

В матрицах H_α и G_β элементы первой наддиагонали равны единице, а все остальные элементы равны нулю. Учитывая эту специфичную структуру матриц H_α и G_β и полагая

$$[X_{\alpha\beta} = \|\xi_{ik}\|] \quad (i=1, 2, \dots, p_\alpha; k=1, 2, \dots, q_\beta),$$

мы заменим матричное уравнение (12) следующей эквивалентной ему системой скалярных соотношений¹⁾:

$$\xi_{i+1, k} = \xi_{i, k-1} (\xi_{i0} = \xi_{p_\alpha+1, k} = 0; i=1, 2, \dots, p_\alpha; k=1, 2, \dots, q_\beta). \quad (13)$$

Равенства (13) означают:

1) В матрице $X_{\alpha\beta}$ на каждой линии, параллельной главной диагонали, стоят равные между собой элементы,

$$2) \xi_{21} = \xi_{31} = \dots = \xi_{p_\alpha 1} = \xi_{p_\alpha 2} = \dots = \xi_{p_\alpha, q_\beta-1} = 0.$$

Пусть $p_\alpha = q_\beta$. В этом случае $X_{\alpha\beta}$ — квадратная матрица. Из 1), 2) следует, что в матрице $X_{\alpha\beta}$ все элементы, расположенные под главной диагональю, равны нулю, все элементы главной диагонали равны некоторому числу $c_{\alpha\beta}$, все элементы первой наддиагонали равны некоторому числу $c'_{\alpha\beta}$ и т. д., т. е.

$$X_{\alpha\beta} = \begin{vmatrix} c_{\alpha\beta} & c'_{\alpha\beta} & \dots & c^{(p_\alpha-1)}_{\alpha\beta} \\ 0 & c_{\alpha\beta} & & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & c'_{\alpha\beta} \\ 0 & \dots & 0 & c_{\alpha\beta} \end{vmatrix} = T_{p_\alpha}; \quad (14)$$

$$(p_\alpha = q_\beta)$$

здесь $c_{\alpha\beta}, c'_{\alpha\beta}, \dots, c^{(p_\alpha-1)}_{\alpha\beta}$ — произвольные параметры (уравнения (12) не накладывают никаких ограничений на значения этих параметров).

Легко видеть, что при $p_\alpha < q_\beta$

$$X_{\alpha\beta} = \begin{pmatrix} & & & q_\beta - p_\alpha \\ 0 & & & \\ & & & \\ & & & \end{pmatrix}, \quad T_{p_\alpha}, \quad (15)$$

а при $p_\alpha > q_\beta$

$$X_{\alpha\beta} = \begin{pmatrix} T_{\alpha\beta} \\ 0 \end{pmatrix}_{\{p_\alpha - q_\beta\}}. \quad (16)$$

Про матрицы (14), (15) и (16) мы будем говорить, что они имеют *правильную верхнюю треугольную форму*. Число произвольных параметров в $X_{\alpha\beta}$ равно наименьшему из чисел p_α и q_β . Приведенная ниже схема показывает структуру матрицы $X_{\alpha\beta}$ при $\lambda_\alpha = \mu_\beta$ (произвольные параметры

¹⁾ Из структуры матриц H_α и G_β следует, что произведение $H_\alpha X_{\alpha\beta}$ получается из $X_{\alpha\beta}$ сдвигом всех строк на одно место вверх и заполнением последней строки нулями; аналогично $X_{\alpha\beta} G_\beta$ получается из $X_{\alpha\beta}$ сдвигом всех столбцов на одно место вправо и заполнением первого столбца нулями (см. гл. I, стр. 25).

Для сокращения обозначений мы не пишем при ξ_{ik} дополнительных индексов α, β .

здесь обозначены через a, b, c, d :

$$X_{\alpha\beta} = \begin{vmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{vmatrix}, \quad X_{\alpha\beta} = \begin{vmatrix} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{vmatrix}, \quad X_{\alpha\beta} = \begin{vmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}.$$

$$(p_\alpha = q_\beta = 4)$$

$$(p_\alpha = 3, q_\beta = 5)$$

$$(p_\alpha = 5, q_\beta = 3)$$

Для того чтобы при подсчете произвольных параметров в матрице \tilde{X} охватить и случай 1, обозначим через $d_{\alpha\beta}(\lambda)$ наибольший общий делитель элементарных делителей $(\lambda - \lambda_\alpha)^{p_\alpha}$ и $(\lambda - \mu_\beta)^{q_\beta}$, а через $\delta_{\alpha\beta}$ — степень многочлена $d_{\alpha\beta}(\lambda)$ ($\alpha = 1, 2, \dots, u; \beta = 1, 2, \dots, v$). В случае 1 $\delta_{\alpha\beta} = 0$; в случае 2 имеем: $\delta_{\alpha\beta} = \min(p_\alpha, q_\beta)$. Таким образом, в обоих случаях число произвольных параметров в $X_{\alpha\beta}$ равно $\delta_{\alpha\beta}$. Число произвольных параметров в \tilde{X} определяется формулой

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta}.$$

В дальнейшем нам удобно будет общее решение уравнения (6) обозначить через $X_{\tilde{A}\tilde{B}}$ (до сих пор мы это решение обозначали буквой \tilde{X}).

Полученные в этом параграфе результаты можно сформулировать в виде следующей теоремы:

Теорема 1. *Общее решение матричного уравнения*

$$AX = XB,$$

где

$$A = \|a_{ih}\|_1^m = U \tilde{A} U^{-1} = U \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)}\} U^{-1},$$

$$B = \|b_{ih}\|_1^n = V \tilde{B} V^{-1} = V \{\mu_1 E^{(q_1)} + H^{(q_1)}, \dots, \mu_v E^{(q_v)} + H^{(q_v)}\} V^{-1},$$

задается формулой

$$X = UX_{\tilde{A}\tilde{B}}V^{-1}. \quad (17)$$

Здесь $X_{\tilde{A}\tilde{B}}$ — общее решение уравнения $\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$ — имеет следующую структуру: $X_{\tilde{A}\tilde{B}}$ разбивается на блоки

$$X_{\tilde{A}\tilde{B}} = (\tilde{X}_{\alpha\beta})_{p_\alpha \times q_\beta} \quad (\alpha = 1, 2, \dots, u; \beta = 1, 2, \dots, v);$$

если $\lambda_\alpha \neq \mu_\beta$, то на месте $X_{\alpha\beta}$ стоит нулевая матрица, если же $\lambda_\alpha = \mu_\beta$, то на месте $X_{\alpha\beta}$ стоит произвольная правильная верхняя треугольная матрица.

$X_{\tilde{A}\tilde{B}}$, а следовательно, и X зависят линейно от N произвольных параметров c_1, c_2, \dots, c_N :

$$X = \sum_{i=1}^N c_i X_j, \quad (18)$$

где N определяется формулой

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta} \quad (19)$$

[здесь $\delta_{\alpha\beta}$ обозначает степень наибольшего общего делителя $(\lambda - \lambda_\alpha)^{p_\alpha}$ и $(\lambda - \mu_\beta)^{q_\beta}$].

Заметим, что матрицы X_1, X_2, \dots, X_N , фигурирующие в формуле (18), суть решения исходного уравнения (1) (матрица X_j получается из X , если параметру c_j дать значение единицы, а остальным параметрам — нулевые значения; $j = 1, 2, \dots, N$). Эти решения линейно независимы, так как в противном случае при некоторых значениях параметров c_1, c_2, \dots, c_N , не равных одновременно нулю, матрица X , а следовательно, и $X_{\widetilde{AB}}$ равнялись бы нулю, что невозможно. Таким образом, равенство (19) показывает, что любое решение исходного уравнения представляет собой линейную комбинацию N линейно независимых решений.

Если матрицы A и B не имеют общих характеристических чисел (характеристические многочлены $|\lambda E - A|$ и $|\lambda E - B|$ взаимно просты), то $N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta} = 0$ и, следовательно, $X = 0$, т. е. в этом случае уравнение (1) имеет только тривиальное нулевое решение $X = 0$.

З а м е ч а н и е. Пусть элементы матриц A и B принадлежат некоторому числовому полю K . Тогда нельзя утверждать, что элементы матриц $U, V, X_{\widetilde{AB}}$, фигурирующих в формуле (17), также принадлежат полю K . Элементы этих матриц можно выбрать в расширенном поле K_1 , которое получается из поля K путем приобщения к последнему корней характеристических уравнений $|\lambda E - A| = 0$ и $|\lambda E - B| = 0$. С такого рода расширением основного поля всегда приходится иметь дело, когда пользуются приведением заданных матриц к нормальной жордановой форме.

Однако матричное уравнение (1) эквивалентно системе m линейных однородных уравнений, где неизвестными служат элементы x_{jk} ($j = 1, 2, \dots, m; k = 1, 2, \dots, n$) искомой матрицы X :

$$\sum_{j=1}^m a_{ij} x_{jk} = \sum_{h=1}^n x_{ih} b_{hk} \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n). \quad (20)$$

Нами доказано, что эта система имеет N линейно независимых решений, где N определяется формулой (19). Но известно, что базисные линейно независимые решения можно выбрать в основном поле K , которому принадлежат коэффициенты уравнений (20). Таким образом, в формуле (18) матрицы X_1, X_2, \dots, X_N можно выбрать так, чтобы их элементы принадлежали полю K . Тогда, придавая в формуле (18) произвольным параметрам все возможные значения из поля K , мы получим все матрицы X с элементами из K , удовлетворяющие уравнению (1)¹⁾.

§ 2. Частный случай: $A = B$. Перестановочные матрицы

Рассмотрим частный случай уравнения (1) — уравнение

$$AX = XA, \quad (21)$$

где $A = \|a_{ik}\|_1^n$ — заданная, а $X = \|x_{ik}\|_1^n$ — искомая матрица. Мы пришли к задаче Фробениуса: определить все матрицы X , перестановочные с данной матрицей A .

1) Матрицы $A = \|a_{ij}\|_1^m$ и $B = \|b_{kl}\|_1^n$ определяют линейный оператор $\tilde{F}(X) = AX - XB$ в пространстве прямоугольных матриц X с размерами $m \times n$. Исследование операторов такого типа проведено в работах [85a, б].

Приведем матрицу A к нормальной жордановой форме:

$$A = U \tilde{A} U^{-1} = U \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \} U^{-1}. \quad (22)$$

Тогда, полагая в формуле (17) $V = U$, $\tilde{B} = \tilde{A}$ и обозначая $X_{\tilde{A}\tilde{A}}$ сокращенно через $X_{\tilde{A}}$, получим все решения уравнения (24), т. е. все матрицы, перестановочные с A , в следующем виде:

$$X = UX_{\tilde{A}}U^{-1}, \quad (23)$$

где $X_{\tilde{A}}$ обозначает произвольную матрицу, перестановочную с \tilde{A} . Как было выяснено в предыдущем параграфе, $X_{\tilde{A}}$ разбивается на u^2 блоков

$$X_{\tilde{A}} = (X_{\alpha\beta})_1^u$$

в соответствии с разбиением на блоки жордановой матрицы \tilde{A} ; $X_{\alpha\beta}$ — нулевая матрица либо произвольная правильная верхняя треугольная матрица в зависимости от того, $\lambda_\alpha \neq \lambda_\beta$ или $\lambda_\alpha = \lambda_\beta$.

