанное с ним пространство, так называемое сопряженное пространство. Для того чтобы сформулировать определение сопряженного пространства, вернемся к понятию линейной функции, введенному нами в п. 1 § 4.

Линейной функцией мы назвали функцию $f(x)$, $x \in R$ удовлетворяющую условиям:

1° $f(x+y) = f(x) + f(y)$,

2° $f(\lambda x) = \lambda f(x)$.

Пусть $e_1, e_2, \ldots, e_n$ — базис в n-мерном пространстве R. Если

$$x = \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n$$

— вектор из R, то линейная функция в R может быть записана в виде (см. § 4)

$$\begin{aligned} f(x) = f(\xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n) = \\ = a_1 \xi^1 + a_2 \xi^2 + \ldots + a_n \xi^n, \end{aligned} \qquad (1)$$

где коэффициенты $a_1, a_2, \ldots, a_n$, определяющие линейную функцию, вычисляются по формулам

$$a_1 = f(e_1),\ a_2 = f(e_2),\ \ldots,\ a_n = f(e_n). \qquad (2)$$

Как это ясно из формулы (1), *при заданном базисе* $e_1, e_2, \ldots, e_n$ *всяким* n *числам* $a_1, a_2, \ldots, a_n$ *отвечает линейная функция, притом только одна.*

Пусть f и g — линейные функции. Их суммой называется функция h, ставящая в соответствие каждому вектору x число $f(x) + g(x)$. Произведением линейной функции f на число α называется функция, ставящая в соответствие каждому вектору x число $\alpha f(x)$.

Очевидно, что сумма линейных функций и произведение линейной функции на число есть снова линейная функция. При этом, если линейная функция f задается числами $a_1, a_2, \ldots, a_n$, а g — числами $b_1, b_2, \ldots, b_n$, то $f + g$ задается числами $a_1 + b_1$, $a_2 + b_2$, $\ldots$, $a_n + b_n$, а αf — числами $\alpha a_1, \alpha a_2, \ldots, \alpha a_n$.

Таким образом, множество заданных в R линейных функций образует линейное пространство.

Определение 1. *Пусть R есть n-мерное пространство. Пространством R', сопряженным к R, мы назовем линейное пространство, векторами которого являются линейные функции, заданные в R. Сумма в R' определяется как сумма линейных функций, а произведение вектора из R' на число — как произведение линейной функции на число.*

Так как при заданном базисе $e_1, e_2, \ldots, e_n$ в пространстве R каждая линейная функция однозначно задается системой n чисел $a_1, a_2, \ldots, a_n$, причем сумме функций отвечает сумма чисел, произведению функции на α произведение чисел a_i на α, то ясно, что R' изоморфно пространству, в котором вектор определен как совокупность n чисел.

Значит, *пространство R', сопряженное к n-мерному пространству R, также n-мерно.*

Если пространства R и R' рассматривают одновременно, то векторы из R называются *контравариантными,* а векторы из R' *ковариантными.* В дальнейшем символы $x, y, \ldots$ будут означать элементы из R, т.е. контравариантные векторы, а $f, g, \ldots$ — элементы из R', т.е. ковариантные векторы.

2. Биортогональные (взаимные) базисы. В дальнейшем мы будем значение линейной функции f в точке x обозначить через (f, x). Таким образом, каждой паре $f \in R'$ и $x \in R$ отнесено число (f, x), причем

1° $(f, x_1 + x_2) = (f, x_1) + (f, x_2)$,
2° $(f, \lambda x) = \lambda(f, x)$,
3° $(\lambda f, x) = \lambda(f, x)$,
4° $(f_1 + f_2, x) = (f_1, x) + (f_2, x)$.

Первое и второе из этих соотношений — это записанные в новых обозначениях равенства

$$f(x_1 + x_2) = f(x_1) + f(x_2) \quad \text{и} \quad f(\lambda x) = \lambda f(x),$$

являющиеся определением линейной функции, а третье и четвертое — определения произведения линейной функции на число и суммы линейных функций. Соотношения 1°–4° напоминают по внешнему виду аксиомы 2° и 3° скалярного произведения (§ 2). Надо лишь подчеркнуть, что в то время, как скалярное произведение есть число, отнесенное паре векторов одного и того же (евклидова) пространства, (f, x) есть число, отнесенное паре векторов, один из которых принадлежит аффинному пространству R, а другой — аффинному пространству R'.

Векторы $x \in R$ и $f \in R'$ мы назовем *ортогональными,* если

$$(f, x) = 0.$$

Таким образом, хотя в аффинном пространстве R (в отличие от евклидова) нет понятия ортогональности двух векторов $x, y \in R$, можно говорить об ортогональности векторов из R к векторам из R'.

Определение 2. *Пусть* $e_1, e_2, \ldots, e_n$ — *базис в* R, *а* $f^1, f^2, \ldots, f^n$ — *базис в* R'. *Мы назовем эти базисы биортогональными (взаимными), если*

$$(f^i, e_k) = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k \end{cases} \quad (i, k = 1, 2, \ldots, n). \tag{3}$$

Введем символ δ_k^i, положив

$$\delta_k^i = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k \end{cases} \quad (i, k = 1, 2, \ldots, n).$$

Тогда

$$(f^i, e_k) = \delta_k^i.$$

Если $e_1, e_2, \ldots, e_n$ — базис в R, то (f, e_k) являются числами a_k, определяющими линейную функцию $f \in R'$ [см. формулу (2)], так как (f, e_k) есть другая форма записи выражения $f(e_k)$.

Из этого замечания следует утверждение:

если $e_1, e_2, \ldots, e_n$ — *произвольный базис в* R, *то в* R' *существует, и притом только один, базис* $f^1, f^2, \ldots, f^n$ *такой, что базисы* $e_1, e_2, \ldots, e_n$ *и* $f^1, f^2, \ldots \ldots, f^n$ *биортогональны (взаимны).*

Действительно, из равенства (3) имеем

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

Таким образом, здесь заданы числа $a_1 = 1$, $a_2 = 0$, $\ldots$, $a_n = 0$. Так как по всяким числам a_i можно построить единственную линейную функцию, то f^1 определено, и при этом однозначно. Аналогично определяется f^2 равенствами

$$(f^2, e_1) = 0, \quad (f^2, e_2) = 1, \quad \ldots, \quad (f^2, e_n) = 0$$

и т. д. Построенные векторы $f^1, f^2, \ldots, f^n$ из R' (линейные функции) линейно независимы, так как отвечающие каждому из них системы чисел $a_1, a_2, \ldots, a_n$ линейно независимы между собой. Мы построили, таким образом, базис, биортогональный базису $e_1, e_2, \ldots, e_n$ и доказали его единственность.

В дальнейшем мы будем пользоваться принятыми в тензорном исчислении обозначениями, а именно, если в некотором выражении один и тот же индекс стоит один раз вверху, а другой раз внизу, то это означает, что по этому индексу производится суммирование (от 1 до n). Сам знак суммирования $\sum$ мы при этом будем опускать.

Например, $\xi^i\eta_i$ означает $\xi^1\eta_1 + \xi^2\eta_2 + \ldots + \xi^n\eta_n$.

Имея в R и R' биортогональные базисы, легко вычислять координаты любого вектора. Пусть e_i и f^k — биортогональные базисы. Найдем координаты ξ^i вектора $x \in R$ в базисе e_i. Мы имеем

$$x = \xi^i e_i.$$

Отсюда

$$(f^k, x) = (f^k, \xi^i e_i) = \xi^i (f^k, e_i) = \xi^i \delta_i^k = \xi^k.$$

Следовательно, *координаты ξ^k вектора x в базисе $e_1, e_2, \ldots, e_n$ вычисляются по формулам*

$$\xi^k = (f^k, x),$$

где f^k — базис, взаимный с базисом e_i.

Аналогично получаем, что *координаты η_i вектора f в базисе f^k вычисляются по формулам*

$$\eta_i = (f, e_i).$$

Пусть $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ — два взаимных (биортогональных) базиса. Выразим величину (f, x) через координаты векторов f и x в базисах $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ соответственно. Пусть

$$x = \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n \quad \text{и} \quad f = \eta_1 f^1 + \eta_2 f^2 + \ldots + \eta_n f^n;$$

тогда

$$(f, x) = (\eta_1 f^1 + \eta_2 f^2 + \ldots + \eta_n f^n,\ \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n) = \\ = (f^i, e_k)\eta_i \xi^k = \delta^i_k \eta_i \xi^k = \eta_i \xi^i.$$

Итак, *если* $e_1, e_2, \ldots, e_n$ — *базис в* R, $f^1, f^2, \ldots, f^n$ — *взаимный с ним базис в* R', *то*

$$(f, x) = \eta_1 \xi^1 + \eta_2 \xi^2 + \ldots + \eta_n \xi^n, \tag{4}$$

где $\xi^1, \xi^2, \ldots, \xi^n$ — *координаты вектора* $x \in R$ *в базисе* $e_1, e_2, \ldots, e_n$, *а* $\eta_1, \eta_2, \ldots, \eta_n$ — *координаты вектора* $f \in R'$ *в базисе* $f^1, f^2, \ldots, f^n$.

Замечание. Если $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ — произвольные базисы в R и R' соответственно, то

$$(f, x) = a^i_k \eta_i \xi^k,$$

где $a^i_k = (f^i, e_k)$.

Мы видим, что во взаимных базисах значение (f, x) записывается особенно просто.

Итак, мы построили соответствие, относящее каждому линейному пространству R другое пространство, а именно сопряженное пространство R'. Мы можем теперь установить соответствие между линейными преобразованиями пространств.

Пусть R_1, R_2 — два линейных пространства и R'_1, R'_2 — пространства, им сопряженные. Каждому линейному преобразованию A пространства R_1 в R_2 мы поставим в соответствие линейное преобразование A' пространства R'_2 в R'_1, которое определим следующим образом.

Пусть $f_2 \in R'_2$, $x_1 \in R_1$. Рассмотрим (f_2, Ax_1); при фиксированном f_2 это линейная функция от x_1, т. е. может быть записана в виде $(f_2, Ax_1) = (f_1, x_1)$, где $f_1 \in R'_1$. Положим по определению $f_1 = A' f_2$. Получаемое преобразование A' называется *сопряженным* к A. Итак, если A — линейное преобразование пространства R_1 в R_2, то сопряженное ему преобразование есть линейное преобразование A' пространства R'_2 в R'_1, задаваемое тождеством

$$(A' f_2, x_1) = (f_2, Ax_1).$$

Установим одно важное свойство операции перехода к сопряженному преобразованию. Пусть A — линейное преобразование пространства R_1 в R_2, B — линейное преобразование пространства R_2 в R_3. Обозначим через BA композицию этих преобразований, т.е. линейное преобразование пространства R_1 в R_3 (по определению $BAx = B(Ax)$ для любого $x \in R_1$).

Покажем, что

$$(BA)' = A'B'.$$

В самом деле, согласно определению имеем:

$$((BA)'f, x) = (f, BAx) \quad \text{для любых} \quad x \in R_1 \quad \text{и} \quad f \in R_3'.$$

С другой стороны, $(A'B'f, x) = (B'f, Ax) = (f, BAx)$. Сопоставляя эти равенства, мы видим, что $(BA)' = A'B'$.

Упражнение. Доказать, что линейное преобразование, сопряженное к A', есть A.

3. Взаимозаменяемость R и R'. В предыдущем изложении R и R' играли различную роль. Мы покажем, что они совершенно равноправны, т.е. что все теоремы останутся справедливыми, если мы поменяем R и R' ролями.

Мы определили R' как совокупность линейных функций в R. Чтобы установить равноправность R и R', докажем, что всякая линейная функция $\varphi(f)$ в R' может быть записана в виде (f, x_0), где x_0 — фиксированный вектор из R.

Пусть $e_1, e_2, \ldots, e_n$ — некоторый базис в R и $f^1, f^2, \ldots, f^n$ — взаимный с ним базис в R'. Линейная функция $\varphi(f)$ может быть записана в виде

$$\varphi(f) = a^1\eta_1 + a^2\eta_2 + \ldots + a^n\eta_n,$$

где $\eta_1, \eta_2, \ldots, \eta_n$ — координаты вектора f в базисе $f^1, f^2, \ldots, f^n$. Рассмотрим вектор x_0, имеющий в базисе $e_1, e_2, \ldots, e_n$ координаты $a^1, a^2, \ldots, a^n$. Тогда, как мы видели в п. 2,

$$(f, x_0) = a^1\eta_1 + a^2\eta_2 + \ldots + a^n\eta_n$$

и, следовательно,

$$\varphi(f) \equiv (f, x_0). \tag{5}$$

Эта формула устанавливает взаимно однозначное соответствие между линейными функциями φ, заданными в R', и векторами $x_0 \in R$.

Мы можем поэтому во всем изложении считать R пространством линейных функций над R', задавая эти линейные функции формулой (5). Этим установлено полное равноправие между R и R'.

Заметим, что при одновременном изучении пространства и сопряженного пространства мы употребляем лишь обычные для векторов операции сложения и умножения на число в каждом пространстве и операцию (f, x), связывающую элементы обоих пространств. Можно поэтому дать другое определение пары сопряженных пространств R и R', при котором их равноправие непосредственно видно. Это определение состоит в следующем: мы рассматриваем пару n-мерных пространств R и R' и каждой паре векторов $x \in R$, $f \in R'$ относим число (f, x), требуя при этом, чтобы выполнялись условия 1°–4° предыдущего пункта и условие

5° *Из $(f, x) = 0$ для любого x следует $f = 0$ и из $(f, x) = 0$ для любого f следует $x = 0$.*

Коротко говоря, пара сопряженных пространств R и R' — это пара n-мерных пространств с введенной дополнительно операцией (f, x), удовлетворяющей перечисленным условиям.

Замечание. В п. 2 мы доказали, что для каждого базиса в R существует и притом единственный взаимный с ним базис в R'. Из равноправия между R и R' следует, что для всякого базиса в R' существует и притом единственный взаимный с ним базис в R.

