ные 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:

$$A^2 = A \cdot A, \quad A^3 = A^2 \cdot A, \quad \ldots \text{ и т. д.}$$

Как и для чисел, полагаем, по определению, $A^0 = E$. Очевидно, что

$$A^{m+n} = A^m \cdot A^n.$$

Пример. 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 $D^3P(t)=P'''(t)$, ... Заметим, что в данном случае $D^n=0$. Действительно, так как векторами пространства являются многочлены степени $\leqslant n-1$, то

$$D^n P(t)=P^{(n)}(t)=0.$$

Упражнение. 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, ...

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 $P(t) = a_0 t^m + a_1 t^{m-1} + \ldots + a_m$ — произвольный многочлен. Тогда под $P(A)$ мы понимаем линейное преобразование, определенное формулой

$$P(A) = a_0 A^m + a_1 A^{m-1} + \ldots + a_m E.$$

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 $P(t) = a_0 t^m + a_1 t^{m-1} + \ldots + a_m$. Тогда $P(D)$ есть линейное преобразование, которое переводит функцию $f(t)$ в

$$P(D) f(t) = a_0 f^{(m)}(t) + a_1 f^{(m-1)}(t) + \ldots + a_m f(t).$$

В соответствии с введенными выше определениями сложения и умножения матриц многочлен от матрицы A мы задаем формулой

$$P(A) = a_0 A^m + a_1 A^{m-1} + \ldots + a_m E.$$

Пример. Предположим, что A — так называемая диагональная матрица, т. е. матрица, у которой на всех местах, кроме главной диагонали, стоят нули. Найдем $P(A)$. Мы имеем

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix};$$

тогда

$$A^2 = \begin{pmatrix} \lambda_1^2 & 0 & \ldots & 0 \\ 0 & \lambda_2^2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n^2 \end{pmatrix}, \ \ldots, \ A^m = \begin{pmatrix} \lambda_1^m & 0 & \ldots & 0 \\ 0 & \lambda_2^m & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n^m \end{pmatrix}.$$

Отсюда следует, что если $P(t) = a_0 t^m + \ldots + a_{m-1} t + a_m$, то

$$P(A) = \begin{pmatrix} P(\lambda_1) & 0 & \ldots & 0 \\ 0 & P(\lambda_2) & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & P(\lambda_n) \end{pmatrix}.$$

Упражнение. Пусть матрица A имеет вид

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Найти $P(A)$.

Можно определить не только многочлен от матрицы, но и вообще функцию от матрицы, например, e^A, $\sin A$ 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, ..., 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 1п. 1), очевидно, не имеет обратного.

С понятием обратного преобразования связано понятие обратной матрицы. Как известно, для каждой матрицы $\boldsymbol{A}$, удовлетворяющей условию $\operatorname{Det}(\boldsymbol{A})\neq 0$, 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, экви-

валентную матричному равенству (9). Элементы ее k-го столбца окажутся равными минорам k-й строки матрицы A, деленным на ее определитель. Легко проверить, что так составленная матрица 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 бы 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.

Пример. Рассмотрим преобразование A, состоящее в проектировании трехмерного пространства R в плоскость XY (пример 1, п. 2). Очевидно, что образ этого преобразования есть плоскость XY.

Упражнение. 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.

Определение 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).

Упражнение. Написать матрицу линейного преобразования A в базисе, первые k векторов которого есть базис ядра.

Пример. 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

$$A^2P(t)=P''(t).$$

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 $f_1, f_2, \ldots, f_n$. 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

$$\left.\begin{aligned} f_1 &= c_{11}e_1 + c_{21}e_2 + \ldots + c_{n1}e_n, \\ f_2 &= c_{12}e_1 + c_{22}e_2 + \ldots + c_{n2}e_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ f_n &= c_{1n}e_1 + c_{2n}e_2 + \ldots + c_{nn}e_n. \end{aligned}\right\} \tag{10}$$

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 $f_1, f_2, \ldots, f_n$ через $\boldsymbol{B} = \|b_{ik}\|$. Иначе говоря,

$$Ae_k = \sum_{i=1}^{n} e_{ik}e_i, \tag{10'}$$

$$Af_k = \sum_{i=1}^{n} b_{ik}f_i. \tag{10''}$$

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 *подобными*.