Для примера выпишем элементы матрицы $X_{\tilde{A}}$ для случая, когда матрица A имеет следующие элементарные делители:

$$(\lambda - \lambda_1)^4, \quad (\lambda - \lambda_1)^3, \quad (\lambda - \lambda_2)^2, \quad \lambda - \lambda_2 \quad (\lambda_1 \neq \lambda_2).$$

В этом случае $X_{\tilde{A}}$ имеет такой вид:

$$\left| \begin{array}{cccc|ccc|cc|c} a & b & c & d & e & f & g & 0 & 0 & 0 \\ 0 & a & b & c & 0 & e & f & 0 & 0 & 0 \\ 0 & 0 & a & b & 0 & 0 & e & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & h & k & l & m & p & q & 0 & 0 & 0 \\ 0 & 0 & h & k & 0 & m & p & 0 & 0 & 0 \\ 0 & 0 & 0 & h & 0 & 0 & m & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & r & s & t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & r & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w & z \end{array} \right| \quad (a, b, \dots, z \text{ — произвольные параметры}).$$

Число параметров в $X_{\tilde{A}}$ равно N , где $N = \sum_{\alpha, \beta=1}^u \delta_{\alpha\beta}$; здесь $\delta_{\alpha\beta}$ обозначает степень наибольшего общего делителя многочленов $(\lambda - \lambda_\alpha)^{p_\alpha}$ и $(\lambda - \lambda_\beta)^{p_\beta}$.

Введем в рассмотрение инвариантные многочлены матрицы A : $i_1(\lambda)$, $i_2(\lambda)$, \dots , $i_t(\lambda)$; $i_{t+1}(\lambda) = \dots = i_n(\lambda) = 1$. Степени этих многочленов обозначим через $n_1 \geq n_2 \geq \dots \geq n_t > n_{t+1} = \dots = 0$. Так как каждый инвариантный многочлен является произведением нескольких попарно взаимно простых элементарных делителей, то формулу для N можно записать и так:

$$N = \sum_{g, j=1}^t x_{gj}, \quad (24)$$

где x_{gj} — степень наибольшего общего делителя многочленов $i_g(\lambda)$ и $i_j(\lambda)$

$(g, j = 1, 2, \dots, t)$. Но наибольшим общим делителем многочленов $i_g(\lambda)$ и $i_j(\lambda)$ является один из этих же многочленов и потому $x_{gj} = \min(n_g, n_j)$. Отсюда получаем:

$$N = n_1 + 3n_2 + \dots + (2t - 1)n_t.$$

Число N является числом линейно независимых матриц, перестановочных с матрицей A (можно считать, что элементы этих матриц принадлежат основному полю K , содержащему элементы матрицы A ; см. замечание в конце предыдущего параграфа). Мы пришли к теореме:

Теорема 2. Число линейно независимых матриц, перестановочных с матрицей $A = \|a_{ik}\|_1^n$, определяется формулой

$$N = n_1 + 3n_2 + \dots + (2t - 1)n_t, \quad (25)$$

где n_1, n_2, \dots, n_t — степени непостоянных инвариантных многочленов $i_1(\lambda), i_2(\lambda), \dots, i_t(\lambda)$ матрицы A .

Заметим, что

$$n = n_1 + n_2 + \dots + n_t. \quad (26)$$

Из (25) и (26) вытекает:

$$N \geq n, \quad (27)$$

причем знак $=$ имеет место в том и только в том случае, когда $t = 1$, т. е. когда все элементарные делители матрицы A попарно взаимно просты.

Пусть $g(\lambda)$ — некоторый многочлен от λ . Тогда матрица $g(A)$ перестановочна с A . Возникает обратный вопрос: в каком случае любая матрица, перестановочная с A , может быть представлена как многочлен от A ? В этом случае любая матрица, перестановочная с A , была бы линейной комбинацией линейно независимых матриц

$$E, A, A^2, \dots, A^{n_1-1}.$$

В рассматриваемом случае $N = n_1 \leq n$; сопоставляя с (27), получаем: $N = n_1 = n$. Таким образом, мы получили

Следствие 1 из теоремы 2. Все матрицы, перестановочные с A , представляются как многочлены от A в том и только в том случае, когда $n_1 = n$, т. е. когда все элементарные делители матрицы A попарно взаимно просты.

Многочлены от матрицы, перестановочной с A , также перестановочны с A . Поставим вопрос: в каком случае все матрицы, перестановочные с A , представляются в виде многочленов от некоторой (одной и той же) матрицы C ? Допустим, что такой случай имеет место. Тогда, так как матрица C удовлетворяет в силу теоремы Гамильтона—Кэли своему характеристическому уравнению, то любая матрица, перестановочная с C , выразится линейно через матрицы

$$E, C, C^2, \dots, C^{n-1}.$$

Поэтому в рассматриваемом случае $N \leq n$. Сопоставляя с (27), находим $N = n$. Но тогда из (25) и (26) и $n_1 = n$.

Следствие 2 из теоремы 2. Все матрицы, перестановочные с A , представляются в виде многочленов от одной и той же матрицы C в том и только в том случае, когда $n_1 = n$, т. е. когда все элементарные делители у $\lambda E - A$ взаимно просты. В этом случае все матрицы, перестановочные с A , представляются и в виде многочленов от A .

Отметим еще одно очень важное свойство перестановочных матриц.

Теорема 3. Если две матрицы $A = \|a_{ik}\|_1^n$, $B = \|b_{ik}\|_1^n$ перестановочны и одна из них, например A , имеет квазидиагональный вид

$$A = \begin{pmatrix} s_1 & s_2 \\ \widehat{A}_1 & \widehat{A}_2 \end{pmatrix}, \quad (28)$$

где матрицы A_1 и A_2 не имеют общих характеристических чисел, то и другая матрица имеет такой же квазидиагональный вид

$$B = \begin{pmatrix} s_1 & s_2 \\ \widehat{B}_1 & \widehat{B}_2 \end{pmatrix}. \quad (29)$$

Доказательство. Разобьем матрицу B на блоки в соответствии с квазидиагональным видом (28):

$$B = \begin{pmatrix} s_1 & s_2 \\ \widehat{B}_1 & X \\ Y & B_2 \end{pmatrix}.$$

Записывая, что $AB = BA$, получим четыре матричных равенства:

$$1. A_1 B_1 = B_1 A_1, \quad 2. A_1 X = X A_2, \quad 3. A_2 Y = Y A_1, \quad 4. A_2 B_2 = B_2 A_2. \quad (30)$$

Второе и третье из уравнений (30), как было выяснено в § 1 (стр. 203), имеют только нулевые решения $X = 0$, $Y = 0$, поскольку матрицы A_1 и A_2 не имеют общих характеристических чисел. Таким образом, наше предложение доказано. Первое и четвертое из равенств (30) выражают перестановочность матриц A_1 и B_1 , A_2 и B_2 .

Доказанное предложение в геометрической формулировке гласит:

Теорема 3'. Если

$$R = I_1 + I_2$$

есть расщепление всего пространства R на инвариантные относительно оператора A подпространства I_1 и I_2 и минимальные многочлены этих подпространств (относительно A) взаимно просты, то эти подпространства I_1 и I_2 инвариантны относительно любого линейного оператора B , перестановочного с A .

Из доказанной теоремы вытекает следующее следствие¹⁾:

Следствие 1. Если линейные операторы A, B, \dots, L попарно перестановочны, то можно расщепить все пространство R на инвариантные относительно всех операторов A, B, \dots, L подпространства

$$R = I_1 + I_2 + \dots + I_w$$

так, чтобы минимальный многочлен любого из этих подпространств относительно любого из операторов A, B, \dots, L был степенью неприводимого многочлена.

Как частный случай отсюда получим

Следствие 2. Если линейные операторы A, B, \dots, L попарно перестановочны и все характеристические числа этих операторов принадлежат основному полю K , то можно расщепить все пространство R на инвариантные относительно всех операторов A, B, \dots, L подпространства I_1, I_2, \dots, I_w , в каждом из которых любой из операторов A, B, \dots, L имеет равные характеристические числа.

¹⁾ См. также теорему 1 гл. VII (стр. 173).

И, наконец, отметим уже частный случай этого предложения:

Следствие 3. Если операторы простой структуры A, B, \dots, L (см. гл. III, § 8) попарно перестановочны, то можно составить базис пространства из общих собственных векторов этих операторов.

Дадим еще матричную формулировку последнего предложения:

Перестановочные матрицы простой структуры можно одновременно, т. е. одним и тем же преобразованием подобия, привести к диагональному виду.

§ 3. Уравнение $AX - XB = C$

Пусть дано матричное уравнение

$$AX - XB = C, \quad (31)$$

где $A = \|a_{ij}\|_1^m$, $B = \|b_{kl}\|_1^n$ — заданные квадратные матрицы порядков m и n , $C = \|c_{jk}\|$ — заданная, а $X = \|x_{jk}\|$ — искомая прямоугольные матрицы размером $m \times n$. Уравнение (31) эквивалентно системе mn скалярных уравнений относительно элементов матрицы X :

$$\sum_{j=1}^m a_{ij}x_{jk} - \sum_{l=1}^n x_{il}b_{lk} = c_{ik} \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n).$$

Соответствующая однородная система уравнений

$$\sum_{j=1}^m a_{ij}x_{jk} - \sum_{l=1}^n x_{il}b_{lk} = 0 \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n)$$

в матричном виде записывается так:

$$AX - XB = 0. \quad (32)$$

Таким образом, если уравнение (32) имеет только нулевое решение, то уравнение (31) имеет одно-единственное решение. Но в § 1 было установлено, что уравнение (32) имеет только нулевое решение тогда и только тогда, когда матрицы A и B не имеют общих характеристических чисел. Следовательно, если матрицы A и B не имеют общих характеристических чисел, то уравнение (31) имеет одно и только одно решение; если же матрицы A и B имеют общие характеристические числа, то в зависимости от «свободного члена» C могут представиться два случая: либо уравнение (31) противоречиво, либо оно имеет бесчисленное множество решений, задаваемых формулой

$$X = X_0 + X_1,$$

где X_0 — фиксированное частное решение уравнения (31), X_1 — общее решение однородного уравнения (32) (структура X_1 была выяснена в § 1).

§ 4. Скалярное уравнение $f(X) = 0$

Рассмотрим сначала уравнение

$$g(X) = 0, \quad (33)$$

где

$$g(\lambda) = (\lambda - \lambda_1)^{a_1} (\lambda - \lambda_2)^{a_2} \dots (\lambda - \lambda_h)^{a_h}$$

— заданный многочлен переменной λ , а X — искомая квадратная матрица порядка n . Так как минимальный многочлен матрицы X , т. е. первый

инвариантный многочлен $i_1(\lambda)$, должен быть делителем многочлена $g(\lambda)$, то элементарные делители матрицы X должны иметь следующий вид:

$$(\lambda - \lambda_{i_1})^{p_{i_1}}, (\lambda - \lambda_{i_2})^{p_{i_2}}, \dots, (\lambda - \lambda_{i_v})^{p_{i_v}} \begin{pmatrix} i_1, i_2, \dots, i_v = 1, 2, \dots, h, \\ p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_h} \leq a_{i_h}, \\ p_{i_1} + p_{i_2} + \dots + p_{i_v} = n \end{pmatrix}$$

(среди индексов i_1, i_2, \dots, i_v могут быть и равные, n — заданный порядок искомой матрицы X).

Искомая матрица X представится в виде

$$X = T \{ \lambda_{i_1} E^{(p_{i_1})} + H^{(p_{i_1})}, \dots, \lambda_{i_v} E^{(p_{i_v})} + H^{(p_{i_v})} \} T^{-1}, \quad (34)$$

где T — произвольная неособенная матрица порядка n . Множество решений уравнения (33) с заданным порядком искомой матрицы распадается согласно формуле (34) на конечное число классов подобных между собой матриц.