4. Преобразования координат в R и R'. Если мы рассматриваем координаты векторов $x \in R$ в некотором базисе $e_1, e_2, \ldots, e_n$, то координаты векторов $f \in R'$ мы будем, как правило, рассматривать в базисе $f^1, f^2, \ldots, f^n$, взаимном к базису $e_1, e_2, \ldots, e_n$. Перейдем в R от базиса $e_1, e_2, \ldots, e_n$ к новому базису $e'_1, e'_2, \ldots, e'_n$, и пусть

$$e'_i = c_i^k e_k \tag{6}$$

— формулы этого перехода.

Обозначая через $f^1, f^2, \ldots, f^n$ базис, взаимный с базисом $e_1, e_2, \ldots, e_n$, а через $f'^1, f'^2, \ldots, f'^n$ — базис, взаимный с базисом $e'_1, e'_2, \ldots, e'_n$, найдем матрицу $\|b_i^k\|$ перехода от базиса f^i к базису f'^i.

Найдем сначала обратную ей матрицу $\|u_i^k\|$ перехода от $f'^1, f'^2, \ldots, f'^n$ к $f^1, f^2, \ldots, f^n$:

$$f^k = u_i^k f'^i. \tag{6'}$$

Для этого вычислим двумя способами выражение (f^k, e'_i):

$$(f^k, e'_i) = (f^k, c_i^\alpha e_\alpha) = c_i^\alpha (f^k, e_\alpha) = c_i^k,$$
$$(f^k, e'_i) = (u_i^k f'^i, e'_i) = u_i^k.$$

Отсюда имеем $c_i^k = u_i^k$, т.е. матрица $\|u_i^k\|$ является транспонированной *) к матрице перехода (6). Следовательно, *матрица перехода*

$$f'^k = b_i^k f^i \tag{7}$$

от $f^1, f^2, \ldots, f^n$ *к* $f'^1, f'^2, \ldots, f'^n$ *равна матрице, транспонированной к матрице, обратной матрице* $\|c_i^k\|$ *перехода от* $e_1, e_2, \ldots, e_n$ *к* $e'_1, e'_2, \ldots, e'_n$.

Выясним теперь, как преобразуются координаты векторов в R и в R'. Пусть ξ^i — координаты вектора $x \in R$ в базисе $e_1, e_2, \ldots, e_n$ и ξ'^i — его координаты в новом базисе $e'_1, e'_2, \ldots, e'_n$.

Тогда

$$(f^i, x) = (f^i,\ \xi^1 e_1 + \xi^2 e_2 + \ldots + \xi^n e_n) = \xi^i$$

и

$$(f'^i, x) = (f'^i,\ \xi'^1 e'_1 + \xi'^2 e'_2 + \ldots + \xi'^n e'_n) = \xi'^i.$$

---

*) Мы говорим, что матрица $\|u_i^k\|$ является транспонированной к матрице перехода (6), так как суммирование в (6′) производится по другому индексу.

Поэтому
$$\xi'^i = (f'^i, x) = (b_k^i f^k, x) = b_k^i (f^k, x) = b_k^i \xi^k.$$
Итак,
$$\xi'^i = b_k^i \xi^k, \tag{8}$$
т. е координаты векторов в R преобразуются по тем же формулам, что и векторы взаимного базиса в R'. Аналогично, координаты векторов в R' преобразуются по тем же формулам, что и векторы взаимного базиса в R, т. е.
$$\eta_i' = c_i^k \eta_k. \tag{9}$$

Мы можем таким образом, сформулировать следующее правило: *при переходе от старой системы координат к новой объекты, имеющие нижний индекс, преобразуются матрицей $\|c_i^k\|$, объекты, имеющие верхний индекс, преобразуются матрицей $\|b_i^k\|$, обратной к $\|c_i^k\|$.*

Тот факт, что матрица $\|b_i^k\|$ является обратной к матрице $\|c_i^k\|$, выражается соотношениями
$$c_i^\alpha b_\alpha^j = \delta_i^j, \quad b_i^\alpha c_\alpha^j = \delta_i^j.$$

5. Пространство, сопряженное к евклидову. Ограничимся для простоты евклидовым пространством над полем действительных чисел.

Лемма. *Пусть R есть n-мерное евклидово пространство. Тогда каждую линейную функцию в нем можно записать в виде*
$$f(x) = (x, y),$$
где y — фиксированный вектор, однозначно определяемый линейной функцией f. Обратно, каждый вектор y определяет линейную функцию $f(x) = (x, y)$.

Доказательство. Выберем в R некоторый ортогональный нормированный базис $e_1, e_2, \ldots, e_n$.

Линейная функция $f(x)$ в этом базисе может быть записана в виде

$$f(x) = a_1\xi^1 + a_2\xi^2 + \ldots + a_n\xi^n.$$

Введем вектор y с координатами $a_1, a_2, \ldots, a_n$. Так как базис $e_1, e_2, \ldots, e_n$ — ортогональный, то

$$(x, y) = a_1\xi^1 + a_2\xi^2 + \ldots + a_n\xi^n.$$

Мы доказали, таким образом, существование такого вектора y, что для любого x имеет место равенство

$$f(x) = (x, y).$$

Докажем теперь, что такой вектор определяется однозначно. Пусть

$$f(x) = (x, y_1) \quad \text{и} \quad f(x) = (x, y_2).$$

Тогда

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

т. е.

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

для любого x. Следовательно, $y_1 - y_2 = 0$. Однозначность доказана.

Таким образом, в случае евклидова пространства мы можем каждый элемент f из R' заменить соответствующим элементом y из R и при этом вместо (f, x) писать (y, x). *Так как при одновременном изучении пространства и сопряженного пространства мы употребляем лишь обычные для векторов операции и операцию (f, x), связывающую элементы $f \in R'$ и $x \in R$, то мы можем в случае евклидова пространства заменить f на y, R' на R и (f, x) на (y, x)*, т. е. отождествить евклидово пространство с сопряженным к нему пространством R' *). Это выражают иногда и так: в евклидовом

*) Если R — аффинное n-мерное пространство, то R' также n-мерно и, следовательно, R и R' изоморфны. Но если бы мы ото-

пространстве можно заменить ковариантные векторы контравариантными.

При таком отождествлении пространства R и сопряженного к нему пространства R' понятие ортогональности векторов $x \in R$ и $f \in R'$, введенное в пункте 2, переходит в обычное для евклидова пространства понятие ортогональности двух векторов из R.

Пусть $e_1, e_2, \ldots, e_n$ — произвольный базис в R, а f^1, $f^2, \ldots, f^n$ — взаимный с ним (биортогональный) базис в R'. Так как в случае евклидова пространства R и R' отождествлены, то мы можем считать векторы биортогонального к e_i базиса f^k также векторами из R.

Выясним, как нам найти в этом случае по базису $e_1, e_2, \ldots, e_n$ базис $f^1, f^2, \ldots, f^n$. Выразим сначала e_i через f^k:

$$e_i = g_{ik} f^k.$$

Нам нужно найти коэффициенты g_{ik}. Для этого умножим скалярно обе части равенства на e_α:

$$(e_i, e_\alpha) = g_{ik}(f^k, e_\alpha).$$

Так как, в силу взаимности (биортогональности) базисов f^k и e_α,

$$(f^k, e_\alpha) = \delta^k_\alpha,$$

то

$$(e_i, e_\alpha) = g_{ik}\delta^k_\alpha = g_{i\alpha}.$$

Итак, если базис f^k биортогонален к базису e_i, то

$$e_i = g_{ik} f^k, \tag{10}$$

где матрица g_{ik} вычисляется по формуле

$$g_{ik} = (e_i, e_k).$$

ждествили R и R', нам пришлось бы вместо (f, x) писать (y, x), где $y, x \in R$, т. е. мы тем самым ввели бы в R скалярное произведение.

Отсюда, разрешив соотношение (10) относительно f^i, имеем:

$$f^i = g^{ik} e_k, \tag{11}$$

где g^{ik} — матрица, обратная к g_{ik}, т. е.

$$g^{i\alpha} g_{\alpha k} = \delta^i_k.$$

Упражнение. Показать, что

$$g^{ik} = (f^i, f^k).$$

§ 24. Тензоры

1. Полилинейные функции. В первой главе мы изучили линейные и билинейные функции в n-мерном аффинном пространстве. Их естественным обобщением являются полилинейные функции, зависящие от произвольного числа векторов. При этом мы будем рассматривать функции, зависящие как от векторов из R, так и от векторов из R'.

Определение 1. *Полилинейной функцией*

$$l(x, y, \ldots; f, g, \ldots),$$

зависящей от p векторов $x, y, \ldots \in R$ и q векторов $f, g, \ldots \in R'$ (R' — пространство, сопряженное к R), *называется функция, линейная относительно каждого из аргументов, когда остальные аргументы фиксированы.*

Например, если зафиксированы все векторы, кроме первого, то

$$l(x' + x'', y, \ldots; f, g, \ldots) = \\ = l(x', y, \ldots; f, g, \ldots) + l(x'', y, \ldots; f, g, \ldots),$$
$$l(\lambda x, y, \ldots; f, g, \ldots) = \lambda l(x, y, \ldots; f, g, \ldots).$$

Аналогично

$$l(x, y, \ldots; f' + f'', g, \ldots) = \\ = l(x, y, \ldots; f', g, \ldots) + l(x, y, \ldots; f'', g, \ldots),$$
$$l(x, y, \ldots; \mu f, g, \ldots) = \mu l(x, y, \ldots; f, g, \ldots).$$

То же самое и для других аргументов.

Полилинейную функцию, зависящую от p векторов из R (контравариантных векторов) и q векторов из R' (ковариантных векторов) мы будем называть *полилинейной функцией типа* (p, q). Рассмотрим некоторые полилинейные функции.

Простейшие полилинейные функции — это функции типа $(1, 0)$ и типа $(0, 1)$.

Полилинейная функция типа $(1, 0)$ — это линейная функция от одного вектора в пространстве R, т. е. вектор пространства R' (ковариантный вектор).

Аналогично, как это было показано в п. 3 предыдущего параграфа, полилинейная функция типа $(0, 1)$ задает вектор из R (контравариантный вектор).

Полилинейные функции, зависящие от двух векторов (билинейные функции), бывают трех типов:

α) функции, зависящие от двух векторов из пространства R, — это введенные в § 4 билинейные функции в пространстве R;

β) функции, зависящие от двух векторов в пространстве R', — это билинейные функции в R';

γ) функции, зависящие от одного вектора из R и одного вектора из R'.

Функции третьего типа тесно связаны с линейными преобразованиями. Действительно, пусть

$$y = Ax$$

— линейное преобразование в R. Построим билинейную функцию

$$(f, Ax),$$

линейно зависящую от векторов $x \in R$ и $f \in R'$.

Мы можем, таким образом, каждому линейному преобразованию в R однозначно сопоставить билинейную функцию типа γ.

Как и в § 11 главы II, можно доказать и обратное, т. е. что каждой билинейной функции типа γ отвечает линейное преобразование в R.

2. Выражения для полилинейной функции в данной системе координат. Переход от одной системы координат к другой. Выясним, как выражается полилинейная функция через координаты тех векторов, от которых она зависит. Для того чтобы не писать слишком длинных формул, проведем рассмотрение на случае полилинейной функции $l(x, y; f)$, зависящей от двух векторов из R и одного вектора из R' [функция типа $(2, 1)$].

Выберем в R некоторый базис $e_1, e_2, \ldots, e_n$, а в R' — взаимный с ним базис $f^1, f^2, \ldots, f^n$. Пусть

$$x = \xi^i e_i, \quad y = \eta^j e_j, \quad f = \zeta_k f^k.$$

Тогда

$$l(x, y; f) = l(\xi^i e_i, \eta^j e_j; \zeta_k f^k) = \xi^i \eta^j \zeta_k l(e_i, e_j; f^k).$$

Итак: при заданных в R и соответственно в R' базисах $e_1, e_2, \ldots, e_n$ и $f^1, f^2, \ldots, f^n$ полилинейная функция $l(x, y; f)$ записывается в виде

$$l(x, y; f) = a_{ij}^k \xi^i \eta^j \zeta_k,$$

где ξ^i, соответственно η^j, соответственно ζ_k — координаты вектора x, соответственно y, соответственно f. Числа a_{ij}^k, определяющие функцию $l(x, y; f)$, задаются формулой

$$a_{ij}^k = l(e_i, e_j; f^k)$$

и зависят, таким образом, от выбора базисов в R и R'.

Аналогичная формула имеем место для полилинейной функции общего вида:

$$l(x, y, \ldots; f, g, \ldots) = a_{ij\ldots}^{rs\ldots} \xi^i \eta^j \ldots \lambda_r \mu_s \ldots, \tag{1}$$

где числа $a_{ij\ldots}^{rs\ldots}$, определяющие полилинейную функцию, вычисляются по формулам

$$a_{ij\ldots}^{rs\ldots} = l(e_i, e_j, \ldots; f^r, f^s, \ldots). \tag{2}$$

Выясним теперь, как изменяется система чисел, определяющая полилинейную форму, при изменении базиса.

Пусть в R задан базис $e_1, e_2, \ldots, e_n$ и в R' — взаимный с ним базис $f^1, f^2, \ldots, f^n$. Перейдем к новому базису $e'_1, e'_2, \ldots, e'_n$ в R и взаимному с ним базису f'^1, $f'^2, \ldots, f'^n$ в R'.

Пусть переход от базиса $e_1, e_2, \ldots, e_n$ к базису e'_1, $e'_2, \ldots, e'_n$ задается формулами

$$e'_\alpha = c_\alpha^\beta e_\beta. \tag{3}$$

Тогда переход от базиса $f^1, f^2, \ldots, f^n$ к базису f'^1, $f'^2, \ldots, f'^n$ задается формулами

$$f'^\beta = b_\alpha^\beta f^\alpha, \tag{4}$$

где $\|b_\alpha^\beta\|$ — матрица, транспонированная к матрице, обратной к $\|c_\alpha^\beta\|$.

Формула (3) показывает, что числа c_α^β при фиксированном α являются координатами вектора e'_α в базисе $e_1, e_2, \ldots, e_n$. Аналогично числа b_α^β при фиксированном β являются координатами вектора f'^β в базисе $f^1, f^2, \ldots, f^n$.