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.

Задача. 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 $x \in R_1$, $Ax \overline{\in} R_1$ *). Например, если R — евклидова плоскость, R_1 — произвольная прямая и A — поворот на угол $\varphi = \frac{\pi}{6}$, то очевидно, что для любого $x \neq 0$ и принадлежащего R_1, $Ax \overline{\in} R_1$. Однако может случиться, 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 пространство).

2. R — плоскость. Преобразование A заключается в растяжении плоскости в λ_1 раз вдоль оси X и в λ_2 раз вдоль

*) Запись $Ax \overline{\in} R_1$ означает, что вектор Ax не принадлежит подпространству R_1.

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

$$\begin{pmatrix} a_{11} & \ldots & a_{1k} & a_{1,\,k+1} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & \ldots & a_{kk} & a_{k,\,k+1} & \ldots & a_{kn} \\ 0 & \ldots & 0 & a_{k+1,\,k+1} & \ldots & a_{k+1,\,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & a_{n,\,k+1} & \ldots & a_{nn} \end{pmatrix}.$$

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. 23). 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.

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1k} & 0 & \ldots & 0 \\ a_{21} & \ldots & a_{2k} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & \ldots & a_{kk} & 0 & \ldots & 0 \\ 0 & \ldots & 0 & a_{k+1,\,k+1} & \ldots & a_{k+1,\,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & a_{n,\,k+1} & \ldots & a_{n,\,n} \end{pmatrix}.$$

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 инвариантно, 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):

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

Условие того, что вектор собственный, т. е. равенство

$$Ax = \lambda x,$$

записывается в следующем виде:

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

или

$$\left.\begin{aligned} (a_{11}-\lambda)\xi_1 + a_{12}\xi_2 + \ldots + a_{1n}\xi_n &= 0, \\ a_{21}\xi_1 + (a_{22}-\lambda)\xi_2 + \ldots + a_{2n}\xi_n &= 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \ldots + (a_{nn}-\lambda)\xi_n &= 0. \end{aligned}\right\} \qquad (1)$$

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).

---

*) Доказательство теоремы пригодно для пространства над любым алгебраическим замкнутым полем, так как использует лишь существование решения у уравнения (2).

Условием существования ненулевого решения однородной системы (1) является равенство нулю ее определителя

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0. \qquad (2)$$

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 $\xi_1^{(0)}, \xi_2^{(0)}, \dots, \xi_n^{(0)}$. Тогда вектор

$$x^{(0)} = \xi_1^{(0)} e_1 + \xi_2^{(0)} e_2 + \dots + \xi_n^{(0)} e_n$$

будет собственным вектором, а λ_0 — собственным значением, так как

$$Ax^{(0)} = \lambda_0 x^{(0)}.$$

Теорема доказана.

Замечание. 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.

$$Ae_i = \lambda_i e_i \quad (i = 1, 2, \ldots, n).$$

Примем $e_1, e_2, \ldots, e_n$ за базис в R. Равенства

$$\begin{aligned} Ae_1 &= \lambda_1 e_1, \\ Ae_2 &= \lambda_2 e_2, \\ &\cdots \\ Ae_n &= \lambda_n e_n \end{aligned}$$

означают, что матрица преобразования A в этом базисе имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

(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.*

---

*) 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.

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

$$\begin{pmatrix} \lambda_0 & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda_0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda_0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 & \lambda_0 \end{pmatrix}.$$

2. Найти характеристический многочлен матрицы

$$\begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_{n-1} & a_n \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Ответ. $(-1)^n(\lambda^n - a_1\lambda^{n-1} - a_2\lambda^{n-2} - \dots - a_n)$.