Пример 1. Дано уравнение

$$X^m = 0. \quad (35)$$

Если некоторая степень матрицы равна нулю, то матрица называется *нильпотентной*. Наименьший из показателей, при которых степень матрицы равна нулю, называется *индексом нильпотентности* данной матрицы.

Очевидно, решениями уравнения (35) являются все нильпотентные матрицы с индексом нильпотентности $\mu \leq m$. Формула, охватывающая все решения данного порядка n , выглядит так:

$$X = T \{ H^{(p_1)}, H^{(p_2)}, \dots, H^{(p_v)} \} T^{-1} \begin{pmatrix} p_1, p_2, \dots, p_v \leq m, \\ p_1 + p_2 + \dots + p_v = n \end{pmatrix} \quad (36)$$

(T — произвольная неособенная матрица).

Пример 2. Дано уравнение

$$X^2 = X. \quad (37)$$

Матрица, удовлетворяющая этому уравнению, называется *идемпотентной*. Элементарными делителями идемпотентной матрицы могут быть только λ либо $\lambda - 1$. Поэтому идемпотентную матрицу можно определить как матрицу простой структуры (т. е. приводящуюся к диагональной форме) с характеристическими числами, равными нулю или единице. Формула, охватывающая все идемпотентные матрицы данного порядка, имеет вид

$$X = T \underbrace{\{ 1, 1, \dots, 1, 0, \dots, 0 \}}_n T^{-1}, \quad (38)$$

где T — произвольная неособенная матрица порядка n .

Рассмотрим теперь более общее уравнение

$$f(X) = 0, \quad (39)$$

где $f(\lambda)$ — регулярная функция в некоторой области G плоскости комплексного аргумента λ . От искомого решения $X = \|x_{ik}\|_1^n$ будем требовать, чтобы характеристические числа его принадлежали области G . Выпишем все нули функции $f(\lambda)$, лежащие в области G , и их кратности:

$$\lambda_1, \lambda_2, \dots,$$

$$a_1, a_2, \dots$$

Как и в предыдущем случае, каждый элементарный делитель матрицы X должен иметь вид

$$(\lambda - \lambda_i)^{p_i} (p_i \leq a_i)$$

и потому

$$X = T \{ \lambda_{i_1} E^{(p_{i_1})} + H^{(p_{i_1})}, \dots, \lambda_{i_v} E^{(p_{i_v})} + H^{(p_{i_v})} \} T^{-1} \quad (40)$$

$$(i_1, i_2, \dots, i_v = 1, 2, \dots; p_{i_1} \leq a_{i_1}, p_{i_2} \leq a_{i_2}, \dots, p_{i_v} \leq a_{i_v};$$

$$p_{i_1} + p_{i_2} + \dots + p_{i_v} = n)$$

(T — произвольная неособенная матрица).

§ 5. Матричное многочленное уравнение

Рассмотрим уравнения

$$A_0 X^m + A_1 X^{m-1} + \dots + A_m = 0, \quad (41)$$

$$Y^m A_0 + Y^{m-1} A_1 + \dots + A_m = 0, \quad (42)$$

где A_0, A_1, \dots, A_m — заданные, а X и Y — искомые квадратные матрицы порядка n . Уравнение (33), рассмотренное в предыдущем параграфе, представляет собой весьма частный (можно сказать, тривиальный) случай уравнений (41), (42) и получается из последних, если положить $A_i = a_i E$, где a_i — число и $i = 1, 2, \dots, m$.

Следующая теорема устанавливает связь между уравнениями (41), (42) и (33).

Теорема 4. Каждое решение матричного уравнения

$$A_0 X^m + A_1 X^{m-1} + \dots + A_m = 0$$

удовлетворяет скалярному уравнению

$$g(X) = 0, \quad (43)$$

где

$$g(\lambda) \equiv |A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m|. \quad (44)$$

Этому же скалярному уравнению удовлетворяет и любое решение Y матричного уравнения

$$Y^m A_0 + Y^{m-1} A_1 + \dots + A_m = 0.$$

Доказательство. Обозначим через $F(\lambda)$ матричный многочлен

$$F(\lambda) = A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m.$$

Тогда уравнения (41) и (42) запишутся так (см. стр. 88):

$$F(X) = 0, \quad \widehat{F}(Y) = 0.$$

Согласно обобщенной теореме Безу (гл. IV, § 2), если X и Y — решения этих уравнений, то матричный многочлен $F(\lambda)$ делится справа на $\lambda E - X$ и слева на $\lambda E - Y$:

$$F(\lambda) = Q(\lambda)(\lambda E - X) = (\lambda E - Y)Q_1(\lambda).$$

Отсюда

$$g(\lambda) = |F(\lambda)| = |Q(\lambda)| \Delta(\lambda) = |Q_1(\lambda)| \Delta_1(\lambda), \quad (45)$$

где $\Delta(\lambda) = |\lambda E - X|$ и $\Delta_1(\lambda) = |\lambda E - Y|$ — характеристические многочлены матриц X и Y . По теореме Гамильтона — Кэли (гл. IV, § 3)

$$\Delta(X) = 0, \quad \Delta_1(Y) = 0.$$

Поэтому из (45) вытекает:

$$g(X) = g(Y) = 0.$$

Теорема доказана.

Мы доказали, что каждое решение уравнения (41) удовлетворяет скалярному уравнению (степени $\leq mn$)

$$g(\lambda) = 0.$$

Но множество матричных решений этого уравнения с заданным порядком n распадается на конечное число классов подобных между собой матриц (см. § 4). Поэтому все решения уравнения (41) приходится искать среди матриц вида

$$T_i D_i T_i^{-1} \quad (46)$$

здесь D_i — известные матрицы; при желании можно считать, что D_i имеют нормальную жорданову форму; T_i — произвольные неособенные матрицы n -го порядка; $i = 1, 2, \dots, h$. Подставим в (41) вместо X матрицу (46) и подберем T_i так, чтобы удовлетворялось уравнение (41). Для каждого T_i получим свое линейное уравнение

$$A_0 T_i D_i^m + A_1 T_i D_i^{m-1} + \dots + A_m T_i = 0 \quad (i = 1, 2, \dots, h). \quad (47)$$

Единственный способ, который мы можем предложить для нахождения решения T_i уравнения (47), заключается в замене матричного уравнения системой линейных однородных скалярных уравнений относительно элементов искомой матрицы T_i . Каждое неособенное решение T_i уравнения (47), будучи подставлено в (46), дает решение данного уравнения (41). Аналогичные рассуждения могут быть проведены для уравнения (42).

В следующих двух параграфах мы рассмотрим частные случаи уравнения (41), связанные с извлечением корня m -й степени из матрицы.

Заметим, что теорема Гамильтона — Кэли является частным случаем теоремы 4. В самом деле, любая квадратная матрица A , будучи подставлена вместо λ , удовлетворяет уравнению

$$\lambda E - A = 0.$$

Поэтому в силу доказанной теоремы

$$\Delta(A) = 0,$$

где $\Delta(\lambda) = |\lambda E - A|$.

Теорема 4 может быть обобщена следующим образом:

Теорема 5 (Филлипса)¹⁾. Если попарно перестановочные между собой квадратные матрицы n -го порядка X_0, X_1, \dots, X_m удовлетворяют матричному уравнению

$$A_0 X_0 + A_1 X_1 + \dots + A_m X_m = 0 \quad (48)$$

(A_0, A_1, \dots, A_m — заданные квадратные матрицы n -го порядка), то эти же матрицы X_0, X_1, \dots, X_m удовлетворяют скалярному уравнению

$$g(X_0, X_1, \dots, X_m) = 0, \quad (49)$$

где

$$g(\xi_0, \xi_1, \dots, \xi_m) = |A_0 \xi_0 + A_1 \xi_1 + \dots + A_m \xi_m|. \quad (50)$$

¹⁾ См. [231].

Доказательство. Положим¹⁾

$$F(\xi_0, \xi_1, \dots, \xi_m) = \|f_{ik}(\xi_0, \xi_1, \dots, \xi_m)\|_1^n = A_0\xi_0 + A_1\xi_1 + \dots + A_m\xi_m;$$

$\xi_0, \xi_1, \dots, \xi_m$ — скалярные переменные.

Обозначим через $\widehat{F}(\xi_0, \xi_1, \dots, \xi_m) = \|\widehat{f}_{ik}(\xi_0, \xi_1, \dots, \xi_m)\|_1^n$ присоединенную матрицу для матрицы F [\widehat{f}_{ik} есть алгебраическое дополнение (адьюнкта) элемента f_{ik} в определителе $|F(\xi_0, \xi_1, \dots, \xi_m)| = |f_{ik}|_1^n$ ($i, k = 1, 2, \dots, n$)]. Тогда каждый элемент \widehat{f}_{ik} ($i, k = 1, 2, \dots, n$) матрицы \widehat{F} есть однородный многочлен относительно $\xi_0, \xi_1, \dots, \xi_{m-1}$ степени $m-1$, и потому матрицу \widehat{F} можно представить в виде

$$\widehat{F} = \sum_{j_0+j_1+\dots+j_m=n-1} F_{j_0j_1\dots j_m} \xi_0^{j_0} \xi_1^{j_1} \dots \xi_m^{j_m},$$

где $F_{j_0j_1\dots j_m}$ — некоторые постоянные матрицы порядка n .

Из определения матрицы \widehat{F} следует тождество

$$\widehat{F}F = g(\xi_0, \xi_1, \dots, \xi_m)E.$$

Запишем это тождество следующим образом:

$$\begin{aligned} \sum_{j_0+j_1+\dots+j_{m-1}=n-1} F_{j_0j_1\dots j_m} (A_0\xi_0 + A_1\xi_1 + \dots + A_m\xi_m) \xi_0^{j_0} \xi_1^{j_1} \dots \xi_m^{j_m} = \\ = g(\xi_0, \xi_1, \dots, \xi_m)E. \end{aligned} \quad (51)$$

Переход от левой части к правой части в тождестве (51) осуществляется путем раскрытия скобок и приведения подобных членов. При этом приходится переставлять местами переменные $\xi_0, \xi_1, \dots, \xi_m$ между собой и не приходится переставлять местами переменные $\xi_0, \xi_1, \dots, \xi_m$ с матричными коэффициентами A_i и $F_{j_0j_1\dots j_m}$. Поэтому равенство (51) не нарушится, если мы вместо переменных $\xi_0, \xi_1, \dots, \xi_m$ подставим попарно перестановочные между собой матрицы X_0, X_1, \dots, X_m :

$$\begin{aligned} \sum_{j_0+j_1+\dots+j_m=n-1} F_{j_0j_1\dots j_m} (A_0X_0 + A_1X_1 + \dots + A_mX_m) X_0^{j_0} X_1^{j_1} \dots X_m^{j_m} = \\ = g(X_0, X_1, \dots, X_m). \end{aligned} \quad (52)$$

Но по условию $A_0X_0 + A_1X_1 + \dots + A_mX_m = 0$. Тогда из (52) находим: $g(X_0, X_1, \dots, X_m) = 0$, что и требовалось доказать.

Замечание 1. Теорема 5 сохраняет свою силу, если уравнение (48) заменить уравнением

$$X_0A_0 + X_1A_1 + \dots + X_mA_m = 0. \quad (53)$$

Действительно, теорему 5 можно применить к уравнению

$$A'_0X_0 + A'_1X_1 + \dots + A'_mX_m = 0$$

и затем перейти в этом уравнении почленно к транспонированным матрицам.