Найдем систему чисел $a'^{rs\dots}_{ij\dots}$, определяющих нашу полилинейную функцию в базисах $e'_1, e'_2, \dots, e'_n$ и $f'^1, f'^2, \dots, f'^n$. Мы знаем, что

$$a'^{rs\dots}_{ij\dots} = l(e'_i, e'_j, \dots; f'^r, f'^s, \dots).$$

Поэтому, чтобы найти $a^{rs\dots}_{ij\dots}$, мы должны в формулу (1) вместо $\xi^i, \eta^j, \dots; \lambda_r, \mu_s, \dots$ подставить координаты векторов $e'_i, e'_j, \dots; f'^r, f'^s, \dots$, т. е. числа $c_i^\alpha, c_j^\beta, \dots; b^r_\sigma, b^s_\tau, \dots$ Мы получаем таким образом:

$$a'^{rs\dots}_{ij\dots} = c_i^\alpha c_j^\beta \dots b^r_\sigma b^s_\tau \dots a^{\sigma\tau\dots}_{\alpha\beta\dots}.$$

Итак, *система чисел $a^{rs\dots}_{ij\dots}$, определяющих полилинейную функцию $l(x, y, \dots;\ f, g, \dots)$ во взаимных базисах $e_1, e_2, \dots, e_n$ и $f^1, f^2, \dots, f^n$, при переходе к новым взаимным базисам $e'_1, e'_2, \dots, e'_n$ и $f'^1, f'^2, \dots, f'^n$ преобразуется по формулам*

$$a'^{rs\dots}_{ij\dots} = c_i^\alpha c_j^\beta \dots b^r_\sigma b^s_\tau \dots a^{\sigma\tau\dots}_{\alpha\beta\dots}, \tag{5}$$

где $\|c_i^j\|$ — матрица, определяющая преобразование базиса $e_1, e_2, \dots, e_n$, а $\|b_i^j\|$ — матрица, определяющая преобразование взаимного с ним базиса $f^1, f^2, \dots, f^n$.

Это можно выразить следующей фразой: *на нижние индексы системы чисел $a^{rs\dots}_{ij\dots}$ действует матрица $\|c_i^j\|$, на верхние индексы — матрица $\|b_i^j\|$* (ср. § 23, п. 4, где рассмотрены формулы для преобразования координат ковариантного и контравариантного векторов).

3. Определение тензора. Объекты, с которыми мы встречались на протяжении этой книги (векторы, линейные функции, линейные преобразования, билинейные функции и т. д.), определялись в каждом базисе своей системой чисел. Например, вектор определялся в каждом базисе системой n чисел — своими координатами. Линейная функция определяется в каждом базисе также системой n чисел — своими коэффициентами.

Линейное преобразование определяется в каждом базисе системой n^2 чисел — матрицей линейного преобразования. Билинейная функция определяется в каждом базисе системой n^2 чисел — матрицей этой билинейной формы. При переходе от одной системы координат (базиса) к другой система чисел, определяющая данный объект, преобразуется определенным образом, причем закон преобразования различен для различных объектов. Например, как вектор из R, так и линейная функция в R задаются системой n чисел, однако при переходе к другому базису они преобразуются по-разному. Для полной характеристики встречающейся величины мы должны задать не только значения соответствующих чисел в какой-либо системе координат, но и закон преобразования соответствующей совокупности чисел при переходе к другой системе координат.

В пунктах 1 и 2 этого параграфа мы ввели понятие полилинейной функции, которая определяется в каждом данном базисе системой n^k чисел (2), преобразующихся при переходе к другому базису по формулам (5). В связи с ним вводится следующее определение, играющее важную роль во многих разделах физики, геометрии и алгебры.

О п р е д е л е н и е 2. *Если каждой системе координат в n-мерном аффинном пространстве отнесена система n^{p+q} чисел $a_{ij\dots}^{rs\dots}$* (число нижних индексов обозначено через p, верхних — через q), *причем при переходе от одной системы координат к другой эти числа преобразуются по формуле*

$$a'^{rs\dots}_{ij\dots} = c_i^\alpha c_j^\beta \dots b_\sigma^r b_\tau^s \dots a_{\alpha\beta\dots}^{\sigma\tau\dots}, \tag{6}$$

где $\|c_i^j\|$ — матрица, задающая переход от одного базиса в R к другому, а $\|b_i^j\|$ — матрица, транспонированная к матрице, обратной к $\|c_i^j\|$, то мы говорим, *что нам задан тензор. Этот тензор называется p раз ковариантным и q раз контравариантным. Число $p+q$ называется рангом (валентностью) тензора. Сами числа $a_{ij\ldots}^{rs\ldots}$ называются компонентами тензора.*

Так как система чисел, определяющих полилинейную функцию от p векторов из R и q векторов из R', при изменении базиса преобразуется как раз по формуле (6), то каждой такой полилинейной функции однозначно соответствует тензор ранга $p+q$, p раз ковариантный и q раз контравариантный. Обратно, каждому тензору однозначно отвечает полилинейная функция. В дальнейшем свойства тензоров и операции над ними мы будем изучать на «модели» полилинейных функций, хотя, конечно, полилинейные функции являются лишь одной из возможных реализаций тензоров.

Приведем некоторые примеры тензоров.

1. Скаляр. Если каждой системе координат отнесено одно и то же фиксированное число a, то его формально можно также считать тензором, а именно — тензором нулевого ранга. Тензор нулевого ранга называется скаляром.

2. Контравариантный вектор. Вектору из R в каждом базисе соответствует совокупность n его координат, которые при переходе к другому базису преобразуются по формулам

$$\eta'^i = b_j^i \eta^j$$

и, следовательно, представляют собой контравариантный тензор ранга 1.

3. Линейная функция (ковариантный вектор). Числа a_i, определяющие линейную функцию, преобразуются по формулам

$$a'_i = c_i^j a_j$$

и, следовательно, образуют ковариантный тензор ранга 1.

4. Билинейная функция. Пусть $A(x; y)$ — билинейная форма в пространстве R. Отнесем каждому базису матрицу данной билинейной формы в этом базисе. Мы получим при этом тензор ранга два, дважды ковариантный.

Аналогично, билинейная форма от векторов $x \in R$, $f \in R'$ определяет тензор ранга два, один раз ковариантный и один раз контравариантный, а билинейная форма от векторов $f, g \in R'$ определяет тензор, дважды контравариантный.

5. Линейные преобразования. Пусть A — линейное преобразование в пространстве R. Отнесем каждому базису матрицу $\|a_i^k\|$ преобразования A в этом базисе, т. е. положим

$$Ae_i = a_i^k e_k.$$

Покажем, что $\|a_i^k\|$ есть тензор ранга два, один раз ковариантный и один раз контравариантный. Действительно, пусть переход к новому базису задается формулой

$$e'_i = c_i^\alpha e_\alpha$$

и, следовательно, обратный переход — формулой

$$e_i = b_i^\alpha e'_\alpha, \quad \text{где} \quad b_i^\alpha c_\alpha^k = \delta_i^k.$$

Тогда

$$Ae'_i = Ac_i^\alpha e_\alpha = c_i^\alpha A e_\alpha = c_i^\alpha a_\alpha^\beta e_\beta = c_i^\alpha a_\alpha^\beta b_\beta^k e'_k.$$

Таким образом, матрица $\|a'^k_i\|$ преобразования A в базисе e'_i имеет вид

$$a'^k_i = a_\alpha^\beta c_i^\alpha b_\beta^k,$$

что и доказывает, что матрица линейного преобразования A есть тензор второго ранга, один раз ковариантный и один раз контравариантный.

В частности, единичному преобразованию E в каждом базисе соответствует единичная матрица, т. е. система чисел

$$\delta_i^k = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k. \end{cases}$$

Таким образом, δ_i^k представляет собой простейший тензор ранга два, один раз ковариантный и один раз контравариантный. Тензор δ_i^k интересен тем, что его компоненты в любой системе координат одни и те же.

Упражнение. Показать непосредственно, что если в каждой системе координат задать систему чисел

$$\delta_i^k = \begin{cases} 1 & \text{при} \quad i = k, \\ 0 & \text{при} \quad i \neq k, \end{cases}$$

то это будет тензор.

Докажем теперь два простых предложения о тензорах.

Пусть имеются два тензора одинакового типа. Тогда для равенства тензоров достаточно, чтобы их компоненты в каком-нибудь базисе были соответственно равны. Другими словами, из того, что компоненты этих двух тензоров равны в какой-либо системе координат, следует, что их компоненты соответственно равны в произвольной системе координат. Это предложение очевидно; действительно, так как оба тензора одинакового типа (т. е. имеют одно и то же число ковариантных и контравариантных индексов), то они преобразуются по одним и тем же формулам, и так как их компоненты в одной системе координат по предположению равны, то они равны и в любой другой системе координат. Заметим, что предположение, что оба тензора одинакового типа, является совершенно обязательным. Например, как билинейная форма, так и линейное преобразование определяются в данной системе координат матрицей. Однако из совпадения матриц линейного преобразования и билинейной формы в одной какой-либо системе координат не следует их совпадение в другой.

При заданных p и q мы можем построить тензор типа (p, q), компоненты которого в каком-нибудь одном базисе равны n^{p+q} наперед заданным числам. Докажем это.

Пусть в некотором базисе нам задана система чисел $a_{ij\dots}^{rs\dots}$. Этими числами задается полилинейная функция $l(x, y, \dots; f, \dots)$ по формуле (1) п. 2 этого параграфа, где ξ^i, соотв. η^j и т. д., — координаты векторов x, соотв. y и т. д. в базисе e_i. Так как с полилинейной функцией однозначно связан тензор, то мы получили тем самым тензор, удовлетворяющий поставленным условиям.

4. Тензоры в евклидовом пространстве. Если R есть n-мерное евклидово пространство, то, как мы видели в п. 5 § 23, можно установить изоморфное соответствие между R и R' так, что если $y \in R$ соответствует элементу $f \in R'$, то

$$(f, x) = (y, x)$$

для любого $x \in R$.

Если мы теперь в полилинейной функции, зависящей от p векторов $x, y, \dots$ из R и q векторов $f, g, \dots$ из R' заменим векторы из R' им соответствующими векторами $u, v, \dots$ из R ,то мы получим полилинейную функцию $l(x, y, \dots; u, v, \dots)$, зависящую от $p + q$ векторов из R.

Найдем коэффициенты функции $l(x, y, \dots; u, v, \dots)$ по коэффициентам функции $l(x, y, \dots; f, g, \dots)$.

Пусть $a_{ij\dots}^{rs\dots}$ — коэффициенты полилинейной функции $l(x, y, \dots; f, g, \dots)$, т. е.

$$a_{ij\dots}^{rs\dots} = l(e_i, e_j, \dots; f^r, f^s, \dots),$$

и пусть $b_{ij\ldots rs\ldots}$ — коэффициенты полилинейной функции $l(x, y, \ldots; u, v, \ldots)$, т. е.

$$b_{ij\ldots rs\ldots} = l(e_i, e_j, \ldots; e_r, e_s, \ldots).$$

Мы доказали в п. 5 § 23, что в евклидовом пространстве векторы e_k базиса, биортогонального f^i, выражаются через векторы базиса f^i по формулам

$$e_r = g_{r\alpha} f^\alpha,$$

где

$$g_{ik} = (e_i, e_k).$$

Подставляя вместо $e_r, \ldots$ их выражения, получаем

$$\begin{aligned} b_{ij\ldots rs\ldots} &= l(e_i, e_j, \ldots; e_r, e_s, \ldots) = \\ &= l(e_i, e_j, \ldots; g_{\alpha r} f^\alpha, g_{\beta s} f^\beta, \ldots) = \\ &= g_{\alpha r} g_{\beta s} \ldots l(e_i, e_j, \ldots; f^\alpha, f^\beta, \ldots) = \\ &= g_{\alpha r} g_{\beta s} \ldots a^{\alpha\beta\ldots}_{ij\ldots}. \end{aligned}$$

Ввиду установленного соответствия между полилинейными функциями и тензорами мы можем сформулировать полученный результат для тензоров:

Если $a^{rs\ldots}_{ij\ldots}$ — тензор, построенный в евклидовом пространстве, p раз ковариантный и q раз контравариантный, то по нему можно построить новый тензор $b_{ij\ldots rs\ldots}$, являющийся $p+q$ раз ковариантным. Эта операция называется операцией опускания индексов. Она определяется формулой

$$b_{ij\ldots rs\ldots} = g_{\alpha r} g_{\beta s} \ldots a^{\alpha\beta\ldots}_{ij\ldots}.$$

g_{ik} является дважды ковариантным тензором. Действительно, $g_{ik} = (e_i, e_k)$ представляют собой в данной системе координат коэффициенты некоторой билинейной формы, а именно скалярного произведения. Ввиду его связи со скалярным произведением (метрикой) пространства *тензор* g_{ik} *называется метрическим тензором.*

Совершенно аналогично операции опускания индексов можно ввести операцию поднимания индексов с помощью формулы

$$b^{ij\ldots rs\ldots} = g^{\alpha i} g^{\beta j} \ldots a^{rs\ldots}_{\alpha\beta\ldots},$$

где g^{ik} имеет смысл, указанный в § 23, п. 5.

Упражнение. Показать, что g^{ik} — дважды ковариантный тензор.

5. Операции над тензорами. Ввиду установленной связи между тензорами и полилинейными функциями мы будем определять операции над полилинейными функциями. Запись полученных результатов в произвольном базисе даст нам соответствующую операцию над тензорами.

Сложение тензоров. Пусть

$$l'(x, y, \ldots; f, g, \ldots), \quad l''(x, y, \ldots; f, g, \ldots)$$

— две полилинейные функции от одного и того же числа векторов из R и одного и того же числа векторов из R'. Определим их сумму $l(x, y, \ldots; f, g, \ldots)$ формулой

$$l(x, y, \ldots; f, g, \ldots) = \\ = l'(x, y, \ldots; f, g, \ldots) + l''(x, y, \ldots; f, g, \ldots).$$

Ясно, что эта сумма есть снова полилинейная функция от того же числа векторов из R и из R'. Сложение тензоров определяется поэтому формулой:

$$a^{rs\ldots}_{ij\ldots} = a'^{rs\ldots}_{ij\ldots} + a''^{rs\ldots}_{ij\ldots}.$$

Умножение тензоров. Пусть

$$l'(x, y, \ldots; f, g, \ldots) \quad \text{и} \quad l''(z, \ldots; h, \ldots)$$

— две полилинейные функции, из которых первая зависит от p' векторов из R и q' векторов из R', а вторая —

от p'' векторов из R и q'' векторов из R'. Определим функцию $l(x, y, \ldots, z, \ldots; f, g, \ldots, h, \ldots)$ формулой

$$l(x, y, \ldots, z, \ldots; f, g, \ldots, h, \ldots) = \\ = l'(x, y, \ldots; f, g, \ldots) l''(z, \ldots; h, \ldots).$$

Функция l называется произведением полилинейных функций l' и l''. Покажем, что l есть полилинейная функция, зависящая от $p' + p''$ векторов из R и $q' + q''$ векторов из R'. Действительно, при проверке того, что

$$l(x, y, \ldots, z, \ldots; f, g, \ldots, h, \ldots)$$

есть полилинейная функция, мы фиксируем, по очереди, все векторы, кроме одного; при этом ясно, что l есть линейная функция от вектора, оставшегося незафиксированным.