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:

$$Q(\lambda) = \begin{vmatrix} a_{11} - \lambda b_{11} & a_{12} - \lambda b_{12} & \dots & a_{1n} - \lambda b_{1n} \\ a_{21} - \lambda b_{21} & a_{22} - \lambda b_{22} & \dots & a_{2n} - \lambda b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} - \lambda b_{n1} & a_{n2} - \lambda b_{n2} & \dots & a_{nn} - \lambda b_{nn} \end{vmatrix}.$$

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

$$q_0 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{4}$$

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 $|A-\lambda E|$. 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* 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* A.

Числа $p_1, p_2, \ldots, p_n$, построенные по матрице 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 A и p_1 — сумма диагональных элементов матрицы A. *Сумма диагональных элементов называется следом матрицы* A. След матрицы A обозначается $\operatorname{tr} A$ (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 A и B — матрицы n-го порядка, то

$$\operatorname{tr} AB=\operatorname{tr} BA.$$

2. Показать, что если C — невырожденная матрица n-го порядка, то для любой матрицы A n-го порядка имеем:

$$\operatorname{tr}(C^{-1}AC)=\operatorname{tr} A.$$

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

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{22} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}, \tag{5}$$

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 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 A связаны соотношением

$$P(\lambda)\, E = (A - \lambda E)\, C(\lambda), \tag{6}$$

где $C(\lambda)$ — многочлен от λ, коэффициенты которого являются матрицами, т. е.

$$C(\lambda) = C_0\lambda^{m-1} + C_1\lambda^{m-2} + \dots + C_{m-1}, \quad C_i \text{ — матрицы.}$$

Тогда $P(A) = 0$.

Заметим, что эта лемма является обобщением на многочлены с матричными коэффициентами теоремы Безу.

Доказательство. Мы имеем:

$$(A - \lambda E)\, C(\lambda) = AC_{m-1} + (AC_{m-2} - C_{m-1})\lambda + \\ + (AC_{m-3} - C_{m-2})\lambda^2 + \dots - C_0\lambda^m. \tag{7}$$

Сравнивая между собой коэффициенты при одинаковых степенях λ в обеих частях равенства (6), мы получаем последовательность

равенств:

$$\begin{aligned} AC_{m-1} \qquad\qquad &= a_m E, \\ AC_{m-2} - C_{m-1} &= a_{m-1} E, \\ AC_{m-3} - C_{m-2} &= a_{m-2} E, \\ \cdots\cdots\cdots\cdots & \\ AC_0 \quad - C_1 \quad &= a_1 E, \\ - C_0 \quad &= a_0 E. \end{aligned} \tag{8}$$

Умножим теперь слева первое равенство на E, второе на A, третье на A^2, ..., последнее на A^m и сложим их. Мы получим справа $P(A) = a_m E + a_{m-1} A + \ldots + a_0 A^m$, а слева 0. Таким образом, $P(A) = 0$, и лемма доказана *).

Теорема. *Если $P(\lambda)$ — характеристический многочлен матрицы A, то $P(A) = 0$.*

Доказательство. Рассмотрим матрицу, обратную матрице $A - \lambda E$. Мы имеем $(A - \lambda E)(A - \lambda E)^{-1} = E$. Как известно, обратная матрица может быть записана в виде

$$(A - \lambda E)^{-1} = \frac{1}{P(\lambda)} C(\lambda),$$

где $C(\lambda)$ — матрица из миноров $(n-1)$-го порядка матрицы $A - \lambda E$, а $P(\lambda)$ — определитель матрицы $A - \lambda E$, т. е. характеристический многочлен матрицы A. Отсюда

$$(A - \lambda E) C(\lambda) = P(\lambda) E.$$

Так как элементами матрицы $C(\lambda)$ являются миноры матрицы $A - \lambda E$, т. е. многочлены степени не выше $n-1$ относительно λ, то согласно доказанной лемме

$$P(A) = 0,$$

и теорема доказана.

Заметим, что если у характеристического многочлена матрицы A нет кратных корней, то не существует многочлена степени ниже n, обращающегося в нуль при подстановке в него матрицы A (см. следующее упражнение).