Замечание 2. Теорема 4 получится как частный случай теоремы 5, если в качестве X_0, X_1, \dots, X_m взять

$$X^m, X^{m-1}, \dots, X, E.$$

¹⁾ $f_{ik}(\xi_0, \xi_1, \dots, \xi_m)$ — линейные формы от $\xi_0, \xi_1, \dots, \xi_m$ ($i, k = 1, 2, \dots, n$).

§ 6. Извлечение корня m -й степени из неособенной матрицы

Этот и следующий параграфы мы посвятим уравнению

$$X^m = A, \quad (54)$$

где A — заданная, а X — искомая матрицы (обе порядка n), m — данное целое положительное число.

В данном параграфе мы рассмотрим случай, когда $|A| \neq 0$ (A — неособенная матрица). В этом случае все характеристические числа матрицы A отличны от нуля (ибо $|A|$ равен произведению этих характеристических чисел).

Обозначим через

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (55)$$

элементарные делители матрицы A и приведем матрицу A к жордановой форме¹⁾:

$$A = U \tilde{A} U^{-1} = U \{ \lambda_1 E_1 + H_1, \dots, \lambda_u E_u + H_u \} U^{-1}. \quad (56)$$

Так как характеристические числа искомой матрицы X при возведении в m -ю степень дают характеристические числа матрицы A , то и у матрицы X все характеристические числа отличны от нуля. Поэтому на этих характеристических числах производная от

$$f(\lambda) = \lambda^m$$

не обращается в нуль. Но в таком случае (см. гл. VI, стр. 159) элементарные делители матрицы X не «расщепляются» при возведении матрицы X в m -ю степень. Из сказанного следует, что элементарными делителями матрицы X будут:

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u}, \quad (57)$$

где $\xi_j^m = \lambda_j$, т. е. ξ_j является одним из корней m -й степени из λ_j ($\xi_j = \sqrt[m]{\lambda_j}$; $j = 1, 2, \dots, u$).

Определим теперь $\sqrt[m]{\lambda_j E_j + H_j}$ следующим образом. Возьмем в λ -плоскости круг с центром в точке λ_j , не захватывающий нуля. В этом круге мы имеем m раздельных ветвей функции $\sqrt[m]{\lambda}$. Эти ветви можно отличать одну от другой по значениям, которые они принимают в центре круга, в точке λ_j . Обозначим через $\sqrt[m]{\lambda}$ ту ветвь, значение которой в точке λ_j совпадает с характеристическим числом ξ_j искомой матрицы X , и, исходя из этой ветви, определим функцию от матрицы $\sqrt[m]{\lambda_j E_j + H_j}$ с помощью обрывающегося ряда

$$\sqrt[m]{\lambda_j E_j + H_j} = \lambda_j^{\frac{1}{m}} E_j + \frac{1}{m} \lambda_j^{\frac{1}{m}-1} H_j + \frac{1}{2!} \frac{1}{m} \left(\frac{1}{m} - 1 \right) \lambda_j^{\frac{1}{m}-2} H_j^2 + \dots \quad (58)$$

Так как производная от рассматриваемой функции $\sqrt[m]{\lambda}$ в точке λ_j не равна нулю, то матрица (58) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$, где $\xi_j = \sqrt[m]{\lambda_j}$ (здесь $j = 1, 2, \dots, u$). Отсюда следует, что квази-

1) Здесь $E_j = E^{(p_j)}$, $H_j = H^{(p_j)}$ ($j = 1, 2, \dots, u$).

диагональная матрица

$$\{\sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u}\}$$

имеет элементарные делители (57), т. е. те же элементарные делители, что и искомая матрица X . Поэтому существует такая неособенная матрица T ($|T| \neq 0$), что

$$X = T \{\sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u}\} T^{-1}. \quad (59)$$

Для определения матрицы T заметим, что, подставляя в обе части тождества

$$(\sqrt[m]{\lambda})^m = \lambda$$

вместо λ матрицу $\lambda_j E_j + H_j$ ($j = 1, 2, \dots, u$), получим:

$$(\sqrt[m]{\lambda_j E_j + H_j})^m = \lambda_j E_j + H_j \quad (j = 1, 2, \dots, u).$$

Теперь из (54) и (59) следует:

$$A = T \{\lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \dots, \lambda_u E_u + H_u\} T^{-1}. \quad (60)$$

Сопоставляя (56) и (60), найдем:

$$T = UX_{\tilde{A}}, \quad (61)$$

где $X_{\tilde{A}}$ — произвольная неособенная матрица, перестановочная с \tilde{A} (структура матрицы $X_{\tilde{A}}$ детально описана в § 2).

Подставляя в (59) вместо T выражение $UX_{\tilde{A}}$, получаем формулу, охватывающую все решения уравнения (54):

$$X = UX_{\tilde{A}} \{\sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u}\} X_{\tilde{A}}^{-1} U^{-1}. \quad (62)$$

Многозначность правой части этой формулы имеет как дискретный, так и континуальный характер: дискретный (в данном случае и конечный) характер этой многозначности получается за счет выбора различных ветвей функции $\sqrt[m]{\lambda}$ в различных клетках квазидиагональной матрицы (при этом даже при $\lambda_j = \lambda_k$ ветви $\sqrt[m]{\lambda}$ в j -й и в k -й диагональных клетках могут быть различными); континуальный характер многозначности получается за счет произвольных параметров, содержащихся в матрице $X_{\tilde{A}}$.

Все решения уравнения (54) мы будем называть *корнями m -й степени из матрицы A* и обозначать многозначным символом $\sqrt[m]{A}$. Обратим внимание на то, что $\sqrt[m]{A}$ в общем случае не является функцией от матрицы A (т. е. не представляется в виде многочлена от A).

З а м е ч а н и е. Если все элементарные делители матрицы A попарно взаимно просты, т. е. числа $\lambda_1, \lambda_2, \dots, \lambda_u$ все различны, то матрица $X_{\tilde{A}}$ имеет квазидиагональный вид

$$X_{\tilde{A}} = \{X_1, X_2, \dots, X_u\},$$

где матрица X_j перестановочна с $\lambda_j E_j + H_j$ и, следовательно, перестановочна с любой функцией от $\lambda_j E_j + H_j$, в частности с $\sqrt[m]{\lambda_j E_j + H_j}$.

$(j = 1, 2, \dots, u)$. Поэтому в рассматриваемом случае формула (62) принимает вид

$$X = U \left\{ \sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u} \right\} U^{-1}.$$

Таким образом, если элементарные делители матрицы A попарно взаимно просты, то в формуле для $X = \sqrt[m]{A}$ имеется только дискретная многозначность. В этом случае любое значение $\sqrt[m]{A}$ можно представить как многочлен от A .

Пример. Пусть требуется найти все квадратные корни из матрицы

$$A = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{vmatrix},$$

т. е. все решения уравнения

$$X^2 = A.$$

В данном случае матрица A уже имеет нормальную жорданову форму. Поэтому в формуле (62) можно положить $A = \tilde{A}$, $U = E$. Матрица $X_{\tilde{A}}$ в данном случае выглядит так (см. стр. 204):

$$X_{\tilde{A}} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{vmatrix},$$

где a, b, c, d, e — произвольные параметры.

Формула (62), дающая все искомые решения X , в данном случае принимает следующий вид:

$$X = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{vmatrix} \begin{vmatrix} \varepsilon & \frac{\varepsilon}{2} & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \eta \end{vmatrix} \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{vmatrix}^{-1} \quad (\varepsilon^2 = \eta^2 = 1). \quad (63)$$

Не изменяя X , мы можем в формуле (62) помножить $X_{\tilde{A}}$ на такой скаляр, чтобы $|X_{\tilde{A}}| = 1$. В данном случае это приведет к равенству $a^2e = 1$, откуда $e = a^{-2}$.

Вычислим элементы матрицы $X_{\tilde{A}}^{-1}$. Для этого выпишем линейное преобразование с матрицей коэффициентов $X_{\tilde{A}}$:

$$\begin{aligned} y_1 &= ax_1 + bx_2 + cx_3, \\ y_2 &= ax_2, \\ y_3 &= dx_2 + a^{-2}x_3. \end{aligned}$$

Разрешим эту систему уравнений относительно x_1, x_2, x_3 . Тогда получим преобразование с обратной матрицей $X_{\tilde{A}}^{-1}$:

$$\begin{aligned} x_1 &= a^{-1}y_1 - (a^{-2}b - cd)y_2 - acy_3, \\ x_2 &= a^{-1}y_2, \\ x_3 &= -ad y_2 + a^2 y_3. \end{aligned}$$

Отсюда находим:

$$X_{\tilde{A}}^{-1} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{vmatrix}^{-1} = \begin{vmatrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ -ad & a^2 & \end{vmatrix}.$$

Формула (63) дает:

$$\begin{aligned} X &= \begin{vmatrix} \varepsilon & (\varepsilon - \eta) a cd + \frac{\varepsilon}{2} & a^2 c (\eta - \varepsilon) \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta) da^{-1} & \eta \end{vmatrix} = \\ &= \begin{vmatrix} \varepsilon & (\varepsilon - \eta) vw + \frac{\varepsilon}{2} & (\eta - \varepsilon) v \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta) w & \eta \end{vmatrix} \quad (v = a^2 c; w = a^{-1} d). \quad (64) \end{aligned}$$

Решение X зависит от двух произвольных параметров v и w и от двух произвольных знаков ε и η .

§ 7. Извлечение корня m -й степени из особенной матрицы

Переходим к разбору случая, когда $|A| = 0$ (A — особенная матрица).

Как и в первом случае, приведем матрицу A к нормальной жордановой форме:

$$A = U \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)}; H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_t)} \} U^{-1}; \quad (65)$$

здесь мы через $(\lambda - \lambda_1)^{p_1}, \dots, (\lambda - \lambda_u)^{p_u}$ обозначили элементарные делители матрицы A , отвечающие ненулевым характеристическим числам, а через $\lambda^{q_1}, \lambda^{q_2}, \dots, \lambda^{q_t}$ — элементарные делители с нулевыми характеристическими числами.

Тогда

$$A = U \{ A_1, A_2 \} U^{-1}, \quad (66)$$

где

$$A_1 = \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \}, \quad A_2 = \{ H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_t)} \}. \quad (67)$$

Заметим, что A_1 — неособенная матрица ($|A_1| \neq 0$), а A_2 — нильпотентная матрица с индексом нильпотентности $\mu = \max(q_1, q_2, \dots, q_t)$ ($A_2^\mu = 0$).