Выразим компоненты тензора, отвечающего произведению полилинейных функций l' и l'', через компоненты тензоров, отвечающих самим этим полилинейным функциям. Так как

$$a'^{rs\ldots}_{ij\ldots} = l'(e_i, e_j, \ldots; f^r, f^s, \ldots),$$

и

$$a''^{rs\ldots}_{ij\ldots} = l''(e_i, e_j, \ldots; f^r, f^s, \ldots),$$

то

$$a^{rstu\ldots}_{ijkl\ldots} = a^{rs\ldots}_{ij\ldots} a^{tu\ldots}_{kl\ldots}$$

Эта формула определяет, таким образом, произведение двух тензоров.

Свертка тензора. Пусть $l(x, y, \ldots; f, g, \ldots)$ — полилинейная функция, зависящая от p векторов $x, y, \ldots$ из R ($p \geqslant 1$) и q векторов $f, g, \ldots$ из R' ($q \geqslant 1$). Мы построим по ней полилинейную функцию, зависящую от $p - 1$ векторов из R и $q - 1$ векторов из R'. Выберем для этого какой-либо базис $e_1, e_2, \ldots, e_n$

в R и взаимный с ним базис $f^1, f^2, \ldots, f^n$ в R'. Будем теперь вместо x и f подставлять соответственно $e_1, f^1; e_2, f^2; \ldots; e_n, f^n$ и рассмотрим сумму *)

$$l'(y, \ldots; g, \ldots) = l(e_\alpha, y, \ldots; f^\alpha, g, \ldots). \tag{7}$$

Ясно, что каждое слагаемое, а значит и вся сумма, есть полилинейная функция от $y, \ldots$ и $g, \ldots$ Покажем, что, хотя каждое слагаемое зависит от выбора базиса, построенная нами сумма от выбора базиса уже не зависит.

Перейдем для этого к другому базису $e'_1, e'_2, \ldots, e'_n$ и соответственно к взаимному с ним базису $f'^1, f'^2, \ldots, f'^n$. Так как мы не меняем при этом векторов $y, \ldots$ и $g, \ldots$, то мы можем их фиксировать и доказывать наше утверждение для билинейной формы $A(x; f)$. Итак, нам нужно доказать, что если $A(x; f)$ — билинейная форма, то

$$A(e_\alpha; f^\alpha) = A(e'_\alpha; f'^\alpha).$$

Если переход от базиса $e_1, e_2, \ldots, e_n$ к базису $e'_1, e'_2, \ldots, e'_n$ задается формулой

$$e'_k = c_i^k e_k,$$

то переход от базиса $f^1, f^2, \ldots, f^n$ к базису $f'^1, f'^2, \ldots, f'^n$ задается формулой

$$f^k = c_i^k f'^i.$$

Поэтому

$$A(e'_\alpha; f'^\alpha) = A(c_\alpha^k e_k; f'^\alpha) = c_\alpha^k A(e_k; f'^\alpha) =$$
$$= A(e_k; c_\alpha^k f'^\alpha) = A(e_k; f^k),$$

*) Напоминаем, что если в некотором выражении один и тот же индекс (в данном случае α) встречается вверху и внизу, то по нему производится суммирование.

т. е. $A(e_\alpha; f^\alpha)$ действительно не зависит от системы координат.

Найдем по коэффициентам формы $l(x, y, \ldots; f, g, \ldots)$ коэффициенты формы (7). Так как

$$a'^{s\ldots}_{j\ldots} = l'(e_j, \ldots; f^s, \ldots)$$

и

$$l'(e_j, \ldots; f^s, \ldots) = l(e_\alpha, e_j, \ldots; f^\alpha, f^s, \ldots),$$

то

$$a'^{s\ldots}_{j\ldots} = a^{\alpha s\ldots}_{\alpha j\ldots}. \tag{8}$$

Тензор $a'^{s\ldots}_{j\ldots}$, полученный из $a^{rs\ldots}_{ij\ldots}$ по формуле (8), *называется сверткой тензора $a^{rs\ldots}_{ij\ldots}$.*

Ясно, что свертку мы можем провести не обязательно по первому верхнему и первому нижнему индексам. Обязательно лишь, чтобы суммирование производилось по одному ковариантному и одному контравариантному индексу. Если бы мы суммировали, например, по двум нижним индексам, то полученная система чисел не образовывала бы тензора (так как при переходе от одной системы координат к другой эти числа не преобразовывались бы по предписанному тензору закону преобразования).

Заметим, что в случае свертки тензора ранга два мы получаем тензор нулевого ранга (скаляр), т. е. число, не зависящее от системы координат.

Рассмотренная нами в п. 4 операция опускания индексов есть не что иное, как свертка произведения данного тензора и метрического тензора g_{ik} (взятого сомножителем соответствующее число раз). Аналогично, поднимание индексов есть свертка произведения данного тензора и тензора g^{ik}.

Приведем еще пример. Пусть a^k_{ij} — тензор ранга три, а b^m_l — тензор ранга два. Их произведение есть тензор $c^{km}_{ijl} = a^k_{ij} b^m_l$ ранга пять. Если теперь свернуть этот тензор, например, по индексам i и m, то мы получим тензор ранга три. Если мы полученный тензор еще раз свернем, например, по индексам j и k, то мы получим тензор ранга один (вектор).

Пусть a_i^j и b_k^l — два тензора ранга два. Умножением и свертыванием можно построить по ним новый тензор ранга два:

$$c_i^l = a_i^\alpha b_\alpha^l.$$

Если тензоры a_i^j и b_k^l трактовать как матрицы линейных преобразований, то полученный тензор есть матрица произведения этих преобразований.

Мы можем также построить по данному тензору ранга два a_i^j ряд инвариантов (т. е. чисел, не зависящих от системы координат, — скаляров), а именно:

$$a_\alpha^\alpha,\ a_\alpha^\beta a_\beta^\alpha,\ \ldots$$

Введенные нами операции над тензорами дают нам возможность по данным тензорам строить ряд новых, инвариантно связанных с ними, тензоров.

Приведем некоторые примеры.

Операцией умножения мы можем из векторов построить тензоры сколь угодно высокого ранга. Пусть, например, ξ^i — координаты контравариантного, а η_j — координаты ковариантного вектора. Тогда $\xi^i\eta_j$ есть тензор ранга два. Аналогично, взяв большее число векторов, можно получить тензоры более высокого ранга. Заметим, что не всякий тензор можно получить умножением векторов. Можно, однако, доказать, что всякий тензор может быть получен из векторов (тензоров ранга один) операциями сложения и умножения.

Целым рациональным инвариантом от данной системы тензоров называется многочлен от компонент тензора, который не меняется при замене компонент тензоров в какой-нибудь системе координат их компонентами в другой системе координат.

Имеет место следующая теорема, которую мы не будем доказывать:

Если задана некоторая система тензоров, то всякий целый рациональный инвариант, построенный по данным тензорам, можно получить из них операциями перемножения тензоров, сложения и умножения на числа и полного свертывания (т. е. свертывания по всем индексам).

6. Симметрические и знакопеременные (антисимметрические) тензоры.

Определение 3. *Тензор называется симметрическим по данным индексам, если при любой перестановке этих индексов компоненты тензора не меняются* *). *Например, симметричность тензора по первым двум индексам означает, что имеет место равенство*

$$a^{st\dots}_{ik\dots} = a^{st\dots}_{ki\dots}.$$

Если $l(x, y, \dots; f, g, \dots)$ — соответствующая тензору $a^{st\dots}_{ik\dots}$ полилинейная форма

$$l(x, y, \dots; f, g, \dots) = a^{st\dots}_{ik\dots}\xi^i\eta^k \dots \lambda_s\mu_t \dots, \tag{9}$$

то симметричность тензора по некоторой группе индексов, как это непосредственно видно из формулы (9), эквивалентна симметричности полилинейной формы по соответствующей группе векторов. Так как для симметричности полилинейной формы по некоторой группе векторов достаточно, чтобы $a^{st\dots}_{ik\dots}$ были симметричны по соответствующим индексам лишь в одной какой-нибудь системе координат, то отсюда следует, что *если компоненты тензора симметричны в одной системе координат, то такая же симметрия будет иметь место и в любой другой системе координат.*

---

*) Само собой разумеется, что речь идет лишь об индексах одной и той же группы (верхний или нижней).

Определение 4. *Знакопеременным (антисимметрическим) называется тензор, который меняет знак при перемене любых двух индексов местами.*

При этом предполагается, конечно, что у этого тензора все индексы одинакового характера, т. е. либо все ковариантные, либо все контравариантные.

Из определения знакопеременного тензора непосредственно следует, что при любой перестановке индексов компоненты тензора не меняются, если перестановка четная, и меняют знак, если перестановка нечетная. Знакопеременным тензорам соответствуют знакопеременные полилинейные функции.

Полилинейная функция $l(x, y, \dots)$, зависящая от p векторов $x, y, \dots$ из R, называется знакопеременной, если при перестановке любой пары из векторов $x, y, \dots$ знак функции меняется.

Для проверки знакопеременности полилинейной функции достаточно проверить знакопеременность компонент соответствующего ей тензора в какой-либо одной системе координат, как это непосредственно следует из формулы (9). С другой стороны, из знакопеременности полилинейной функции следует знакопеременность соответствующего ей тензора (в любой системе координат). Следовательно, если компоненты тензора знакопеременны в какой-либо одной системе координат, то это же имеет место и в любой другой системе координат, и, значит, тензор является знакопеременным (антисимметрическим).

Выясним число независимых компонент антисимметрического тензора. Пусть, например, a_{ik} есть знакопеременный тензор ранга 2. Тогда $a_{ik} = -a_{ki}$ и, следовательно, число различных компонент равно $\dfrac{n(n-1)}{2}$. Аналогично, для знакопеременного тензора a_{ijk} число различных компонент равно $\dfrac{n(n-1)(n-2)}{3!}$, так как компоненты с одинаковыми индексами равны нулю, а компоненты, отличающиеся лишь порядком индексов, определяются одна через другую.

Аналогично, число независимых компонент знакопеременного тензора с k индексами ($k \leqslant n$) равно C_n^k. (Отличных от нуля знакопеременных тензоров с числом индексов больше, чем n, не существует, так как у знакопеременного тензора компоненты хотя бы с двумя одинаковыми индексами равны нулю, а если число индексов превышает n, то у каждой компоненты совпадают, по крайней мере, два индекса.)

Рассмотрим более подробно знакопеременный тензор с n индексами. Так как все группы по n различных индексов, принимающих значение от 1 до n, отличаются лишь порядком, то у такого тензора есть лишь одна независимая компонента и он имеет, таким образом, следующий вид.

Пусть $i_1, i_2, \ldots, i_n$ — некоторая перестановка чисел 1, 2, … , n. Положим $a_{12\ldots n} = a$. Тогда

$$a_{i_1 i_2 \ldots i_n} = \pm a, \tag{10}$$

где «$+$» отвечает четной подстановке, а «$-$» — нечетной.

Упражнение. Показать, что при переходе к другой системе координат число $a_{12\ldots n} = a$ умножится на определитель матрицы перехода.

Напишем полилинейную функцию, соответствующую знакопеременному тензору с n индексами. В силу формулы (10) она имеет вид:

$$l(x, y, \ldots, z) = a_{i_1 i_2 \ldots i_n} \xi^{i_1} \eta^{i_2} \ldots \zeta^{i_n} = a \begin{vmatrix} \xi_1 & \xi_2 & \ldots & \xi_n \\ \eta_1 & \eta_2 & \ldots & \eta_n \\ \ldots & \ldots & \ldots & \ldots \\ \zeta_1 & \zeta_2 & \ldots & \zeta_n \end{vmatrix}.$$

Мы доказали, таким образом, что определитель из координат векторов есть, с точностью до множителя, единственная знакопеременная полилинейная функция от n векторов в n-мерном линейном пространстве.

Операция симметрирования. Мы можем по всякому тензору построить новый тензор, симметрический по некоторой наперед заданной группе индексов. Эта операция называется симметрированием и состоит в следующем.

Пусть задан некоторый тензор, например, $a_{i_1 i_2 \ldots i_n}$; симметрирование его, например, по первым k индексам, состоит в построении тензора

$$a_{(i_1 i_2 \ldots i_k) i_{k+1} \ldots} = \frac{1}{k!} \sum a_{j_1 j_2 \ldots j_k i_{k+1} \ldots},$$

где сумма распространяется по всем перестановкам $j_1, j_2, \ldots, j_k$ индексов $i_1, i_2, \ldots, i_k$. Например,

$$a_{(i_1, i_2)} = \frac{1}{2}(a_{i_1 i_2} + a_{i_2 i_1}).$$

Операция симметрирования тензора по группе из k индексов $i_1, i_2, \ldots, i_k$ обозначается следующим образом:

$$a_{j_1 j_2 \ldots j_l (i_1 i_2 \ldots i_k) \ldots}.$$

Операция альтернирования вводится аналогично операции симметрирования и дает возможность по данному тензору построить тензор, знакопеременный по данной группе индексов. Она определяется следующим образом:

$$a_{[i_1 i_2 \ldots i_k] i_{k+1} \ldots} = \frac{1}{k!} \sum \pm a_{j_1 j_2 \ldots j_k i_{k+1} \ldots},$$

где сумма распространяется по всем перестановкам $j_1, j_2, \ldots, j_k$ индексов $i_1, i_2, \ldots, i_k$, а знак определяется честностью или нечетностью этой перестановки. Например,

$$a_{[i_1, i_2]} = \frac{1}{2}(a_{i_1 i_2} - a_{i_2 i_1}).$$

Операция альтернирования обозначается скобками [ ]; в них заключаются те индексы, по которым тензор альтернируется.