Упражнение. Пусть A — диагональная матрица вида

$$A = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

где все числа λ_i различны. Найти многочлен $P(t)$ возможно более низкой степени, для которого $P(A) = 0$. (См. пример в § 9, п. 3.)

---

*) В алгебре теорема Безу доказывается тем, что в равенство (6) подставляется A вместо λ. Здесь, конечно, мы не вправе этого сделать непосредственно, так как λ есть число, а 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 A^k и складывая затем все равенства, мы, по существу, подставляем A вместо λ.

§ 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 $A(x;\ y)$, задаваемая формулой

$$A(x;\ y) \equiv (Ax,\ y).$$

Действительно, функция $A(x;\ y) \equiv (Ax,\ y)$ удовлетворяет условиям, определяющим билинейную форму. Имеем:

$$1^\circ\ (A(x_1+x_2),\ y) = (Ax_1+Ax_2,\ y) = (Ax_1,\ y) + (Ax_2,\ y),$$
$$(A\lambda x,\ y) = (\lambda Ax,\ y) = \lambda(Ax,\ y).$$

$$2^\circ\ (x,\ A(y_1+y_2)) = (x,\ Ay_1+Ay_2) = (x,\ Ay_1) + (x,\ Ay_2),$$
$$(x,\ A\mu y) = (x,\ \mu Ay) = \bar{\mu}(x,\ Ay).$$

Покажем, что преобразование A определяется соответствующей билинейной формой $A(x;\ y)$ однозначно.

---

*) 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 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.

Пусть

$$A(x;\ y)=(Ax,\ y)$$

и

$$A(x;\ y)=(Bx,\ y).$$

Тогда

$$(Ax,\ y)\equiv(Bx,\ y),$$

т. е.

$$(Ax-Bx,\ y)=0$$

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

$$x=\xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n \text{ и } y=\eta_1e_1+\eta_2e_2+\ldots+\eta_ne_n,$$

то $A(x;\ y)$ можно записать в виде

$$\begin{aligned}A(x;\ y)=a_{11}\xi_1\bar{\eta}_1+a_{12}\xi_1\bar{\eta}_2+\ldots+a_{1n}\xi_1\bar{\eta}_n+\\ +a_{21}\xi_2\bar{\eta}_1+a_{22}\xi_2\bar{\eta}_2+\ldots+a_{2n}\xi_2\bar{\eta}_n+\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ +a_{n1}\xi_n\bar{\eta}_1+a_{n2}\xi_n\bar{\eta}_2+\ldots+a_{nn}\xi_n\bar{\eta}_n.\end{aligned}\tag{1}$$

Постараемся представить это выражение в виде некоторого скалярного произведения. Для этого перепишем его следующим образом:

$$\begin{aligned}A(x;\ y)=(a_{11}\xi_1+a_{21}\xi_2+\ldots+a_{n1}\xi_n)\,\bar{\eta}_1+\\ +(a_{12}\xi_1+a_{22}\xi_2+\ldots+a_{n2}\xi_n)\,\bar{\eta}_2+\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ +(a_{1n}\xi_1+a_{2n}\xi_2+\ldots+a_{nn}\xi_n)\,\bar{\eta}_n.\end{aligned}$$

Введем в рассмотрение вектор z с координатами

$$\begin{aligned}\zeta_1&=a_{11}\xi_1+a_{21}\xi_2+\ldots+a_{n1}\xi_n,\\ \zeta_2&=a_{12}\xi_1+a_{22}\xi_2+\ldots+a_{n2}\xi_n,\\ &\cdots\cdots\cdots\cdots\\ \zeta_n&=a_{1n}\xi_1+a_{2n}\xi_2+\ldots+a_{nn}\xi_n.\end{aligned}$$

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

$$A(x;\ y) = \zeta_1\bar{\eta}_1 + \zeta_2\bar{\eta}_2 + \ldots + \zeta_n\bar{\eta}_n = (z,\ y) = (Ax,\ y).$$

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.