Из исходного уравнения (54) следует перестановочность матрицы A с искомой матрицей X , а следовательно, и перестановочность подобных им матриц

$$U^{-1}AU = \{ A_1, A_2 \} \quad \text{и} \quad U^{-1}XU. \quad (68)$$

Как было доказано в § 2 (теорема 3), из перестановочности матриц (68) и из того факта, что матрицы A_1 и A_2 не имеют общих характеристических чисел, вытекает, что и вторая из матриц (68) имеет соответствующую квазидиагональную форму

$$U^{-1}XU = \{ X_1, X_2 \}. \quad (69)$$

Заменяя в уравнении (54) матрицы A и X подобными им матрицами

$$\{ A_1, A_2 \} \quad \text{и} \quad \{ X_1, X_2 \},$$

мы заменим уравнение (54) двумя уравнениями:

$$X_1^m = A_1, \quad (70)$$

$$X_2^m = A_2. \quad (71)$$

Так как $|A_1| \neq 0$, то к уравнению (70) применимы результаты предыдущего параграфа. Поэтому X_1 находим по формуле (62):

$$X_1 = X_{A_1} \left\{ \sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \dots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}} \right\} X_{A_1}^{-1}. \quad (72)$$

Таким образом, остается рассмотреть уравнение (71), т. е. заняться нахождением всех корней m -й степени из нильпотентной матрицы A_2 , уже имеющей нормальную жорданову форму:

$$A_2 = \{H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_t)}\}. \quad (73)$$

$\mu = \max(q_1, q_2, \dots, q_t)$ — индекс нильпотентности матрицы A_2 . Из $A_2^\mu = 0$ и из (71) находим:

$$X_2^{m\mu} = 0.$$

Последнее равенство показывает, что искомая матрица X_2 также является нильпотентной с индексом нильпотентности v , где $m(\mu - 1) < v \leq m\mu$. Приведем матрицу X_2 к жордановой форме:

$$\begin{aligned} X_2 = T \{H^{(v_1)}, H^{(v_2)}, \dots, H^{(v_s)}\} T^{-1} \\ (v_1, v_2, \dots, v_s \leq v). \end{aligned} \quad (74)$$

Возведем теперь обе части последнего равенства в m -ю степень. Получим:

$$A_2 = X_2^m = T \{[H^{(v_1)}]^m, [H^{(v_2)}]^m, \dots, [H^{(v_s)}]^m\} T^{-1}. \quad (75)$$

Выясним теперь, какие элементарные делители имеет матрица $[H^{(v)}]^m$ ¹⁾. Обозначим через H линейный оператор, задаваемый матрицей $H^{(v)}$ в v -мерном векторном пространстве с базисом e_1, e_2, \dots, e_v . Тогда из вида матрицы $H^{(v)}$ (в матрице $H^{(v)}$ все элементы первой наддиагонали равны единице и все остальные элементы равны нулю) следует, что

$$He_1 = 0, \quad He_2 = e_1, \quad \dots, \quad He_v = e_{v-1}. \quad (76)$$

Эти равенства показывают, что для оператора H векторы e_1, e_2, \dots, e_v образуют жорданову цепочку векторов, соответствующую элементарному делителю λ^v .

Равенства (76) запишем так:

$$He_j = e_{j-1} \quad (j = 1, 2, \dots, v; e_0 = 0).$$

Очевидно, что

$$H^m e_j = e_{j-m} \quad (j = 1, 2, \dots, v; e_0 = e_{-1} = \dots = e_{-m+1} = 0). \quad (77)$$

Представим число v в виде

$$v = km + r \quad (r < m),$$

1) Ответ на этот вопрос дает теорема 9 гл. VI (стр. 158—159). Здесь мы вынуждены другим методом исследовать этот вопрос, так как нам нужно отыскать не только элементарные делители матрицы $[H^{(v)}]^m$, но и матрицу $P_{v,m}$, преобразующую $[H^{(v)}]^m$ к жордановой форме.

где k, r — целые неотрицательные числа. Расположим базисные векторы e_1, e_2, \dots, e_v следующим образом:

$$\begin{aligned} &e_1, &e_2, &\dots, &e_m, \\ &e_{m+1}, &e_{m+2}, &\dots, &e_{2m}, \\ &\dots &\dots &\dots &\dots \\ &e_{(k-1)m+1}, &e_{(k-1)m+2}, &\dots, &e_{km}, \\ &e_{km+1}, &\dots, &e_{km+r}. \end{aligned} \quad (78)$$

В этой таблице мы имеем m столбцов: первые r содержат по $(k+1)$ векторов в каждом, остальные — по k векторов. Равенство (78) показывает, что векторы каждого столбца образуют жорданову цепочку векторов относительно оператора H^m . Если вместо последовательной нумерации векторов (78) по строкам занумеровать их по столбцам, то в полученном таким образом новом базисе матрица оператора H^m будет иметь следующую нормальную жорданову форму:

$$\underbrace{\{H^{(k+1)}, \dots, H^{(k+1)}\}}_r, \underbrace{\{H^{(k)}, \dots, H^{(k)}\}}_{m-r}^1,$$

и следовательно,

$$[H^{(v)}]^m = P_{v,m} \underbrace{\{H^{(k+1)}, \dots, H^{(k+1)}\}}_r, \underbrace{\{H^{(k)}, \dots, H^{(k)}\}}_{m-r} P_{v,m}^{-1}, \quad (79)$$

где матрица $P_{v,m}$ (матрица перехода от одного базиса к другому) имеет вид (см. гл. III, § 4)

$$P_{v,m} = \left| \begin{array}{cccccc|c} & & & & & & m \\ & \overbrace{1 \ 0 \ \dots \ 0} & & & & & \\ & 0 \ 0 \ \dots \ 0 & & 1 & & & \\ & \vdots & \ddots & \vdots & \ddots & \ddots & \\ & 0 \ 0 \ \dots \ 0 & & & & & \\ & 0 \ 1 \ \dots \ 0 & & & & & \\ & \ddots & & & & & \end{array} \right\} m \quad (80)$$

Матрица $H^{(v)}$ имеет один элементарный делитель λ^v . При возведении матрицы $H^{(v)}$ в m -ю степень этот элементарный делитель «расщепляется». Как показывает формула (79), матрица $[H^{(v)}]^m$ имеет элементарные делители:

$$\underbrace{\lambda^{k+1}, \dots, \lambda^{k+1}}_r, \underbrace{\lambda^k, \dots, \lambda^k}_{m-r}.$$

Возвращаясь теперь к равенству (75), положим:

$$v_i = k_i m + r_i (0 \leq r_i \leq m, k_i \geq 0; i = 1, 2, \dots, s). \quad (81)$$

1) В случае $k=0$ клетки $\underbrace{\{H^{(k)}, \dots, H^{(k)}\}}_{m-r}$ отсутствуют, и эта матрица имеет вид $\underbrace{\{H^{(1)}, \dots, H^{(1)}\}}_r$.

Тогда в силу (79) равенство (75) перепишется так:

$$A_2 = X_2^m = TP \{ \underbrace{H^{(k_1+1)}, \dots, H^{(k_1+1)}}_{r_1}, \underbrace{H^{(k_1)}, \dots, H^{(k_1)}}_{m-r_1}, \\ \underbrace{H^{(k_2+1)}, \dots, H^{(k_2+1)}}_{r_2}, H^{(k_2)}, \dots \} P^{-1} T^{-1}, \quad (82)$$

где $P = \{P_{v_1, m}, P_{v_2, m}, \dots, P_{v_s, m}\}$.

Сопоставляя (82) с (73), видим, что клетки

$$H^{(k_1+1)}, \dots, H^{(k_1+1)}, H^{(k_1)}, \dots, H^{(k_1)}, H^{(k_2+1)}, \dots, H^{(k_2+1)}, \dots \quad (83)$$

с точностью до порядка должны совпасть с клетками

$$H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_t)}. \quad (84)$$

Условимся систему элементарных делителей $\lambda^{v_1}, \lambda^{v_2}, \dots, \lambda^{v_s}$ называть *возможной* для X_2 , если после возведения матрицы в m -ю степень эти элементарные делители, расщепляясь, порождают заданную систему элементарных делителей матрицы A_2 : $\lambda^{q_1}, \lambda^{q_2}, \dots, \lambda^{q_t}$. Число возможных систем элементарных делителей всегда конечно, поскольку

$$\max(v_1, v_2, \dots, v_s) \leq m, \quad v_1 + v_2 + \dots + v_s = n_2 \quad (85)$$

$(n_2$ — степень матрицы $A_2)$.

В каждом конкретном случае возможные системы элементарных делителей для X_2 могут быть легко определены путем конечного числа испытаний.

Покажем, что для каждой возможной системы элементарных делителей $\lambda^{v_1}, \lambda^{v_2}, \dots, \lambda^{v_s}$ существуют соответствующие решения уравнения (71), и определим все эти решения. В этом случае существует преобразующая матрица Q такая, что

$$\{H^{(k_1+1)}, \dots, H^{(k_1+1)}, H^{(k_1)}, \dots, H^{(k_1)}, H^{(k_2+1)}, \dots\} = Q^{-1} A_2 Q. \quad (86)$$

Матрица Q осуществляет перестановку клеток в квазидиагональной матрице, что достигается надлежащей перенумерацией базисных векторов. Поэтому матрицу Q можно считать известной. Используя (86), мы из (82) получим:

$$A_2 = TPQ^{-1}A_2QP^{-1}T^{-1}.$$

Отсюда

$$TPQ^{-1} = X_{A_2}$$

или

$$T = X_{A_2} Q P^{-1}, \quad (87)$$

где X_{A_2} — произвольная матрица, перестановочная с A_2 .

Подставляя выражение (87) для T в (74), будем иметь:

$$X_2 = X_{A_2} Q P^{-1} \{H^{(v_1)}, H^{(v_2)}, \dots, H^{(v_s)}\} P Q^{-1} X_{A_2}^{-1}. \quad (88)$$

Из (69), (72) и (88) получим общую формулу, охватывающую все искомые решения:

$$X = U \{X_{A_1}, X_{A_2} Q P^{-1}\} \{ \sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \dots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}}, \\ H^{(v_1)}, \dots, H^{(v_s)} \} \cdot \{X_{A_1}^{-1}, P Q^{-1} X_{A_2}^{-1}\} U^{-1}. \quad (89)$$

Обратим внимание читателя на то, что корень m -й степени из особенной матрицы не всегда существует. Его существование связано с существованием системы возможных элементарных делителей для матрицы X_2 .

Легко видеть, например, что уравнение

$$X^m = H^{(p)}$$

не имеет решений при $m > 1$, $p > 1$.

Пример. Требуется извлечь корень квадратный из матрицы

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

т. е. найти все решения уравнения

$$X^2 = A.$$

В данном случае $A = A_2$, $X = X_2$, $m = 2$, $t = 2$, $q_1 = 2$, $q_2 = 1$. Матрица X может иметь только один элементарный делитель λ^3 . Поэтому $s = 1$, $v_1 = 3$, $k_1 = 1$, $r_1 = 1$ и [см. (80)]

$$P = P_{3,2} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = P^{-1}, \quad Q = E.$$

Кроме того, как и в примере на стр. 214, можно положить в формуле (88):

$$X_{A_2} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{vmatrix}, \quad X_{A_2}^{-1} = \begin{vmatrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^2 \end{vmatrix}.$$

Из этой формулы получим:

$$X = X_2 = X_{A_2} P^{-1} H^{(3)} P X_{A_2}^{-1} = \begin{vmatrix} 0 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & \beta^{-1} & 0 \end{vmatrix},$$

где $\alpha = ca^{-1} - a^2d$ и $\beta = a^3$ — произвольные параметры.

§ 8. Логарифм матрицы

1. Рассмотрим матричное уравнение

$$e^X = A. \quad (90)$$

Все решения этого уравнения будем называть логарифмами (натуральными) матрицы A и обозначать через $\ln A$.

Характеристические числа λ_j матрицы A связаны с характеристическими числами ξ_j матрицы X формулой $\lambda_j = e^{\xi_j}$; поэтому, если уравнение (90) имеет решение, то все характеристические числа матрицы A отличны от нуля и матрица A является неособенной ($|A| \neq 0$). Таким образом, условие $|A| \neq 0$ является необходимым для существования решения уравнения (90). Ниже мы увидим, что это условие является и достаточным.