По всяким k векторам $\xi^{i_1}, \eta^{i_2}, \ldots, \zeta^{i_k}$ можно построить антисимметрический тензор

$$a^{i_1 i_2 \ldots i_k} = \xi^{[i_1} \eta^{i_2} \ldots \zeta^{i_k]}, \tag{11}$$

где через $\xi^{[i_1} \eta^{i_2} \ldots \zeta^{i_k]}$ обозначен тензор, полученный альтернированием тензора $\xi^{i_1} \eta^{i_2} \ldots \zeta^{i_k}$. Как нетрудно усмотреть из написанной формулы, компонентами этого тензора являются миноры k-го порядка следующей матрицы из n столбцов:

$$\begin{pmatrix} \xi^1 & \xi^2 & \ldots & \xi^n \\ \eta^1 & \eta^2 & \ldots & \eta^n \\ \ldots & \ldots & \ldots & \ldots \\ \zeta^1 & \zeta^2 & \ldots & \zeta^n \end{pmatrix}.$$

Построенный тензор (11) обладает тем свойством, что если к какому-либо из векторов $\xi^{i_1}, \eta^{i_2}, \ldots$ добавить линейную комбинацию остальных, то тензор $a^{i_1 \ldots i_k}$ от этого не изменится.

Рассмотрим k-мерное подпространство n-мерного пространства R. Поставим вопрос о том, чтобы охарактеризовать это k-мерное подпространство системой чисел, т. е. ввести координаты подпространства.

k-мерное подпространство порождается k линейно независимыми векторами $\xi^{i_1}, \eta^{i_2}, \ldots, \zeta^{i_k}$. При этом разные системы из k векторов могут породить одно и то же подпространство. Однако нетрудно показать, и мы предоставляем это читателю, что если две системы векторов порождают одно и то же подпространство, то построенные по каждой из них тензоры

$$a^{i_1 i_2 \ldots i_k} = \xi^{[i_1} \eta^{i_2} \ldots \zeta^{i_k]}$$

совпадают с точностью до множителя.

Таким образом, тензор $a^{i_1 i_2 \ldots i_k}$, построенный по векторам $\xi^{i_1}, \eta^{i_2}, \ldots, \zeta^{i_k}$, порождающим некоторое подпространство, определяет это подпространство.

§ 25. Тензорное произведение

1. Тензорное произведение $\boldsymbol{R} \otimes \boldsymbol{R}$. В первой главе мы изучали билинейные функции в аффинном пространстве R. Здесь мы покажем, что билинейные функции можно трактовать и как линейные функции в некотором новом пространстве. Это пространство, играющее очень важную роль, называется тензорным произведением R и R (по-другому, тензорым квадратом R) и обозначается $R \otimes R$ или $\overset{2}{\bigotimes} R$. Дадим его определение.

Рассмотрим всевозможные упорядоченные пары x, y элементов из R. Каждую такую пару будем называть *тензорным произведением* x и y и обозначать $x \otimes y$. Образуем формальные конечные суммы таких пар:

$$X = x_1 \otimes y_1 + \ldots + x_k \otimes y_k. \tag{1}$$

При этом формальные суммы, отличающиеся только порядком слагаемых, мы не будем различать между собой. Запись (1) означает, таким образом, только то, что нам задано множество k пар $x_1, y_1; \ldots; x_k, y_k$.

Введем для выражений вида (1) операции сложения и умножения на число. Сумму двух таких выражений определим как результат формального дописывания к первому выражению второго:

$$(x_1 \otimes y_1 + \ldots + x_k \otimes y_k) + (x_{k+1} \otimes y_{k+1} + \ldots + x_{k+l} \otimes y_{k+l}) = \\ = x_1 \otimes y_1 + \ldots + x_{k+l} \otimes y_{k+l}. \tag{2}$$

Произведение на число λ определим так:

$$\lambda(x_1 \otimes y_1 + \ldots + x_k \otimes y_k) = (\lambda x_1) \otimes y_1 + \ldots + (\lambda x_k) \otimes y_k. \tag{3}$$

Элемент $0 \otimes 0$ будем называть *нулем* тензорного произведения и обозначать коротко через 0.

Мы должны еще объяснить, какие выражения вида (1) считаются равными. Будем предполагать, что

1) $(x_1 + x_2) \otimes y - x_1 \otimes y - x_2 \otimes y = 0$ *);

2) $x \otimes (y_1 + y_2) - x \otimes y_1 - x \otimes y_2 = 0$;

3) $(\lambda x) \otimes y - x \otimes (\lambda y) = 0$.

Кроме того, *приравняем нулю любое выражение, получающееся из выражений* 1), 2), *и* 3) *сложением и умножением на число.*

Теперь *два выражения X и X' вида* (1) *будем считать равными, если их можно превратить в одинаковые, прибавляя к X и X' выражения, равные нулю, т. е. если существуют такие выражения $Z = 0$ и $Z' = 0$, что $X + Z$ совпадает с $X' + Z'$.* Очевидно, что введенное отношение равенства рефлексивно (т. е. $X = X$) и симметрично (т. е. из $X = Y$ следует $Y = X$); легко проверить, что оно обладает также и свойством транзитивности (т. е. из $X = Y$ и $Y = Z$ следует $X = Z$).

В результате мы получаем пространство, элементы которого — классы равных между собой выражений вида (1), а сложение и умножение на число определены по формулам (2) и (3).

Заметим, что операции сложения и умножения на число мы ввели до того, как было бы определено отношение равенства двух выражений (1). Поэтому нам нужно было убедиться в корректности определений этих операций; именно нужно показать, что сумма выражений (1) и произведение на число не меняются при замене этих выражений на равные. Эта простая проверка предоставляется читателю.

Нетрудно убедиться, что построенное пространство является линейным пространством. Покажем, напри-

---

*) Более подробно, это означает, что $(x_1+x_2)\otimes y+(-x_1)\otimes y+ + (-x_2)\otimes y = 0$.

мер, что для каждого X существует элемент, ему противоположный. Это достаточно проверить для элементов вида $X = x \otimes y$. Из условия 3) при $\lambda = 1$ получаем $x \otimes y + (-x) \otimes y = 0$, т. е. $Y = (-x) \otimes y$ является элементом, противоположным X.

Построенное линейное пространство называется *тензорным произведением* R на R и обозначается через $R \otimes R$.

Итак, *мы определили тензорное произведение* $R \otimes R$ *как линейное пространство, элементами которого являются формальные выражения вида* $x_1 \otimes y_1 + \ldots \ldots + x_k \otimes y_k$, *где* x_i, y_i — *элементы из* R. *Точнее, элементами пространства* $R \otimes R$ *являются классы равных между собой выражений вида* $x_1 \otimes y_1 + \ldots + x_k \otimes y_k$ (условие равенства дано выше). *Сложение в* $R \otimes R$ *и умножение на число определяются по формулам* (2) *и* (3).

Отметим, что из условий 1)–3) следует:

$$(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y;$$
$$x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2;$$
$$\lambda(x \otimes y) = (\lambda x) \otimes y = x \otimes (\lambda y).$$

Поэтому в тензорном произведении можно раскрывать скобки по обычному правилу:

$$(\lambda_1 x_1 + \ldots + \lambda_m x_m) \otimes (\mu_1 y_1 + \ldots + \mu_n y_n) = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_i \mu_j (x_i \otimes y_j).$$

2. Связь между билинейными формами в пространстве R и линейными функциями в $R \otimes R$. Покажем теперь, как по билинейной форме на R можно построить линейную функцию на тензорном произведении $R \otimes R$. Пусть задана билинейная форма $f(x, y)$ на R. Сопоставим ей линейную функцию $F(X)$ на $R \otimes R$. Для элементов $X = x \otimes y$ положим

$$F(x \otimes y) = f(x, y);$$

для произвольного $X = x_1 \otimes y_1 + \ldots + x_k \otimes y_k$ полагаем

$$F(X) = f(x_1, y_1) + \ldots + f(x_k, y_k).$$

Чтобы определение $F(X)$ было корректным, нужно, чтобы на равных выражениях функция F принимала одинаковые значения. Убедимся, что это так. Для этого достаточно показать, что $F(X) = 0$ на выражениях вида 1), 2), 3) (стр. 294). Проверим, например, что

$$F((x_1 + x_2) \otimes y - x_1 \otimes y - x_2 \otimes y) = 0.$$

В самом деле, имеем:

$$\begin{aligned} F((x_1 + x_2) \otimes y - x_1 \otimes y - x_2 \otimes y) &= \\ = f(x_1 + x_2, y) + f(-x_1, y) + f(-x_2, y) &= \\ = f(x_1, y) + f(x_2, y) - f(x_1, y) - f(x_2, y) &= 0. \end{aligned}$$

Очевидно, что построенная по $f(x, y)$ функция $F(X)$ на $R \otimes R$ линейна, т.е. $F(X + Y) = F(X) + F(Y)$ и $F(\lambda X) = \lambda F(X)$. Обратно, если $F(X)$ — линейная функция на $R \otimes R$, то ей соответствует билинейная форма на R:

$$f(x, y) = F(x \otimes y).$$

Итак, мы установили естественное взаимно однозначное соответствие между билинейными формами на R и линейными функциями на $R \otimes R$.

Заметим, что это соответствие линейно; именно, если билинейным формам f_1, f_2 отвечают линейные функции F_1 и F_2, то их линейной комбинации $\lambda_1 f_1 + \lambda_2 f_2$ отвечает функция $\lambda_1 F_1 + \lambda_2 F_2$. Таким образом, построенное соответствие является изоморфизмом между пространством $B(R)$ билинейных форм на R и пространством $(R \otimes R)'$ линейных функций на $R \otimes R$.

3. Размерность тензорного произведения $R\otimes R$. Докажем, что $R\otimes R$ — *конечномерное пространство размерности* n^2, *где* n — *размерность* R.

Зададим базис $e_1, \ldots, e_n$ в пространстве R. Пусть x, y — произвольные векторы из R; разложим их по векторам базиса:

$$x = \xi_1 e_1 + \ldots + \xi_n e_n, \quad y = \eta_1 e_1 + \ldots + \eta_n e_n.$$

Тогда

$$x \otimes y = \sum_{i,j=1}^{n} \xi_i \eta_j (e_i \otimes e_j).$$

Таким образом, $x \otimes y$, а значит, и любой другой вектор из $R \otimes R$ является линейной комбинацией n^2 векторов $e_i \otimes e_j$.

Убедимся, что векторы $e_i \otimes e_j$ линейно независимы. Для этого воспользуемся следующей простой леммой, доказательство которой предоставляется читателю.

Лемма 1. *Пусть* L — *линейное пространство и* X_α $(\alpha = 1, \ldots, N)$ — *векторы из* L. *Если для каждого* $\alpha = 1, \ldots, N$ *существует линейная функция* $F_\alpha(X)$ *на* L *такая, что* $F_\alpha(X_\alpha) = 1$ *и* $F_\alpha(X_\beta) = 0$ *при* $\beta \neq \alpha$, *то векторы* X_α *линейно независимы.*

Зададим для каждого $i = 1, \ldots, n$ линейную функцию $f_i(x)$ на R такую, что $f_i(e_i) = 1$ и $f_i(e_j) = 0$ при $j \neq i$. Так как векторы $e_1, \ldots, e_n$ образуют базис в R, то такая функция существует и единственна. Положим

$$f_{ij}(x, y) = f_i(x) f_j(y), \quad i, j = 1, \ldots, n.$$

Функция $f_{ij}(x, y)$ является билинейной формой на R; значит, согласно п. 2 ей соответствует линейная функция $F_{ij}(X)$ на $R \otimes R$ такая, что

$$F_{ij}(x \otimes y) = f_i(x) f_j(y).$$

Убедимся, что эта линейная функция $F_{ij}(X)$ равна 1 на $e_i \otimes e_j$ и равна 0 на остальных векторах $e_{i'} \otimes e_{j'}$. В самом деле, $F_{ij}(e_{i'} \otimes e_{j'}) = f_i(e_{i'}) f_j(e_{j'})$, и наше утверждение сразу следует из определения функций $f_i(x)$ и $f_j(x)$.

В силу леммы этим доказано, что векторы $e_i \otimes e_j$ линейно независимы. Так как, с другой стороны, по ним раскладывается любой вектор в $R \otimes R$, то векторы $e_i \otimes e_j$ образуют базис в $R \otimes R$. Таким образом, размерность $R \otimes R$ равна числу векторов $e_i \otimes e_j$, т. е. равна n^2.

4. Тензорное произведение $\boldsymbol{R_1 \otimes \ldots \otimes R_m}$. Определяя тензорное произведение $R \otimes R$, мы фактически нигде не пользовались тем, что векторы x и y в произведении $x \otimes y$ берутся из одного и того же пространства. Поэтому, повторяя дословно определения п. 1, можно определить также тензорное произведение $R_1 \otimes R_2$ двух различных пространств R_1 и R_2.

Если $e_1, \ldots, e_m$ — базис в R_1, а $f_1, \ldots, f_n$ — базис в R_2, то базисом в $R_1 \otimes R_2$ служат mn векторов $e_i \otimes f_j$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$.

Отметим, что пространства $R_1 \otimes R_2$ и $R_2 \otimes R_1$ различны по определению.

Аналогично определяется тензорное произведение любого числа линейных пространств. Так, например, элементами тензорного произведения $R_1 \otimes R_2 \otimes R_3$ являются формальные суммы

$$x_1 \otimes y_1 \otimes z_1 + \ldots + x_k \otimes y_k \otimes z_k, \tag{4}$$

где x_i — элементы из R_1, y_i — элементы из R_2 и z_i — элементы из R_3. Операции сложения и умножения на число определяются так же, как и в случае двух сомножителей. Читателю предлагается установить, какие выражения вида (4) при этом следует считать равными.

Тензорное произведение m линейных пространств $R_1, \ldots, R_m$ часто обозначают так: $\bigotimes\limits_{i=1}^{m} R_i$. В случае, когда все сомножители R_i совпадают с одним и тем же пространством R, их тензорное произведение называется m-й тензорной степенью и обозначается так: $\bigotimes\limits^{m} R$.