Теорема 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)

$$\begin{aligned} A(x;\ y) = {} & a_{11}\xi_1\bar{\eta}_1 + a_{12}\xi_1\bar{\eta}_2 + \ldots + a_{1n}\xi_1\bar{\eta}_n + \\ & + a_{21}\xi_2\bar{\eta}_1 + a_{22}\xi_2\bar{\eta}_2 + \ldots + a_{2n}\xi_2\bar{\eta}_n + \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ & + a_{n1}\xi_n\bar{\eta}_1 + a_{n2}\xi_n\bar{\eta}_2 + \ldots + a_{nn}\xi_n\bar{\eta}_n \end{aligned}$$

мы будем выносить за скобки координаты $\xi_1, \xi_2, \ldots, \xi_n$ вектора x. Повторяя снова прежние рассуждения, мы

получаем:

$$\begin{aligned}
A(x;\ y) = {} & \xi_1(a_{11}\bar{\eta}_1 + a_{12}\bar{\eta}_2 + \ldots + a_{1n}\bar{\eta}_n) + \\
& + \xi_2(a_{21}\bar{\eta}_1 + a_{22}\bar{\eta}_2 + \ldots + a_{2n}\bar{\eta}_n) + \\
& \ldots\ldots\ldots\ldots\ldots\ldots \\
& + \xi_n(a_{n1}\bar{\eta}_1 + a_{n2}\bar{\eta}_2 + \ldots + a_{nn}\bar{\eta}_n) = \\
& = \xi_1\overline{(\bar{a}_{11}\eta_1 + \bar{a}_{12}\eta_2 + \ldots + \bar{a}_{1n}\eta_n)} + \\
& + \xi_2\overline{(\bar{a}_{21}\eta_1 + \bar{a}_{22}\eta_2 + \ldots + \bar{a}_{2n}\eta_n)} + \\
& \ldots\ldots\ldots\ldots\ldots\ldots \\
& + \xi_n\overline{(\bar{a}_{n1}\eta_1 + \bar{a}_{n2}\eta_2 + \ldots + \bar{a}_{nn}\eta_n)} = (x,\ A^*y).
\end{aligned}$$

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° $(\lambda A)^* = \bar{\lambda} A^*$.

5° $E^* = E$.

Докажем, например, первые два 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 равенства 3—5.

2. Доказать равенства 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 $\bar{\alpha}$. 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 $\bar{\alpha}=\alpha$. 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 всех 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

$$A^*=A \text{ и } B^*=B.$$

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 $z\bar{z}=1$, 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.

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, $\bar{\alpha}\alpha$ всегда равно $\alpha\bar{\alpha}$.

Нетрудно убедиться, что как самосопряженные, так и унитарные преобразования являются частными случаями нормальных преобразований.

---

*) В n-мерном пространстве условия $U^*U=E$ и $UU^*=E$ эквивалентны. В бесконечномерном пространстве это — два различных условия.

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:

$$\lambda (x, x) = \bar{\lambda} (x, x),$$

и так как $(x, x) \neq 0$, то $\lambda = \bar{\lambda}$, 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 A 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

$$\begin{aligned} Ae_1 &= \lambda_1 e_1, \\ Ae_2 &= \lambda_2 e_2, \\ &\cdots\cdots \\ Ae_n &= \lambda_n e_n, \end{aligned}$$

т. е. в этом базисе матрица преобразования A имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \tag{1}$$

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 заменою 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

$$(e_1, e_2) = 0.$$

Замечание. 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 $|\lambda_l|$ раз и, кроме того, зеркальное отражение в плоскости, ортогональной к данному направлению, если соответствующее λ_l отрицательно.

Параллельно с понятием самосопряженного преобразования вводится понятие эрмитовой матрицы.

Матрица $\|a_{ik}\|$ называется *эрмитовой*, если $a_{ik}=\bar{a}_{ki}$.