Итак, пусть $|A| \neq 0$. Выпишем элементарные делители матрицы A :

$$\begin{aligned} & (\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \\ & (\lambda_1 \lambda_2 \dots \lambda_u \neq 0, p_1 + p_2 + \dots + p_u = n). \end{aligned} \quad (91)$$

В соответствии с этими элементарными делителями приведем матрицу A к нормальной жордановой форме:

$$A = U \tilde{A} U^{-1} = \\ = U \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \} U^{-1}. \quad (92)$$

Так как производная от функции e^{ξ} отлична от нуля при всех значениях ξ , то (см. гл. VI, стр. 159) при переходе от матрицы X к матрице $A = e^X$ элементарные делители не расщепляются, т. е. матрица X имеет элементарные делители

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u}, \quad (93)$$

где $e^{\xi_j} = \lambda_j$ ($j = 1, 2, \dots, u$), т. е. ξ_j есть одно из значений $\ln \lambda_j$ ($j = 1, 2, \dots, u$).

Возьмем в плоскости комплексного переменного λ круг с центром в точке λ_j радиуса $< |\lambda_j|$ и обозначим через $f_j(\lambda) = \ln \lambda$ ту из ветвей функции $\ln \lambda$ в рассматриваемом круге, которая в точке λ_j принимает значение, равное характеристическому числу ξ_j матрицы X ($j = 1, 2, \dots, u$). После этого полагаем:

$$\ln(\lambda_j E^{(p_j)} + H^{(p_j)}) = f_j(\lambda_j E^{(p_j)} + H^{(p_j)}) = \ln \lambda_j E^{(p_j)} + \lambda_j^{-1} H^{(p_j)} + \dots \quad (94)$$

Так как производная от $\ln \lambda$ нигде не обращается в нуль (в конечной части плоскости λ), то матрица (94) имеет только один элементарный делитель $(\lambda - \xi_j)^{p_j}$. В силу этого квазидиагональная матрица

$$\{\ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln(\lambda_2 E^{(p_2)} + (\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)})\} \quad (95)$$

имеет те же элементарные делители, что и искомая матрица X . Поэтому существует такая матрица T ($|T| \neq 0$), что

$$X = T \{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \} T^{-1}. \quad (96)$$

Для определения матрицы T заметим, что

$$A = e^X = T \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \} T^{-1}. \quad (97)$$

Сопоставляя (97) с (92), находим:

$$T = UX_{\tilde{A}}, \quad (98)$$

где $X_{\tilde{A}}$ — произвольная матрица, перестановочная с матрицей \tilde{A} . Подставляя выражение для T из (98) в (96), получим общую формулу, охватывающую все логарифмы матрицы:

$$X = UX_{\tilde{A}} \{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \\ \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \} X_{\tilde{A}}^{-1} U^{-1}. \quad (99)$$

Замечание. Если все элементарные делители матрицы A взаимно просты, то в правой части формулы (99) можно выбросить множители $X_{\tilde{A}}$ и $X_{\tilde{A}}^{-1}$ (см. аналогичное замечание на стр. 213).

2. Выясним, когда вещественная неособенная матрица A имеет вещественный логарифм X . Пусть искомая матрица имеет несколько элементарных делителей, отвечающих характеристическому числу вида $\varrho + i\pi$: $(\lambda - \varrho - i\pi)^{q_1}, \dots, (\lambda - \varrho - i\pi)^{q_t}$. Поскольку матрица X вещественна, то она имеет и сопряженные элементарные делители: $(\lambda - \varrho + i\pi)^{q_1}, \dots, (\lambda - \varrho + i\pi)^{q_t}$. При переходе от матрицы X к матри-

це A элементарные делители не расщепляются, но характеристические числа $\varrho + i\pi$, $\varrho - i\pi$ заменяются в них числами $e^{\varrho+i\pi} = -\mu$, $e^{\varrho-i\pi} = -\mu$, где $\mu = e^\varrho > 0$. Поэтому в системе элементарных делителей матрицы A каждый элементарный делитель, соответствующий отрицательному характеристическому числу (если такие существуют), повторяется четное число раз. Докажем теперь, что это необходимое условие является и достаточным, т. е. что *вещественная неособенная матрица A тогда и только тогда имеет вещественный логарифм X , когда у матрицы A либо совсем нет элементарных делителей, соответствующих отрицательным характеристическим числам¹⁾, либо каждый такой элементарный делитель повторяется четное число раз²⁾.*

Действительно, пусть это условие выполнено. Тогда в квазидиагональной матрице (95) в соответствии с формулой (94) в тех клетках, где λ_i вещественно и положительно, возьмем для $\ln \lambda_i$ вещественное значение; если же в какой-либо клетке имеется комплексное λ_n , то найдется другая клетка такого же размера с $\lambda_g = \bar{\lambda}_n$. В этих клетках возьмем комплексно сопряженные значения для $\ln \lambda_n$ и $\ln \lambda_g$. Каждая же клетка по условию повторяется в (98) четное число раз с сохранением размера клетки. Тогда в половине этих клеток положим $\ln \lambda_k = \ln |\lambda_k| + i\pi$, а в другой половине возьмем $\ln \lambda_k = \ln |\lambda_k| - i\pi$. Тогда в квазидиагональной матрице (98) диагональные клетки либо будут вещественными, либо будут попарно комплексно сопряженными. Но *такая квазидиагональная матрица всегда подобна вещественной матрице³⁾.* Поэтому существует такая неособенная матрица T_1 ($|T_1| \neq 0$), что матрица

$$X_1 = T_1 \{ \ln (\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln (\lambda_n E^{(p_n)} + H^{(p_n)}) \} T_1^{-1}$$

вещественна. Но тогда будет вещественной и матрица

$$A_1 = e^{X_1} = T_1 \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_n E^{(p_n)} + H^{(p_n)} \} T_1^{-1}. \quad (100)$$

Сопоставляя формулу (100) с формулой (92), заключаем, что матрицы A и A_1 подобны между собой (поскольку они подобны одной и той же жордановой матрице). Но две подобные вещественные матрицы могут быть преобразованы друг в друга с помощью некоторой неособенной вещественной матрицы W ($|W| \neq 0$):

$$A = WA_1W^{-1} = We^{X_1}W^{-1} = e^{WX_1W^{-1}}.$$

Тогда матрица $X = WX_1W^{-1}$ будет искомым вещественным логарифмом матрицы A .

1) В этом случае существует вещественный $\ln A = r(A)$, где $r(\lambda)$ — надлежащий интерполяционный многочлен для $\ln_0 \lambda$ (см. стр. 104).

2) Это условие, в частности, выполняется, когда $A = B^2$, где B — вещественная матрица.

3) Для того чтобы убедиться в этом, достаточно показать, что квазидиагональная матрица

$$D = \begin{pmatrix} B & 0 \\ 0 & \bar{B} \end{pmatrix}$$

всегда подобна некоторой вещественной матрице. Здесь $B = U + iV$, $\bar{B} = U - iV$, где U и V — вещественные матрицы. Обозначая через E единичную матрицу тех же размеров, что и B , и полагая

$$T = \begin{pmatrix} E & iE \\ E & -iE \end{pmatrix},$$

легко проверим, что

$$T^{-1} = \begin{pmatrix} \frac{1}{2}E & \frac{1}{2}E \\ \frac{1}{2i}E & -\frac{1}{2i}E \end{pmatrix}, \quad T^{-1}DT = \begin{pmatrix} U & -V \\ V & U \end{pmatrix}$$

ГЛАВА IX

ЛИНЕЙНЫЕ ОПЕРАТОРЫ В УНИТАРНОМ ПРОСТРАНСТВЕ

§ 1. Общие соображения

В главах III и VII мы изучали линейные операторы в произвольном n -мерном векторном пространстве. Все базисы такого пространства равноправны между собой. Данному линейному оператору в каждом базисе отвечает некоторая матрица. Матрицы, отвечающие одному и тому же оператору в различных базисах, подобны между собой. Таким образом, изучение линейных операторов в n -мерном векторном пространстве давало возможность выявить свойства матрицы, присущие одновременно всему классу подобных между собой матриц.

В начале этой главы мы введем метрику в n -мерное векторное пространство, относя специальным образом каждым двум векторам некоторое число — их «скалярное произведение». С помощью скалярного произведения определяется «длина» вектора и «косинус угла» между двумя векторами. Такая метризация приводит нас к унитарному пространству, если основное поле K — поле всех комплексных чисел, и к евклидову пространству, если K — поле всех вещественных чисел.

В настоящей главе мы будем изучать свойства линейных операторов, связанные с метрикой пространства. По отношению к метрике пространства уже не все базисы равноправны. Однако равноправными являются все ортонормированные базисы. Переход от одного ортонормированного базиса к другому осуществляется в унитарном (соответственно евклидовом) пространстве при помощи специального — унитарного (соответственно ортогонального) — преобразования. Поэтому две матрицы, отвечающие одному и тому же линейному оператору в двух различных базисах унитарного (евклидова) пространства, унитарно-подобны (ортогонально-подобны) между собой. Таким образом, изучая линейные операторы в n -мерном метризованном пространстве, мы изучаем те свойства матрицы, которые остаются инвариантными при переходе от данной матрицы к матрице унитарно- или ортогонально-подобной. Это приводит нас естественным образом к исследованию свойств специальных классов матриц (нормальных, эрмитовых, унитарных, симметрических, кососимметрических, ортогональных).

§ 2. Метризация пространства

Рассмотрим векторное пространство R над полем комплексных чисел. Пусть каждым двум векторам x и y из R , заданным в определенном порядке, отнесено некоторое комплексное число, называемое *скалярным произведением* этих векторов и обозначаемое через (xy) или (x, y) . Пусть при этом имеют место следующие свойства «склалярного умножения».

- Для любых векторов x, y, z из R и любого комплексного числа a
1. $(xy) = (\bar{yx})^1)$,
 2. $(ax, y) = a(xy)$,
 3. $(x+y, z) = (xz) + (yz)$.
}
- (1)

В этом случае говорят, что в пространство R внесена *эрмитова метрика*.

Заметим еще, что из 1, 2 и 3 следует для любых x, y, z из R :

- 2'. $(x, ay) = a(x\bar{y})$,
- 3'. $(x, y+z) = (xy) + (xz)$.

Из 1 заключаем, что для любого вектора x скалярное произведение (xx) является вещественным числом.

Если для любого вектора x из R

$$4. \quad (xx) \geq 0, \quad (2)$$

то эрмитова метрика называется *неотрицательной*. Если же при этом

$$5. \quad (xx) > 0 \text{ при } x \neq 0, \quad (3)$$

то эрмитова метрика называется *положительно определенной*.

Определение 1. Векторное пространство R с положительно определенной эрмитовой метрикой мы будем называть *унитарным пространством*²⁾.

В настоящей главе мы будем рассматривать конечномерные унитарные пространства³⁾.

Под длиной вектора x понимают⁴⁾ $|x| = \sqrt{(x, x)}$. Из 2 и 5 следует, что каждый вектор, отличный от нуля, имеет положительную длину и лишь вектор-нуль имеет длину, равную нулю. Вектор x называется *нормированным* (также *единичным вектором* или *ортом*), если $|x|=1$. Для «нормировки» произвольного вектора $x \neq 0$ достаточно умножить этот вектор на любое комплексное число λ , у которого $|\lambda| = \frac{1}{|x|}$.

По аналогии с обычным трехмерным векторным пространством два вектора x и y называются *ортогональными* (обозначение: $x \perp y$), если $(xy) = 0$. В этом случае из 1, 3, 3' следует:

$$(x+y, x+y) = (xx) + (yy)$$

т. е. (теорема Пифагора!)

$$|x+y|^2 = |x|^2 + |y|^2 \quad (x \perp y).$$

Пусть унитарное пространство R имеет конечное число измерений n . Рассмотрим в R произвольный базис e_1, e_2, \dots, e_n . Обозначим через x_i и y_i ($i = 1, 2, \dots, n$) соответственно координаты векторов x и y в этом базисе:

$$x = \sum_{i=1}^n x_i e_i, \quad y = \sum_{i=1}^n y_i e_i.$$

1) Чертка над числом означает переход к комплексно сопряженному числу.