5. Связь между тензорами и элементами тензорных произведений. Мы покажем, что любой тензор в пространстве R можно рассматривать как элемент некоторого тензорного произведения. Сначала убедимся в этом для тензоров ранга 2, дважды ковариантных.

Согласно § 24, п. 3, тензор ранга 2, дважды ковариантный, задается билинейной формой на R. Но мы уже знаем, что между билинейными формами на R и линейными функциями на $R \otimes R$ имеется естественное взаимно однозначное линейное соответствие. Значит, в силу этого соответствия любой тензор ранга 2, дважды ковариантный, можно рассматривать как линейную функцию на $R \otimes R$, т. е. как элемент сопряженного пространства $(R \otimes R)'$.

С другой стороны, мы установим сейчас естественный изоморфизм $(R \otimes R)' \cong R' \otimes R'$. В силу этого изоморфизма любой тензор ранга 2, дважды ковариантный, можно рассматривать как элемент из тензорного произведения $R' \otimes R'$.

Построим изоморфизм $(R \otimes R)' \cong R' \otimes R'$. Пусть $F \in R' \otimes R'$, т. е.

$$F = f^1 \otimes g^1 + \ldots + f^l \otimes g^l,$$

где f^i, g^i — линейные функции на R. Мы должны сопоставить F элемент из $(R \otimes R)'$, т. е. линейную функцию $F(X)$ на $R \otimes R$. Определим эту функцию по формулам

$$F(x \otimes y) = f^1(x)g^1(y) + \ldots + f^l(x)g^l(x),$$
$$F(x_1 \otimes y_1 + \ldots + x_k \otimes y_k) = F(x_1 \otimes y_1) + \ldots + F(x_k \otimes y_k).$$

Читателю предлагается убедиться, что эти формулы действительно определяют линейную функцию на $R\otimes R$ и что построенное соответствие — изоморфизм.

Рассмотрим теперь тензоры ранга 2, дважды контравариантные. Каждый из них задается билинейной формой на R'. Но между билинейными формами на R' и линейными функциями на $R'\otimes R'$ имеется естественное взаимно однозначное линейное соответствие. Значит, тензоры ранга 2, дважды контравариантные, можно рассматривать как элементы пространства $(R'\otimes R')'\cong$ $\cong R\otimes R$.

Перейдем к общему случаю. Рассмотрим тензоры ранга $p+q$, p раз ковариантные и q раз контравариантные. Из § 23, п. 3, мы знаем, что им однозначно отвечают полилинейные функции $l(x, y, \dots; f, g, \dots)$ от p векторов $x, y, \dots$ из R и q векторов $f, g, \dots$ из R'.

Подобно тому, как это делалось в п. 2 для билинейных форм, можно установить естественное взаимно однозначное линейное соответствие между такими полилинейными функциями $l(x, y, \dots; f, g, \dots)$ и линейными функциями на тензорном произведении $\underbrace{R\otimes\dots\otimes R}_{p \text{ раз}}\otimes\underbrace{R'\otimes\dots\otimes R'}_{q \text{ раз}}$.

Именно, если $F(X)$ — линейная функция на тензорном произведении $\underbrace{R\otimes\dots\otimes R}_{p \text{ раз}}\otimes\underbrace{R'\otimes\dots\otimes R'}_{q \text{ раз}}$, то ей отвечает полилинейная функция $l(x, y, \dots; f, g, \dots)$ от p векторов из R и q векторов из R':

$$l(x, y, \dots; f, g, \dots) = F(x\otimes y\otimes\dots\otimes f\otimes g\otimes\dots). \tag{5}$$

Обратно, если задана полилинейная функция $l(x, y, \dots; f, g, \dots)$, то существует (и притом единственная) линейная функция на тензорном произведении $\underbrace{R\otimes\dots\otimes R}_{p \text{ раз}}\otimes\underbrace{R'\otimes\dots\otimes R'}_{q \text{ раз}}$, удовлетворяющая соотношению (5). (Доказать.)

Значит, тензоры ранга $p+q$, p раз ковариантные и q раз контравариантные, можно рассматривать как линейные функции на $\underbrace{R\otimes\ldots\otimes R}_{p\ \text{раз}}\otimes\underbrace{R'\otimes\ldots\otimes R'}_{q\ \text{раз}}$, т. е. как элементы из сопряженного пространства

$$(\underbrace{R\otimes\ldots\otimes R}_{p\ \text{раз}}\otimes\underbrace{R'\otimes\ldots\otimes R'}_{q\ \text{раз}})'\cong$$

$$\cong\underbrace{R'\otimes\ldots\otimes R'}_{p\ \text{раз}}\otimes\underbrace{R\otimes\ldots\otimes R}_{q\ \text{раз}}.$$

6. Тензорное произведение линейных преобразований. Мы научились по каждой паре линейных пространств R_1, R_2 строить новое линейное пространство — их тензорное произведение $R_1\otimes R_2$. Однако этим задача не заканчивается. Можно еще по линейным преобразованиям в каждом из пространств R_1, R_2 построить линейное преобразование в их тензорном произведении.

Итак, пусть заданы линейное преобразование A пространства R_1 в R_1 и линейное преобразование B пространства R_2 в R_2 *). Мы построим по ним линейное преобразование пространства $R_1\otimes R_2$ в $R_1\otimes R_2$, которое будем называть тензорным произведением преобразований A и B и обозначать через $A\otimes B$.

Рассмотрим тензорное произведение $R_1\otimes R_2$ пространств R_1 и R_2. Напомним, что элементами $R_1\otimes R_2$ являются формальные суммы

$$x_1\otimes y_1+\ldots+x_k\otimes y_k,$$

где $x_i\in R_1$, $y_i\in R_2$, $i=1,\ldots,k$.

Тензорным произведением $A\otimes B$ линейного преобразования A пространства R_1 в R_1 и линейного преобразования B пространства R_2 в R_2 называется линейное

*) Иногда это записывают так: $A\colon R_1\longrightarrow R_1$ и $B\colon R_2\longrightarrow R_2$.

преобразование C пространства $R_1 \otimes R_2$ в $R_1 \otimes R_2$, определяемое следующим образом:

$$C(x_1 \otimes y_1 + \ldots + x_k \otimes y_k) = \\ = (Ax_1) \otimes (By_1) + \ldots + (Ax_k) \otimes (By_k) \text{ *)}.$$

Более общо, если имеются два линейных пространства R_1, S_1, два других линейных пространства R_2, S_2 и линейные преобразования $A\colon R_1 \longrightarrow S_1$ и $B\colon R_2 \longrightarrow S_2$, то можно аналогично определить линейное преобразование

$$A \otimes B\colon R_1 \otimes R_2 \longrightarrow S_1 \otimes S_2.$$

Отметим, что каждому линейному преобразованию A пространства R_1 естественным образом отвечает линейное преобразование пространства $R_1 \otimes R_2$, а именно $A \otimes 1$, где 1 — единичное преобразование; аналогично каждому линейному преобразованию B пространства R_2 можно поставить в соответствие линейное преобразование $1 \otimes B$ пространства $R_1 \otimes R_2$.

Установим, как выражается матрица линейного преобразования $C = A \otimes B$ через матрицы преобразований A и B. Зададим базис $e_1, \ldots, e_m$ в пространстве R_1 и базис $f_1, \ldots, f_n$ в пространстве R_2. Тогда векторы $e_i \otimes f_j$ образуют базис в тензорном произведении $R_1 \otimes R_2$.

Пусть $\boldsymbol{A} = \|a_{ij}\|$ — матрица преобразования A в базисе $e_1, \ldots, e_m$; $\boldsymbol{B}$ — матрица преобразования B в базисе $f_1, \ldots, f_n$, т. е.

$$Ae_k = \sum_{i=1}^{m} a_{ik} e_i, \quad Bf_l = \sum_{j=1}^{n} b_{jl} f_j.$$

Тогда $C = A \otimes B$ преобразует базисные векторы $e_k \otimes f_l$

---

*) Легко проверить, что определение корректно, т. е. равные выражения преобразуются в равные.

по следующей формуле:

$$C(e_k \otimes f_l) = (Ae_k) \otimes (Bf_l) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ik} b_{jl} (e_k \otimes f_l).$$

Таким образом, матрица линейного преобразования C есть матрица $\boldsymbol{C} = \|c_{ij,kl}\|$ порядка mn, строки и столбцы которой занумерованы парами индексов (i, j), $i = 1, \ldots, m$; $j = 1, \ldots, n$. При этом $c_{ij,kl} = a_{ik} b_{jl}$. Такая матрица $\boldsymbol{C}$ называется *кронекеровским произведением* матриц $\boldsymbol{A}$ и $\boldsymbol{B}$.

Упражнение. Доказать, что определитель кронекеровского произведения матриц $\boldsymbol{A}$ и $\boldsymbol{B}$ равен произведению определителей матриц $\boldsymbol{A}$ и $\boldsymbol{B}$.

7. Понятие функтора. В этой главе мы рассмотрели несколько типов операций над линейными пространствами, как, например, операции перехода к сопряженному пространству или тензорное умножение. Дадим общее определение таких операций.

Мы говорим, что задан ковариантный функтор (или более подробно, *ковариантный функтор в категории линейных пространств* *)), *если задано правило, сопоставляющее каждому линейному пространству R некоторое линейное пространство $F(R)$ и каждому линейному преобразованию $A\colon R_1 \longrightarrow R_2$ некоторое линейное преобразование $F(A)$ пространства $F(R_1)$ в $F(R_2)$* (в наших обозначениях, $F(A)\colon F(R_1) \longrightarrow F(R_2)$)*. При этом предполагаются выполненными следующие условия*:

1) *если 1 — единичное преобразование в R, то $F(1)$ — единичное преобразование в пространстве $F(R)$*;

2) *если $A\colon R_1 \longrightarrow R_2$ и $B\colon R_2 \longrightarrow R_3$ — два линейных преобразования, то*

$$F(BA) = F(B)F(A).$$

Примером ковариантного функтора является тензорное умножение. Именно, пусть S — фиксированное пространство. Отнесем каждому линейному пространству R пространство $F(R) = R \otimes S$ и каждому линейному преобразованию $A\colon R_1 \longrightarrow R_2$ линейное

---

*) Понятие функтора можно ввести для произвольной категории. Общие определения категории и функтора см., например, в книге: С. Ленг, Алгебра, «Мир», 1968.

преобразование $F(A) = A \otimes 1$ пространства $R_1 \otimes S$ в $R_2 \otimes S$. Нетрудно проверить, что при этом свойства 1) и 2) выполняются; таким образом, F — ковариантный функтор.

Аналогично определяется контравариантный функтор. *Мы говорим, что задан контравариантный функтор F, если задано правило, сопоставляющее каждому линейному пространству R некоторое линейное пространство $F(R)$ и каждому линейному преобразованию $A\colon R_1 \longrightarrow R_2$ некоторое линейное преобразование $F(A)\colon F(R_2) \longrightarrow F(R_1)$. При этом предполагается выполненными условие* 1) *и условие*

2$'$) *если $A\colon R_1 \longrightarrow R_2$ и $B\colon R_2 \longrightarrow R_3$ — два линейных преобразования, то*

$$F(BA) = F(A)F(B).$$

Примером контравариантного функтора является операция перехода к сопряженным пространствам. Именно, отнесем каждому линейному пространству R сопряженное ему пространство $F(R) = R'$ и каждому линейному преобразованию $A\colon R_1 \longrightarrow R_2$ сопряженное преобразование $F(A) = A'$. Нетрудно проверить (см. п. 2 § 23), что при этом свойства 1) и 2$'$) выполняются; таким образом, F — контравариантный функтор.

З а д а ч а. Пусть S — фиксированное линейное пространство. Обозначим через $\mathrm{Hom}(R, S)$ пространство всех линейных преобразований $A\colon R \longrightarrow S$. Для любого линейного пространства R положим $F(R) = \mathrm{Hom}(R, S)$. Мы определили, таким образом, операцию F на множестве линейных пространств. Требуется определить F также на множестве линейных преобразований таким образом, чтобы F стало контравариантным функтором.

8. Симметрическая и внешняя степени. Наряду с тензорным произведением $R \otimes R$ полезно также рассматривать симметрическую степень и внешнюю степень пространства R; особенно важным понятием является внешняя степень. Эти пространства строятся аналогично тензорному произведению.

Начнем с определения симметрического квадрата $S^2(R)$. Напомним, что элементами пространства $R \otimes R$ являются выражения

$$x_1 \otimes y_1 + \ldots + x_k \otimes y_k, \tag{6}$$

где x_i, y_i — элементы из R. При этом предполагается, что

1) $$(x_1 + x_2) \otimes y - x_1 \otimes y - x_2 \otimes y = 0;$$
2) $$x \otimes (y_1 + y_2) - x \otimes y_1 - x \otimes y_2 = 0;$$
3) $$(\lambda x) \otimes y - x \otimes (\lambda y) = 0.$$

Элементы $x \otimes y$ и $y \otimes x$ в $R \otimes R$ являются при $y \neq x$, по определению, различными.

Однако иногда удобно ввести пространство, в котором $x \otimes y = y \otimes x$.

Для этого дополним условия 1)–3) следующим:

4) $$x \otimes y - y \otimes x = 0.$$

Приравняем также нулю и все линейные комбинации выражений 1), 2), 3) и 4). Два выражения X и X' вида (6) будем теперь считать равными, если для них существуют такие выражения $Z = 0$ и $Z' = 0$, что $X + Z$ и $X' + Z'$ совпадают.

В результате мы получим линейное пространство, элементы которого — классы равных между собой выражений вида (6), а операции сложения и умножения на число определены, как и для тензорного произведения $R \otimes R$, по формулам (2) и (3). (Читателю предлагается убедиться в корректности определений этих операций и в том, что все аксиомы линейного пространства здесь выполнены.) Это пространство называется *симметрическим квадратом пространства* R и обозначается через $S^2(R)$.

Упражнение. Доказать, что размерность $S^2(R)$ равна $\frac{n(n+1)}{2}$, где n — размерность R.