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

$$Ae_1=\lambda_1 e_1,\quad Ae_2=\lambda_2 e_2,\ \ldots,\ Ae_n=\lambda_n e_n.$$

Пусть

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

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.

Теорема 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.*

Доказательство. 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

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

В нормированном ортогональном базисе скалярное произведение имеет вид

$$(x, x) = |\xi_1|^2 + |\xi_2|^2 + \ldots + |\xi_n|^2,$$

т. е.

$$B(x; x) = |\xi_1|^2 + |\xi_2|^2 + \ldots + |\xi_n|^2. \tag{3}$$

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 $\lambda_1, \lambda_2, \ldots, \lambda_n$.

В каноническом виде матрицы квадратичных форм A и B имеют вид

$$\boldsymbol{A} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Следовательно,

$$\operatorname{Det}(\boldsymbol{A} - \lambda \boldsymbol{B}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda). \tag{4}$$

При переходе к другому базису матрицы эрмитовых квадратичных форм A и B переходят в $\boldsymbol{A}_1 = \boldsymbol{C}^* \boldsymbol{A} \boldsymbol{C}$ и

*) Относительно введенного нами скалярного произведения $(x, y) = B(x; y)$.

$B_1 = C^*BC$. Поэтому, если $e_1, e_2, \ldots, e_n$ — произвольный базис, то в этом базисе

$$\operatorname{Det}(A_1 - \lambda B_1) = \operatorname{Det} C^* \cdot \operatorname{Det}(A - \lambda B) \cdot \operatorname{Det} C,$$

т. е. отличается лишь постоянным множителем от выражения (4). Отсюда следует, что *числа* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *являются корнями следующего уравнения:*

$$\begin{vmatrix} a_{11} - \lambda b_{11} & a_{12} - \lambda b_{12} & \ldots & a_{1n} - \lambda b_{1n} \\ a_{21} - \lambda b_{21} & a_{22} - \lambda b_{22} & \ldots & a_{2n} - \lambda b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} - \lambda b_{n1} & a_{n2} - \lambda b_{n2} & \ldots & a_{nn} - \lambda b_{nn} \end{vmatrix} = 0,$$

где $\|a_{ik}\|$ и $\|b_{ik}\|$ — матрицы форм $A(x;\ x)$ и $B(x;\ x)$ в каком-нибудь базисе $e_1, e_2, \ldots, e_n$.

Замечание. Требование положительной определенности одной из форм является существенным, о чем свидетельствует следующий пример: две квадратичные формы

$$A(x;\ x) = |\xi_1|^2 - |\xi_2|^2, \quad B(x;\ x) = \xi_1\bar{\xi}_2 + \xi_2\bar{\xi}_1,$$

из которых ни одна не является положительно определенной, не могут быть одновременно приведены к сумме квадратов. В самом деле, первой форме соответствует матрица

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

а второй — матрица

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Рассмотрим матрицу $A - \lambda B$, 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 $e_1,\ e_2,\ \ldots,\ e_n$. Пусть в этом базисе преобразованию U соответствует матрица

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}. \tag{2}$$

Тогда сопряженному преобразованию U^* соответствует матрица

$$\begin{pmatrix} \bar{a}_{11} & \bar{a}_{21} & \dots & \bar{a}_{n1} \\ \bar{a}_{12} & \bar{a}_{22} & \dots & \bar{a}_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ \bar{a}_{1n} & \bar{a}_{2n} & \dots & \bar{a}_{nn} \end{pmatrix}. \tag{3}$$

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:

$$\sum_{\alpha=1}^{n} a_{i\alpha}\bar{a}_{i\alpha} = 1, \quad \sum_{\alpha=1}^{n} a_{i\alpha}\bar{a}_{k\alpha} = 0 \quad (i \neq k). \tag{4}$$

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:

$$\sum_{\alpha=1}^{n} a_{\alpha i}\bar{a}_{\alpha i} = 1; \quad \sum_{\alpha=1}^{n} a_{\alpha i}\bar{a}_{\alpha k} = 0 \quad (i \neq k). \tag{5}$$

Это условие аналогично предыдущему, но вместо строк в нем участвуют столбцы матрицы.