2) Исследование n -мерных векторных пространств с произвольной (не положительно определенной) метрикой проведено в статье [117], а также в книге [24], гл. IX и X.

3) В §§ 2—7 этой главы во всех случаях, когда конечномерность пространства не будет особо оговорена, все рассуждения сохраняют свою силу и для бесконечномерных пространств.

4) Здесь знаком $\sqrt{}$ обозначаем неотрицательное (арифметическое) значение корня.

Тогда в силу 2, 3, 2' и 3'

$$(xy) = \sum_{i,k=1}^n h_{ik} x_i \bar{y}_k, \quad (4)$$

где

$$h_{ik} = (e_i e_k) \quad (i, k = 1, 2, \dots, n). \quad (5)$$

В частности,

$$(xx) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k. \quad (6)$$

Из 1 и (5) следует:

$$h_{ii} = \bar{h}_{ii} \quad (i, k = 1, 2, \dots, n). \quad (7)$$

Форма $\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$, где $h_{ik} = \bar{h}_{ik}$ ($i, k = 1, 2, \dots, n$), называется *эрмитовой*¹⁾. Таким образом, квадрат длины вектора представляется в виде эрмитовой формы его координат. Отсюда и название «эрмитова метрика». Форма, стоящая в правой части равенства (6), является в силу 4 *неотрицательной*:

$$\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k \geq 0, \quad (8)$$

при всех значениях переменных x_1, x_2, \dots, x_n . В силу же дополнительного условия 5 эта форма будет *положительно определенной*, т. е. знак = в (8) будет иметь место только при равенстве нулю всех x_i ($i = 1, 2, \dots, n$).

Определение 2. Систему векторов e_1, e_2, \dots, e_n будем называть ортонормированной, если

$$(e_i e_k) = \delta_{ik} = \begin{cases} 0 & \text{при } i \neq k, \\ 1 & \text{при } i = k \end{cases} \quad (i, k = 1, 2, \dots, n). \quad (9)$$

При $m = n$, где n — число измерений пространства, получаем *ортонормированный базис* пространства.

В § 7 будет доказано, что в каждом n -мерном унитарном пространстве существует ортонормированный базис.

Пусть x_i и y_i ($i = 1, 2, \dots, n$) — соответственно координаты векторов x и y в ортонормированном базисе. Тогда в силу (4), (5) и (9)

$$\left. \begin{aligned} (xy) &= \sum_{i=1}^n x_i \bar{y}_i, \\ (xx) &= \sum_{i=1}^n |x_i|^2. \end{aligned} \right\} \quad (10)$$

Фиксируем произвольно некоторый базис в n -мерном пространстве R . При этом базисе каждая метризация пространства связана с некоторой положительно определенной эрмитовой формой $\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$, и наоборот, согласно (4) каждая такая форма

1) В соответствии с этим выражение, стоящее в правой части равенства (4), называется *билинейной эрмитовой формой* (относительно величин x_1, x_2, \dots, x_n и y_1, y_2, \dots, y_n).

определяет некоторую положительно определенную эрмитову метрику в \mathbf{R} . Однако все эти метрики не дают существенно различных унитарных n -мерных пространств. Действительно, возьмем две такие метрики со скалярным произведением соответственно (xy) и (xy') . По отношению к этим метрикам определим ортонормированные базисы в \mathbf{R} : e_i и e_i ($i=1, 2, \dots, n$). Отнесем друг другу векторы x и x' из \mathbf{R} ($x \rightarrow x'$), имеющие в этих базисах одинаковые координаты. Это соответствие является аффинным¹⁾. Кроме того, в силу (10)

$$(xy) = (x'y')'.$$

Таким образом, с точностью до аффинного преобразования пространства все положительно определенные эрмитовы метризации n -мерного векторного пространства совпадают друг с другом.

Если основным числовым полем K является поле вещественных чисел, то метрика, удовлетворяющая постулатам 1, 2, 3, 4 и 5, называется евклидовой.

Определение 3. Векторное пространство \mathbf{R} над полем вещественных чисел с положительно евклидовой метрикой называется евклидовым пространством.

Если x_i и y_i ($i=1, 2, \dots, n$) суть координаты векторов x и y в некотором базисе e_1, e_2, \dots, e_n n -мерного евклидова пространства, то

$$(xy) = \sum_{i, k=1}^n s_{ik} x_i y_k, \quad |x|^2 = \sum_{i, k=1}^n s_{ik} x_i x_k.$$

Здесь $s_{ik} = s_{ki}$ ($i, k=1, 2, \dots, n$) — вещественные числа²⁾. Выражение $\sum_{i, k=1}^n s_{ik} x_i x_k$ называется квадратичной формой относительно x_1, x_2, \dots, x_n . Из положительности определенности метрики вытекает, что квадратичная форма $\sum_{i, k=1}^n s_{ik} x_i x_k$, задающая аналитически эту метрику, является положительно определенной, т. е. $\sum_{i, k=1}^n s_{ik} x_i x_k > 0$, если $\sum_{i=1}^n x_i^2 > 0$.

При ортонормированном базисе

$$(xy) = \sum_{i=1}^n x_i y_i, \quad |x|^2 = \sum_{i=1}^n x_i^2. \quad (11)$$

При $n=3$ получаем известные формулы для скалярного произведения двух векторов и для квадрата длины вектора в трехмерном евклидовом пространстве.

§ 3. Критерий Грама линейной зависимости векторов

Пусть векторы x_1, x_2, \dots, x_m унитарного или евклидова пространства \mathbf{K} линейно зависимы, т. е. существуют такие не равные одновременно нулю числа c_1, c_2, \dots, c_m ³⁾, что

$$c_1 x_1 + c_2 x_2 + \dots + c_m x_m = 0. \quad (12)$$

1) То есть оператор A , относящий вектору x из \mathbf{R} вектор x' из \mathbf{R} , является линейным и неособенным.

2) $s_{ik} = (e_i e_k)$ ($i, k=1, 2, \dots, n$).

3) В случае евклидова пространства c_1, c_2, \dots, c_m — вещественные числа.

Умножив последовательно обе части этого равенства слева скалярно на x_1, x_2, \dots, x_m , получим:

Рассматривая $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_m$ как ненулевое решение системы линейных однородных уравнений (13) с определителем

$$\Gamma(x_1, x_2, \dots, x_m) = \begin{vmatrix} (x_1x_1) & (x_1x_2) & \dots & (x_1x_m) \\ (x_2x_1) & (x_2x_2) & \dots & (x_2x_m) \\ \dots & \dots & \dots & \dots \\ (x_mx_1) & (x_mx_2) & \dots & (x_mx_m) \end{vmatrix}, \quad (14)$$

заключаем, что этот определитель равен нулю:

$$\Gamma(x_1, x_2, \dots, x_m) = 0.$$

Определитель $\Gamma(x_1, x_2, \dots, x_m)$ называется *определителем Грама*, составленным для векторов x_1, x_2, \dots, x_m .

Пусть, обратно, определитель Грама (14) равен нулю. Тогда система уравнений (13) имеет ненулевое решение c_1, c_2, \dots, c_m . Равенства (13) можно записать так:

$$\left. \begin{array}{l} (x_1, c_1x_1 + c_2x_2 + \dots + c_mx_m) = 0, \\ (x_2, c_1x_1 + c_2x_2 + \dots + c_mx_m) = 0, \\ \dots \dots \dots \dots \dots \dots \\ (x_m, c_1x_1 + c_2x_2 + \dots + c_mx_m) = 0. \end{array} \right\} \quad (13')$$

Умножая почленно эти равенства соответственно на c_1, c_2, \dots, c_m и складывая, получим:

$$|(c_1x_1 + c_2x_2 + \dots + c_mx_m)|^2 = 0;$$

отсюда в силу положительной определенности метрики

$$c_1x_1 + c_2x_2 + \dots + c_mx_m = 0,$$

т. е. векторы x_1, x_2, \dots, x_m линейно зависимы.

Нами доказана

Теорема 1. Для того чтобы векторы x_1, x_2, \dots, x_m были линейно независимы, необходимо и достаточно, чтобы определитель Грама, составленный из этих векторов, не был равен нулю.

Отметим следующее свойство определителя Грама.

Если какой-либо главный минор определителя Грама равен нулю, то равен нулю и сам определитель Грама.

Действительно, главный минор является определителем Грама для части векторов. Из равенства нулю этого главного минора следует линейная зависимость между этими векторами, а значит, и между векторами полной системы.

Пример. Даны n комплексных функций $f_1(t), f_2(t), \dots, f_n(t)$ вещественного аргумента t , кусочно непрерывных в замкнутом интервале $[a, b]$. Требуется определить, при каком условии они будут линейно зависимы. Для этого мы в прост-

ранстве кусочно непрерывных в $[\alpha, \beta]$ функций введем положительно определенную метрику, полагая

$$(f, g) = \int_a^b f(t) \overline{g(t)} dt.$$

Тогда критерий Грама (теорема 1) в применении к данным функциям даст искомое условие

$$\left| \int_a^b f_1(t) \overline{f_1(t)} dt \dots \int_a^b f_n(t) \overline{f_n(t)} dt \right| = 0.$$

§ 4. Ортогональное проектирование

Пусть в унитарном или в евклидовом пространстве R даны произвольный вектор x и некоторое m -мерное подпространство S с базисом x_1, x_2, \dots, x_m . Мы покажем, что вектор x можно (и притом единственным способом) представить в виде суммы

$$\left. \begin{array}{l} x = x_S + x_N, \\ x_S \in S, \text{ а } x_N \perp S \end{array} \right\} \quad (15)$$

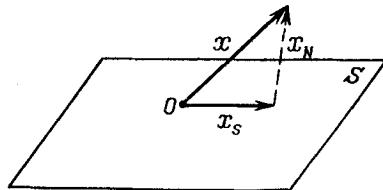


Рис. 5.

(знаком \perp мы обозначаем ортогональность векторов; под ортогональностью к подпространству понимаем ортогональность ко всем векторам из этого подпространства); x_S — ортогональная проекция вектора x на подпространство S , x_N — проектирующий вектор¹⁾.

Пример. R —трехмерное евклидово векторное пространство, а $m=2$. Все векторы будем строить из фиксированной точки O . Тогда S —плоскость, проходящая через O ; x_S —ортогональная проекция вектора x на плоскость S ; x_N —перпендикуляр, опущенный из конца вектора x на плоскость S (рис. 5); $h=|x_N|$ —расстояние конца вектора x от плоскости S .

Для установления разложения (15) искомое x_s представим в виде

$$x_s = c_1 x_1 + c_2 x_2 + \dots + c_m x_m, \quad (16)$$

где c_1, c_2, \dots, c_m — некоторые комплексные числа²⁾.

Для определения этих чисел будем исходить из соотношений

$$(x - x_s, x_k) = 0 \quad (k = 1, 2, \dots, m). \quad (17)$$

Подставляя в (17) вместо x_s его выражение из (16), получим:

$$\left. \begin{aligned} (x_1 x_1) c_1 + \dots + (x_m x_1) c_m + (x x_1) \cdot (-1) &= 0, \\ \vdots &\quad \vdots \\ (x_1 x_m) c_1 + \dots + (x_m x_m) c_m + (x x_m) \cdot (-1) &= 0, \\ x_1 c_1 + \dots + x_m c_m + x s \cdot (-1) &= 0. \end{aligned} \right\} \quad (18)$$

¹⁾ В данном случае x_S — проекция вектора x на подпространство S параллельно подпространству T , состоящему из всех векторов из R , ортогональных к S (см. стр. 69).