Другим важным понятием является внешний квадрат R. Чтобы это пространство построить, дополним условия 1), 2) и 3) следующим условием:

4′) $$x \otimes x = 0.$$

После этого мы определим равенство двух выражений вида (6) подобно тому, как это уже делалось для тензорного произведения $R \otimes R$ и для симметрического квадрата $S^2(R)$. Получаемое линейное пространство, элементы которого — классы равных между собой выражений вида (6), называется *внешним квадратом пространства* R и обозначается через $R \wedge R$ (по-другому, $\overset{2}{\bigwedge} R$).

Л е м м а 2. *В пространстве* $R \wedge R$ *имеет место равенство*

$$x \otimes y + y \otimes x = 0. \tag{7}$$

В самом деле, имеем:

$$x \otimes y + y \otimes x = (x + y) \otimes (x + y) - x \otimes x - y \otimes y.$$

Таким образом, выражение $x \otimes y + y \otimes x$ является линейной комбинацией выражений вида 4′), и, значит, оно равно нулю.

Выражение $x \otimes y$, рассматриваемое как элемент из $R \wedge R$, называют *внешним произведением векторов* x *и* y и обозначают так: $x \wedge y$. Равенство (7) означает, что внешнее произведение векторов антисимметрично: $x \wedge y = -y \wedge x$.

Покажем, что $S^2(R)$ и $R \wedge R$ можно определить и как подпространства в $R \otimes R$; точнее, в $R \otimes R$ имеются подпространства, естественным образом изоморфные $S^2(R)$ и $R \wedge R$.

Для этого зададим в пространстве $R \otimes R$ линейное преобразование σ, определяемое по формуле

$$\sigma(x_1 \otimes y_1 + \ldots + x_k \otimes y_k) = y_1 \otimes x_1 + \ldots + y_k \otimes x_k.$$

Очевидно, что его квадрат есть единичное преобразование: $\sigma^2 = 1$.

Рассмотрим два подпространства в $R \otimes R$ — подпространство H_1 элементов X, для которых $\sigma X = X$ и подпространство H_2 элементов X, для которых $\sigma X = -X$. Эти подпространства имеют нулевое пересечение, так как из условий $\sigma X = X$ и $\sigma X = -X$ следует, что $X = 0$. Покажем, что их прямая сумма есть все пространство $R \otimes R$. В самом деле, представим любой элемент X из $R \otimes R$ в виде суммы $X = X_1 + X_2$, где $X_1 = \frac{1}{2}(X + \sigma X)$ и $X_2 = \frac{1}{2}(X - \sigma X)$. Очевидно, что $\sigma X_1 = X_1$, т. е. $X_1 \in H_1$, и $\sigma X_2 = -X_2$, т. е. $X_2 \in H_2$.

Покажем теперь, что при естественном отображении $R \otimes R$ на $S^2(R)$ в нуль переходят все элементы из H_2, и притом только они. В самом деле, пусть $X \in R \otimes R$ переходит при этом отображении в нуль; тогда X равно линейной комбинации выражений вида 4), т. е. выражений $x \otimes y - y \otimes x$; следовательно, $\sigma X = -X$, т. е. $X \in H_2$. Обратно, пусть $X \in H_2$, т. е. $\sigma X = -X$; тогда $X = \frac{1}{2}(X - \sigma X)$; следовательно, X равно линейной комбинации выражений вида 4) и, значит, переходит в нуль при отображении $R \otimes R$ на $S^2(R)$.

Поскольку $R \otimes R$ является прямой суммой H_1 и H_2, то этим доказано, что при отображении $R \otimes R$ на $S^2(R)$ подпространство H_1 изоморфно отображается на $S^2(R)$.

Итак, мы установили изоморфизм между $S^2(R)$ и подпространством $H_1 \subset R \otimes R$ элементов X, для которых $\sigma X = X$.

Аналогично устанавливается изоморфизм между $R \wedge R$ и подпространством $H_2 \subset R \otimes R$ элементов X, для которых $\sigma X = -X$.

У п р а ж н е н и е. Пусть $e_1, \ldots, e_n$ — базис в R. Доказать, что элементы $e_i \wedge e_j$, где $i < j$, образуют базис в $R \wedge R$.

9. Внешняя степень $\overset{m}{\bigwedge} \boldsymbol{R}$. Теперь дадим определение внешней степени $\overset{m}{\bigwedge} R$ пространства R для произвольного m. Рассмотрим m-ю тензорную степень $\overset{m}{\bigotimes} R$. Напомним, что элементами пространства $\overset{m}{\bigotimes} R$ являются формальные суммы выражений вида

$$x_1 \otimes x_2 \otimes \ldots \otimes x_m, \tag{8}$$

где $x_i \in R$, причем некоторые из таких сумм считаются равными между собой. Приравняем дополнительно нулю все выражения вида (8), у которых совпадают хотя бы два сомножителя, а также любые их линейные комбинации. То линейное пространство, которое при этом получается, называется *внешней m-й степенью пространства* R и обозначается через $\overset{m}{\bigwedge} R$.

Выражение $x_1 \otimes x_2 \otimes \ldots \otimes x_m$, рассматриваемое как элемент из $\overset{m}{\bigwedge} R$, называется *внешним произведением векторов* $x_1, \ldots, x_m$ и обозначается $x_1 \wedge x_2 \wedge \ldots \wedge x_m$. Нетрудно убедиться (подобно тому, как это уже делалось для случая двух сомножителей), что *внешнее произведение векторов антисимметрично,* т. е. оно меняет знак при перестановке любых двух сомножителей.

Среди внешних степеней пространства R имеется лишь конечное число отличных от нуля. Именно, покажем, что $\overset{m}{\bigwedge} R = 0$ при $m > n$, где n — размерность R. Для этого зададим базис $e_1, \ldots, e_n$ в пространстве R. Разлагая векторы из R по элементам базиса, мы убеждаемся, что любое внешнее произведение $x_1 \wedge \ldots \wedge x_m$, а значит, и любой элемент из $\overset{m}{\bigwedge} R$, является линейной комбинацией выражений $e_{i_1} \wedge \ldots \wedge e_{i_m}$. Но если $m > n$, то в каждом выражении $e_{i_1} \wedge \ldots \wedge e_{i_m}$ совпадают хотя бы два сомножителя; значит, всегда $e_{i_1} \wedge \ldots \wedge e_{i_m} = 0$. Итак, $\overset{m}{\bigwedge} R = 0$ при $m > n$.

Покажем также, что *пространство $\overset{n}{\bigwedge} R$, где n — размерность R, является одномерным пространством.* В самом деле, среди элементов $e_{i_1} \wedge \ldots \wedge e_{i_n}$ отличны от нуля только те, у которых индексы $i_1, \ldots, i_n$ попарно различны и, значит, являются перестановками индексов $1, \ldots, n$. Так как внешнее произведение векторов антисимметрично, то такие отличные от нуля элементы совпадают, с точностью до знака, с элементом $e_1 \wedge \ldots \wedge e_n$. Поскольку любой элемент из $\overset{n}{\bigwedge} R$ является линейной комбинацией векторов $e_{i_1} \wedge \ldots \wedge e_{i_n}$, то тем самым он кратен вектору $e_1 \wedge \ldots \wedge e_n$.

Упражнения. 1. Пусть $e_1, \ldots, e_n$ — базис в R и $e'_i = \sum\limits_{j=1}^{n} a_{ji} e_j$ — любые n векторов из R. Доказать, что $e'_1 \wedge \ldots \wedge e'_n = a e_1 \wedge \ldots \wedge e_n$, где a — определитель матрицы $\|a_{ij}\|$.

2. Пусть $e_1, \ldots, e_n$ — базис в R. Доказать, что выражения $e_{i_1} \wedge \ldots \wedge e_{i_m}$, где $i_1 < i_2 < \ldots < i_m$, образуют базис в $\bigwedge\limits^{m} R$. На основании этого вычислить размерность пространства $\bigwedge\limits^{m} R$.

Задача. Дать (по аналогии со случаем $m = 2$) определение m-й симметрической степени $S^m(R)$ пространства R для любого m.

10. Тензорное произведение евклидовых пространств. Пусть R_1 — евклидово пространство со скалярным произведением $(x, x')_1$, R_2 — другое евклидово пространство со скалярным произведением $(y, y')_2$. Тогда в их тензорном произведении $R_1 \otimes R_2$ можно естественным образом ввести скалярное произведение.

Сначала определим его для пары векторов $x \otimes y$ и $x' \otimes y'$, полагая

$$(x \otimes y, x' \otimes y') = (x, x')_1 \cdot (y, y')_2.$$

Если теперь

$$X = x_1 \otimes y_1 + \ldots + x_k \otimes y_k,$$
$$X' = x'_1 \otimes y'_1 + \ldots + x'_l \otimes y'_l$$

— произвольные векторы из $R_1 \otimes R_2$, то положим:

$$(X, X') = \sum_{i=1}^{k} \sum_{j=1}^{l} (x_i \otimes y_i, x'_j \otimes y'_j). \tag{9}$$

Читателю предлагается убедиться, что выражение (9) действительно задает скалярное произведение на $R_1 \otimes R_2$. Именно, оно имеет смысл на $R_1 \otimes R_2$ (т. е. сохраняется при замене выражений X и X' на равные) и удовлетворяет всем аксиомам скалярного произведения.

Пространство $R_1 \otimes R_2$ с введенным в нем так скалярным произведением называется *тензорным произведением евклидовых пространств* R_1 и R_2.

Заметим, что если $e_1, \ldots, e_m$ — ортонормированный базис в R_1, а $f_1, \ldots, f_n$ — ортонормированный базис в R_2, то векторы $e_i \otimes f_j$ образуют ортонормированый базис в тензорном произведении $R_1 \otimes R_2$. В самом деле,

$$(e_i \otimes f_j, e_{i'} \otimes f_{j'}) = (e_i, e_{i'})_1 (f_j, f_{j'})_2.$$

Значит, это выражение равно 1 при $i = i'$, $j = j'$ и равно нулю во всех остальных случаях.

Д О Б А В Л Е Н И Е

ТЕОРИЯ ВОЗМУЩЕНИЙ

Точное вычисление собственных значений и собственных векторов самосопряженного линейного преобразования часто наталкивается на значительные вычислительные трудности. Одним из распространенных методов приближенного вычисления собственных значений в квантовой механике и во многих задачах теории колебаний является так называемый метод возмущений. Этот метод, применимый к линейным преобразованиям как в вещественном, так и в комплексном пространстве, грубо говоря, состоит в следующем: пусть известны собственные значения и собственные векторы некоторого самосопряженного линейного преобразования A. Рассмотрим преобразование $A + \varepsilon B$, где B — произвольное самосопряженное преобразование. Тогда собственные значения $A+\varepsilon B$ суть функции от ε. Можно показать, что при $\varepsilon \longrightarrow 0$ собственные значения и векторы $A+\varepsilon B$ стремятся к собственным значениям и векторам A. Задача состоит в нахождении «поправок» к собственным значениям и векторам при замене преобразования A на $A+\varepsilon B$.

§ 1. Случай некратных собственных значений

Пусть A имеет различные собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ и пусть $e_1, e_2, \ldots, e_n$ — соответствующие им нормированные собственные векторы. Пусть, далее, B — какое-либо другое самосопряженное линейное преобразование. Собственные значения преобразования $A+\varepsilon B$ обозначим через $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \ldots, \lambda_n(\varepsilon)$, а соответствующие собственные векторы — через $e_1(\varepsilon)$, $e_2(\varepsilon), \ldots, e_n(\varepsilon)$. Можно доказать, что $\lambda_k(\varepsilon)$ и $e_k(\varepsilon)$ являются непрерывными и дифференцируемыми функциями от ε, причем $\lambda_k(0)=\lambda_k$, а $e_k(0)=e_k$. Представим эти функции в виде

$$\lambda_k(\varepsilon)=\lambda_k+\varepsilon\lambda_k^{(1)}+\ldots$$

и

$$e_k(\varepsilon)=e_k+\varepsilon e_k^{(1)}+\ldots\ ^{*)}$$

и будем сначала искать $\lambda_k^{(1)}$ и $e_k^{(1)}$, т. е. «главную часть» поправки к $e_k=e_k(0)$ и $\lambda_k=\lambda_k(0)$. Мы имеем

$$(A+\varepsilon B)e_k(\varepsilon)=\lambda_k(\varepsilon)e_k(\varepsilon),$$

т. е.

$$(A+\varepsilon B)(e_k+\varepsilon e_k^{(1)}+\ldots)=(\lambda_k+\varepsilon\lambda_k^{(1)}+\ldots)(e_k+\varepsilon e_k^{(1)}+\ldots).$$

Сравним члены первой степени относительно ε в обеих частях равенства. Мы получим

$$Ae_k^{(1)}+Be_k=\lambda_k e_k^{(1)}+\lambda_k^{(1)}e_k. \tag{1}$$

Умножим обе части (1) скалярно на e_k:

$$(Ae_k^{(1)},e_k)+(Be_k,e_k)=\lambda_k(e_k^{(1)},e_k)+\lambda_k^{(1)}(e_k,e_k).$$

Так как, в силу самосопряженности преобразования A,

$$(Ae_k^{(1)},e_k)=(e_k^{(1)},Ae_k)=\lambda_k(e_k^{(1)},e_k),$$

$^{*)}$ Многоточие здесь и дальнейшем означает, что отброшено слагаемое порядка выше первого по сравнению с ε. Мы не пишем вместо многоточия $o(\varepsilon)$, чтобы не загромождать изложения.

то
$$(Be_k, e_k) = \lambda_k^{(1)}(e_k, e_k) = \lambda_k^{(1)}.$$
Отсюда
$$\lambda_k^{(1)} = (Be_k, e_k), \tag{2}$$
и первая половина нашей задачи таким образом решена.

Вычислим теперь главный член поправки к собственному вектору $e_k(\varepsilon)$, т. е. $e_k^{(1)}$. Для этого умножим скалярно обе части равенства (1) на e_i, где $i \neq k$. Так как векторы e_k и e_i ортогональны, т. е. $(e_k, e_i) = 0$ при $i \neq k$, то мы получим
$$(Ae_k^{(1)}, e_i) + (Be_k, e_i) = \lambda_k(e_k^{(1)}, e_i).$$
Но, аналогично предыдущему, мы имеем
$$(Ae_k^{(1)}, e_i) = (e_k^{(1)}, Ae_i) = \lambda_i(e_k^{(1)}, e_i),$$
поэтому
$$(e_k^{(1)}, e_i) = \frac{(Be_k, e_i)}{\lambda_k - \lambda_i}, \quad i \neq k. \tag{3}$$
Совокупность этих равенств и определяет вектор $e_k^{(1)}$.