Условие (5) имеет простой геометрический смысл. Действительно, скалярное произведение векторов

$$Ue_i = a_{1i}e_1 + a_{2i}e_2 + \dots + a_{ni}e_n$$

и

$$Ue_k = a_{1k}e_1 + a_{2k}e_2 + \dots + a_{nk}e_n$$

равно $\sum a_{\alpha i}\bar{a}_{\alpha k}$ (так как $e_1, e_2, \dots, e_n$ — 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 U 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 базиса 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

$$(x, x) = (Ux, Ux) = (\lambda x, \lambda x) = \lambda\bar{\lambda}(x, x),$$

т. е. $\lambda\bar{\lambda} = 1$, 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 $\bar{\lambda}(Ux, e) = 0$. 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 R. 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 также 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:*

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \tag{7}$$

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,

$$\begin{aligned} Ue_1 &= \lambda_1 e_1, \\ Ue_2 &= \lambda_2 e_2, \\ &\cdot\ \cdot\ \cdot\ \cdot \\ Ue_n &= \lambda_n e_n \end{aligned}$$

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 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

$$x \in R_\lambda, \text{ т. е. } Ax = \lambda x,$$

то и

$$Bx \in R_\lambda, \text{ т. е. } ABx = \lambda Bx.$$

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 одновремен-*

но 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

$$Ae_1 = \lambda_1 e_1, \qquad Be_1 = \mu_1 e_1.$$

$(n-1)$-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:

$$Ae_2 = \lambda_2 e_2, \qquad Be_2 = \mu_2 e_2.$$

Совокупность векторов из R_1, ортогональных к e_2, образует $(n-2)$-мерное подпространство, 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:

$$Ae_i = \lambda_i e_i, \qquad Be_i = \mu_i e_i \qquad (i = 1, \ldots, n).$$

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 перестановочных 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, ..., то все доказано. Предположим поэтому, что хотя бы один вектор не является собственным для какого-либо из наших преобразований, например для A.

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, ... (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, ... 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.

Теорема 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.*)

Доказательство. Необходимость. Пусть в некотором ортогональном нормированном базисе матрица преобразования A диагональна, т. е. имеет вид

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Так как базис ортогональный и нормированный, то матрица преобразования A^* имеет вид

$$\begin{pmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ 0 & \bar{\lambda}_2 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \bar{\lambda}_n \end{pmatrix}.$$

*) Это означает, что каждое из преобразований A, B, C, ... кратно единичному преобразованию.

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. $Ax \in R_1$. 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 перестановочны. 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

$$AA^* = A^*A = A^2,$$

*) Упражнение: доказать, что $\mu_1 = \overline{\lambda_1}$.

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*

$$A = HU \quad (\text{либо } A = U_1H_1),$$

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 определенно. 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 $\|\bar{a}_{ki}\|$ 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

$$(Be,\ e) = \lambda (e,\ e)$$

итак 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

$$\begin{aligned}(Bx, x) &= \\ &= (\xi_1 Be_1 + \xi_2 Be_2 + \ldots + \xi_n Be_n,\ \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) = \\ &= (\xi_1 \lambda_1 e_1 + \xi_2 \lambda_2 e_2 + \ldots + \xi_n \lambda_n e_n,\ \xi_1 e_1 + \xi_2 e_2 + \ldots + \xi_n e_n) = \\ &= \lambda_1 |\xi_1|^2 + \lambda_2 |\xi_2|^2 + \ldots + \lambda_n |\xi_n|^2 \qquad (1)\end{aligned}$$

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:

$$B = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$ — собственные значения B. Согласно лемме 2 все $\lambda_i \geqslant 0$. Положим

$$H = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \ldots & 0 \\ 0 & \sqrt{\lambda_2} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \sqrt{\lambda_n} \end{pmatrix},$$

где числа $\sqrt{\lambda_i}$ выбираются неотрицательными. В силу той же леммы 2 преобразование H положительно опре-