Запишем формулы (2) и (3) в координатной форме. Для этого удобнее всего выбрать в качестве базиса собственные векторы $e_1, \ldots, e_k$ «невозмущенного» преобразования A. Матрицу преобразования B в этом базисе обозначим через b_{ij}, т. е. $Be_j = \sum b_{jk}e_k$ и, следовательно,
$$(Be_j, e_i) = b_{ij}.$$
Координаты вектора $e_k^{(1)}$ — главного члена «поправки» — обозначим через $\xi_1, \ldots, \xi_k, \ldots, \xi_n$, т. е.
$$e_k^{(1)} = \xi_1 e_1 + \ldots + \xi_n e_n \tag{4}$$
и, значит,
$$\xi_i = (e_k^{(1)}, e_i).$$

Формулы (2) и (3) приобретут вид

$$\lambda_k^{(1)} = b_{kk}, \tag{2$'$}$$

$$\xi_i = \frac{b_{ik}}{\lambda_k - \lambda_i}. \tag{3$'$}$$

Сам вектор $e_k^{(1)}$ определяется числами $\xi_i = (e_k^{(1)}, e_i)$ по формуле (4). У нас осталась неопределенной k-я координата ξ_k. Она определяется из условия нормировки собственного вектора, т. е. из условия, чтобы длина вектора $e_k + \varepsilon e_k^{(1)} + \ldots$ была равна единице. Мы имеем

$$(e_k + \varepsilon e_k^{(1)} + \ldots,\ e_k + \varepsilon e_k^{(1)} + \ldots) = 1,$$

т. е.

$$(e_k, e_k) + \varepsilon[(e_k^{(1)}, e_k) + (e_k, e_k^{(1)})] + \ldots = 1.$$

Сравнивая члены при первых степенях ε, имеем $(e_k^{(1)}, e_k) + (e_k, e_k^{(1)}) = 0$. Этому условию можно удовлетворить, полагая *)

$$\xi_k = (e_k^{(1)}, e_k) = 0. \tag{5}$$

Окончательно имеем

$$\lambda_k^{(1)} = b_{kk}, \tag{I}$$

$$e_k^{(1)} = \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{b_{ik}}{\lambda_k - \lambda_i} e_i, \tag{II}$$

где $b_{ik} = (Be_k, e_i)$, а λ_k — собственные значения «невозмущенного» преобразования A.

*) В комплексном случае $(e_k^{(1)}, e_k) + (e_k, e_k^{(1)}) = 2\,\mathrm{Re}(e_k^{(1)}, e_k)$, и мы могли бы считать $(e_k^{(1)}, e_k)$ не только нулем, но и произвольным чисто мнимым числом. Это связано с тем, что нормировка собственного вектора определяет его в комплексном случае с точностью до множителя, по модулю равного единице.

Для получения формул (I) и (II) мы выбрали базис, состоящий из собственных векторов преобразования A. При произвольном базисе формулы (2) и (3) также определяют $\lambda_k^{(1)}$ и $e_k^{(1)}$. Чтобы получить формулы, аналогичные (I) и (II) в произвольном ортогональном базисе, надо знать только координаты векторов e_k и матрицу преобразования B в этом базисе. Пусть матрица B есть $\|\beta_{\mu\nu}\|$, а $e_k^{(1)} = (c_1^{(k)}, \ldots, c_n^{(k)})$. Тогда из (2) получаем

$$\lambda_k = \sum_{\mu,\nu=1}^{n} \beta_{\mu\nu} c_\mu^{(k)} c_\nu^{(k)},$$

а из (3) получаем систему уравнений для определения координат $\zeta_1, \ldots, \zeta_n$ вектора $e_k^{(1)}$:

$$\zeta_1 c_1^{(i)} + \zeta_2 c_2^{(i)} + \ldots + \zeta_n c_n^{(i)} = \\ = \frac{\sum\limits_{\mu,\nu} \beta_{\mu\nu} c_\mu^{(i)} c_\nu^{(k)}}{\lambda_k - \lambda_i} \quad (i = 1, 2, \ldots, k-1, k+1, \ldots, n).$$

Недостающее уравнение снова получаем из условия (5) нормировки вектора $e_k(\varepsilon)$:

$$\zeta_1 c_1^{(k)} + \zeta_2 c_2^{(k)} + \ldots + \zeta_n c_n^{(k)} = 0.$$

Так как векторы $e_1, \ldots, e_n$ линейно независимы, то определитель этой системы отличен от нуля, и числа $\zeta_1, \zeta_2, \ldots, \zeta_n$ определяются из нее однозначно.

Найдем теперь собственные значения во втором приближении, т.е. с точностью до членов порядка ε^2.

Мы видели, что, для того чтобы найти собственное значение в первом приближении [формула (2)], достаточно было знать собственный вектор в нулевом приближении. Аналогично, для того чтобы найти второе приближение к собственному значению, нам достаточно будет знать собственные векторы в первом приближении. Мы имеем

$$(A + \varepsilon B) e_k(\varepsilon) = \lambda_k(\varepsilon) e_k(\varepsilon).$$

Подставим в это равенство:

$$e_k(\varepsilon) = e_k + \varepsilon e_k^{(1)} + \varepsilon^2 e_k^{(2)} + \dots,$$
$$\lambda_k(\varepsilon) = \lambda_k + \varepsilon \lambda_k^{(1)} + \varepsilon^2 \lambda_k^{(2)} + \dots$$

и сравним члены при ε^2. Получим

$$Be_k^{(1)} + Ae_k^{(2)} = \lambda_k^{(2)} e_k + \lambda_k^{(1)} e_k^{(1)} + \lambda_k e_k^{(2)}. \qquad (6)$$

Для того чтобы найти $\lambda_k^{(2)}$, умножим скалярно обе части этого равенства на e_k. Учитывая, что $(Ae_k^{(2)}, e_k) = (e_k^{(2)}, Ae_k) = \lambda_k (e_k^{(2)}, e_k)$, мы получим

$$(Be_k^{(1)}, e_k) = \lambda_k^{(2)} + \lambda_k^{(1)} (e_k^{(1)}, e_k).$$

Так как, в силу (5) $(e_k^{(1)}, e_k) = 0$, то

$$\lambda_k^{(2)} = (Be_k^{(1)}, e_k).$$

Но $e_k^{(1)} = \sum \xi_i e_i$. Подставляя, получаем в силу формулы (3) для первого приближения

$$\lambda_k^{(2)} = \sum_{i=1}^{n} \xi_i (Be_i, e_k) = \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{(Be_k, e_i)(Be_i, e_k)}{\lambda_k - \lambda_i},$$

а так как $(Be_k, e_i) = \overline{(Be_i, e_k)}$, то окончательно имеем

$$\lambda_k^{(2)} = \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{|(Be_k, e_i)|^2}{\lambda_k - \lambda_i},$$

где e_k — собственные векторы, а λ_k — собственные значения преобразования A, или

$$\lambda_k^{(2)} = \sum_{\substack{i=1 \\ i \neq k}}^{n} \frac{|b_{ik}|^2}{\lambda_k - \lambda_i}.$$

З а д а ч а. Найти поправки для собственного вектора во втором приближении, умножая скалярно обе части равенства (6) на e_i.

Ответ.

$$e_k^{(2)} = \sum_{i\neq k}\sum_{j\neq k} \frac{b_{ij}b_{jk}}{(\lambda_i-\lambda_k)(\lambda_j-\lambda_k)}e_i - \\ - \sum_{i\neq k}\frac{b_{kk}b_{ik}}{(\lambda_i-\lambda_k)^2}e_i - \sum_{i\neq k}\frac{|b_{ik}|^2}{(\lambda_i-\lambda_k)^2}\frac{e_k}{2}.$$

Кроме указанных уравнений надо воспользоваться также условием нормировки, т. е. равенством

$$(e_k + \varepsilon e_k^{(1)} + \varepsilon^2 e_k^{(2)} + \ldots,\ e_k + \varepsilon e_k^{(1)} + \varepsilon^2 e_k^{(2)} + \ldots) = 1.$$

§ 2. Случай кратных собственных значений

Рассмотрим теперь случай, когда λ есть r-кратное собственное значение преобразования A. Обозначим через

$$f_1, f_2, \ldots, f_r \tag{1}$$

какие-либо r попарно ортогональных собственных векторов преобразования A, отвечающих этому собственному значению λ. Заметим, что, так как e_i $(i = 1, 2, \ldots, r)$ отвечают одному и тому же собственному значению λ, то линейная комбинация этих векторов также будет собственным вектором, отвечающим этому собственному значению. Этими линейными комбинациями исчерпываются все собственные значения преобразования A, отвечающие собственному значению λ.

При замене преобразования A на $A + \varepsilon B$ собственное значение перестанет, вообще говоря, быть кратным, и вместо λ мы получим r различных собственных значений $\lambda_1(\varepsilon), \lambda_2(\varepsilon), \ldots, \lambda_r(\varepsilon)$. Соответствующие нормированные собственные векторы обозначим через $e_1(\varepsilon), e_2(\varepsilon), \ldots, e_r(\varepsilon)$.

Например, если A — единичное преобразование, т. е. $A = E$, в n-мерном пространстве, а B — произвольное самосопряженное преобразование с попарно различными собственными значениями $\lambda_1, \dots, \lambda_n$, то $A = E$ имеет n-кратное собственное значение 1, а $A+\varepsilon B = E+\varepsilon B$ имеет n различных собственных значений $1+\lambda_1\varepsilon$, $1 + \lambda_2\varepsilon$, …, $1 + \lambda_n\varepsilon$.

Аналогично случаю простых собственных значений $\lambda_i(\varepsilon)$ и $e_i(\varepsilon)$ являются непрерывными и дифференцируемыми функциями от ε. При $\varepsilon \longrightarrow 0$ $\lambda_i(\varepsilon)$ стремятся к собственному значению A, т.е. к λ_i. Мы имеем

$$\lambda_i(\varepsilon) = \lambda_i + \varepsilon\lambda_i^{(1)} + \varepsilon^2\lambda_i^{(2)} + \dots \tag{2}$$

Для собственных векторов $e_i(\varepsilon)$ преобразования $A+\varepsilon B$ мы имеем аналогичное равенство

$$e_i(\varepsilon) = e_i + \varepsilon e_i^{(1)} + \varepsilon^2 e_i^{(2)} + \dots, \tag{3}$$

$e_i = \lim\limits_{\varepsilon\to 0} e_i(\varepsilon)$, и значит, e_i есть собственный вектор преобразования A, отвечающий собственному значению λ. Следовательно, e_i есть какая-то, заранее нам неизвестная, линейная комбинация векторов (1). Таким образом, в отличие от случая некратных собственных значений, сами e_i также подлежат определению. Подставим в равенство $(A + \varepsilon B)e_i(\varepsilon) = \lambda_i(\varepsilon)e_i(\varepsilon)$ выражения (2) и (3). Сравнивая, как и в предыдущем случае, коэффициенты при ε, получаем

$$Be_i + Ae_i^{(1)} = \lambda e_i^{(1)} + \lambda_i^{(1)} e_i, \quad i = 1, \dots, r. \tag{4}$$

Здесь вектор e_i является, как было указано, линейной комбинацией собственных векторов $f_1, \dots, f_r$:

$$e_i = \eta_1 f_1 + \eta_2 f_2 + \dots + \eta_r f_r.$$

Наша цель — найти число $\lambda_i^{(1)}$ и вектор e_i, т.е. числа $\eta_1, \dots, \eta_r$.

Умножая обе части (4) скалярно на f_k, получим

$$(Be_i, f_k) + (Ae_i^{(1)}, f_k) = (\lambda e_i^{(1)}, f_k) + \lambda_i^{(1)}(e_i, f_k),$$

или, так как $(Ae_i^{(1)}, f_k) = (e_i^{(1)}, Af_k) = \lambda_k(e_i^{(1)}, f_k)$,

$$(Be_i, f_k) = \lambda_i^{(1)}(e_i, f_k).$$

Подставляя в левую часть этого равенства вместо e_i его выражение и замечая, что $(e_i, f_k) = \eta_k$, получим

$$\sum_{p=1}^{r}(Bf_p, f_k)\eta_p = \lambda_i^{(1)}\eta_k, \tag{5}$$

или

$$\sum_{p=1}^{r} b_{kp}\eta_p = \lambda_i^{(1)}\eta_k, \tag{I}$$

где

$$(Bf_p, f_k) = b_{kp}.$$

Итак, числа $\lambda_i^{(1)}$ являются собственными значениями матрицы $\|b_{kp}\|$, $k, p = 1, 2, \ldots, r$, т. е. определяются из уравнения

$$\mathrm{Det}\,\|b_{ik} - \lambda\sigma_{ik}\| = 0,$$

а вектор e_i определяется формулой

$$e_i = \eta_1 f_1 + \ldots + \eta_r f_r,$$

где числа η_i находятся из уравнений (I).

Аналогично можно было бы найти поправки к собственным векторам, т. е. $e_k^{(1)}, e_k^{(2)}$, и следующие поправки к собственным значениям, т. е. $\lambda_i^{(2)}$.

Израиль Моисеевич Гельфанд

ЛЕКЦИИ ПО ЛИНЕЙНОЙ АЛГЕБРЕ

пятое издание

Редактор *В. В. Ященко*
Технический редактор *В. Ю. Радионов*
Художник *Ю. Винецкий*
Оригинал-макет подготовлен МЦНМО
Набор *Ю. Е. Галямина*

ООО „Добросвет“
Редакция: тел. (095) 238–23–33
ЛР № 064934
Подписано в печать 22.09.98
Формат 84 × 108/32. Бумага офсетная № 1
Печать офсетная. Объем 10 печ. л
Тираж 5000 экз. Заказ №
Цена договорная
По вопросам закупок обращаться в ТОО „ЧеРо“.
Отдел реализации: тел. (095) 939–34–39,
(095) 939–47–09

Отпечатано с готовых диапозитивов
Самарский дом печати. Просп. Карла Маркса, 201