2) В случае евклидова пространства c_1, c_2, \dots, c_m — вещественные числа.

Рассматривая эту систему равенств как систему линейных однородных уравнений, имеющую ненулевое решение $c_1, c_2, \dots, c_m, -1$, приравниваем определитель этой системы нулю (предварительно транспонировав его относительно главной диагонали)¹⁾:

$$\begin{vmatrix} (x_1 x_1) & \dots & (x_1 x_m) & x_1 \\ \vdots & \ddots & \vdots & \vdots \\ (x_m x_1) & \dots & (x_m x_m) & x_m \\ (x x_1) & \dots & (x x_m) & x_s \end{vmatrix} = 0. \quad (19)$$

Выделяя из этого определителя член, содержащий x_s , получим (в легко понятных условных обозначениях):

$$x_s = -\frac{\begin{vmatrix} & & x_1 \\ & \Gamma & \vdots \\ & & x_m \\ (x x_1) & \dots & (x x_m) & 0 \end{vmatrix}}{\Gamma}, \quad (20)$$

где $\Gamma = \Gamma(x_1, x_2, \dots, x_m)$ — определитель Грама для векторов x_1, x_2, \dots, x_m (в силу линейной независимости этих векторов $\Gamma \neq 0$). Из (15) и (20) находим:

$$x_N = x - x_s = \frac{\begin{vmatrix} & & x_1 \\ & \Gamma & \vdots \\ & & x_m \\ (x x_1) & \dots & (x x_m) & x \end{vmatrix}}{\Gamma}. \quad (21)$$

Формулы (20) и (21) выражают проекции x_s вектора x на подпространство S и проектирующий вектор x_N через данный вектор x и базис подпространства S .

Обратим внимание еще на одну важную формулу. Обозначим через h линию вектора x_N . Тогда в силу (15) и (21)

$$h^2 = (x_N x_N) = (x_N x) = \frac{\begin{vmatrix} (x_1 x) \\ \vdots \\ (x_m x) \\ (x x_1) & \dots & (x x_m) & (x x) \end{vmatrix}}{\Gamma},$$

т. е.

$$h^2 = \frac{\Gamma(x_1, x_2, \dots, x_m, x)}{\Gamma(x_1, x_2, \dots, x_m)}. \quad (22)$$

Величину h можно еще интерпретировать следующим образом:

Построим векторы x_1, x_2, \dots, x_m, x из одной точки и построим на этих векторах, как на ребрах, $(m+1)$ -мерный параллелепипед. h будет

1) Определитель, стоящий в левой части равенства (19), представляет собой вектор, i -я координата которого получается, если в последнем столбце все векторы x_1, \dots, x_m, x_s заменить их i -ми координатами ($i=1, 2, \dots, n$); координаты берутся в некотором произвольном базисе. Для оправдания перехода от (18) к (19) достаточно в последнем равенстве (18) и в последнем столбце в (19) заменить векторы x_1, \dots, x_m, x_s их i -ми координатами.

высотой этого параллелепипеда, опущенной из конца ребра x на основание S , проходящее через ребра x_1, x_2, \dots, x_m .

Пусть y — произвольный вектор в S , а x — произвольный вектор в R . Если все векторы построить из начала координат n -мерного точечного пространства, то $|x - y|$ и $|x - x_s|$ будут соответственно равны величинам наклонной и высоты, проведенным из конца вектора x к гиперплоскости S ¹⁾. Поэтому, записывая, что *высота короче наклонной*, будем иметь²⁾:

$$h = |x - x_s| \leq |x - y|$$

(знак равенства лишь при $y = x_s$). Таким образом, среди всех векторов $y \in S$ вектор x_s наименее уклоняется от заданного вектора $x \in R$. Величина $h = \sqrt{|x - x_s|^2 - |x - x_s|^2}$ является *квадратичной погрешностью* при приближении $x \approx x_s$ ³⁾.

§ 5. Геометрический смысл определителя Грама и некоторые неравенства

1. Рассмотрим произвольные векторы x_1, x_2, \dots, x_m . Допустим сначала, что эти векторы линейно независимы. В этом случае определитель Грама, составленный для любых из этих векторов, будет отличен от нуля. Тогда, полагая согласно (22)

$$\frac{\Gamma(x_1, x_2, \dots, x_{p+1})}{\Gamma(x_1, x_2, \dots, x_p)} = h_p^2 > 0 \quad (p = 1, 2, \dots, m-1) \quad (23)$$

и перемножая почленно эти неравенства и неравенство

$$\Gamma(x_1) = (x_1 x_1) > 0, \quad (24)$$

получим:

$$\Gamma(x_1, x_2, \dots, x_m) > 0.$$

Таким образом, определитель Грама для линейно независимых векторов положителен, для линейно зависимых равен нулю. Отрицательным определитель Грама никогда не бывает.

Обозначим для сокращения $\Gamma_p = \Gamma(x_1, x_2, \dots, x_p)$ ($p = 1, 2, \dots, m$). Тогда из (23) и (24)

$$\begin{aligned} \sqrt{\Gamma_1} &= |x_1| = V_1, \\ \sqrt{\Gamma_2} &= V_1 h_1 = V_2, \end{aligned}$$

где V_2 — площадь параллелограмма, построенного на x_1 и x_2 . Далее,

$$\sqrt{\Gamma_3} = V_2 h_2 = V_3,$$

где V_3 — объем параллелепипеда, построенного на векторах x_1, x_2, x_3 . Продолжая далее, найдем:

$$\sqrt{\Gamma_4} = V_3 h_3 = V_4,$$

и, наконец,

$$\sqrt{\Gamma_m} = V_{m-1} h_{m-1} = V_m. \quad (25)$$

Естественно V_m назвать объемом m -мерного параллелепипеда, построенного на векторах x_1, x_2, \dots, x_m , как на ребрах⁴⁾.

1) Ср. с примером на стр. 227.

2) $|x - y|^2 = |x_N + x_s - y|^2 = |x_N|^2 + |x_s - y|^2 \geq |x_N|^2 = h^2$.

3) Относительно использования метризованных функциональных пространств в задачах аппроксимации функций см. [2].

4) Формула (25) дает индуктивное определение объема m -мерного параллелепипеда.

Обозначим через $x_{1k}, x_{2k}, \dots, x_{nk}$ координаты вектора x_k ($k = 1, 2, \dots, m$) в некотором ортонормированном базисе в R , и пусть

$$X = \|x_{ik}\| \quad (i = 1, 2, \dots, n; k = 1, 2, \dots, m).$$

Тогда на основании (14)

$$\Gamma_m = |X' \bar{X}|$$

и потому [см. формулу (25)]

$$V_m^2 = \Gamma_m = \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} \text{mod} \begin{vmatrix} x_{i_11} & x_{i_12} & \dots & x_{i_1m} \\ x_{i_21} & x_{i_22} & \dots & x_{i_2m} \\ \dots & \dots & \dots & \dots \\ x_{i_m1} & x_{i_m2} & \dots & x_{i_mm} \end{vmatrix}^2. \quad (26)$$

Это равенство имеет следующий геометрический смысл:

Квадрат объема параллелепипеда равен сумме квадратов объемов его проекций на все координатные m -мерные подпространства. В частности, при $m = n$ из (26) следует:

$$V_n = \text{mod} \begin{vmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{vmatrix}. \quad (26')$$

При помощи формул (20), (21), (22), (26), (26') решается ряд основных метрических задач n -мерной унитарной и евклидовой аналитической геометрии.

2. Вернемся к разложению (15). Из него непосредственно следует:

$$(xx) = (x_s + x_N, x_s + x_N) = (x_s, x_s) + (x_N, x_N) \geq (x_N x_N) = h^2,$$

что в сочетании с (22) дает неравенство (для произвольных векторов x_1, x_2, \dots, x_m, x)

$$\Gamma(x_1, x_2, \dots, x_m, x) \leq \Gamma(x_1, x_2, \dots, x_m) \Gamma(x); \quad (27)$$

при этом знак равенства имеет место тогда и только тогда, когда вектор x ортогонален к векторам x_1, x_2, \dots, x_n .

Отсюда нетрудно получить так называемое *неравенство Адамара*

$$\Gamma(x_1, x_2, \dots, x_m) \leq \Gamma(x_1) \Gamma(x_2) \dots \Gamma(x_m), \quad (28)$$

где знак равенства имеет место тогда и только тогда, когда векторы x_1, x_2, \dots, x_m попарно ортогональны. Неравенство (29) выражает собой следующий геометрически очевидный факт:

Объем параллелепипеда не превосходит произведения длин его ребер и равен этому произведению лишь тогда, когда параллелепипед прямоугольный.

Неравенству Адамара можно придать его обычный вид, полагая в (28) $m = n$ и вводя в рассмотрение определитель Δ , составленный из координат $x_{1k}, x_{2k}, \dots, x_{nk}$ векторов x_k ($k = 1, 2, \dots, n$), в некотором ортонормированном базисе:

$$\Delta = \begin{vmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{nn} \end{vmatrix}.$$

Тогда из (26') и (28) следует:

$$|\Delta|^2 \leq \sum_{i=1}^n |x_{i1}|^2 \sum_{i=1}^n |x_{i2}|^2 \dots \sum_{i=1}^n |x_{in}|^2. \quad (28')$$

3. Установим теперь обобщенное неравенство Адамара, охватывающее как неравенство (27), так и неравенство (28):

$$\Gamma(x_1, x_2, \dots, x_m) \leq \Gamma(x_1, \dots, x_p) \Gamma(x_{p+1}, \dots, x_m), \quad (29)$$

причем знак равенства имеет место тогда и только тогда, когда каждый из векторов x_1, x_2, \dots, x_p ортогонален к любому из векторов x_{p+1}, \dots, x_m либо один из определителей $\Gamma(x_1, \dots, x_p)$, $\Gamma(x_{p+1}, \dots, x_m)$ равен нулю.

Неравенство (28') имеет следующий геометрический смысл:

Объем параллелепипеда не превосходит произведения объемов двух дополнительных «граней» и равен этому произведению в том и только в том случае, когда эти «грани» взаимно ортогональны либо хотя бы одна из них имеет нулевой объем.

Справедливость неравенства (29) установим индуктивно относительно числа векторов x_{p+1}, \dots, x_m . Неравенство справедливо, когда это число равно 1 [см. формулу (27)].

Введем в рассмотрение два подпространства S и S_1 соответственно с базисами x_1, \dots, x_{m-1} и x_{p+1}, \dots, x_{m-1} . Очевидно, $S_1 \subset S$. Рассмотрим ортогональные разложения

$$\begin{aligned} x_m &= x_{S_1} + x_{N_1} \quad (x_{S_1} \in S_1, x_{N_1} \perp S_1), \\ x_{N_1} &= x'_S + x_N \quad (x'_S \in S, x_N \perp S). \end{aligned}$$

Отсюда

$$x_m = x_S + x_N \quad (x_S = x_{S_1} + x'_S, \quad x_N \perp S).$$

Заменяя квадрат объема параллелепипеда произведением квадрата объема основания на квадрат высоты [см. формулу (22)], найдем:

$$\Gamma(x_1, \dots, x_{m-1}, x_m) = \Gamma(x_1, \dots, x_{m-1}) \Gamma(x_N), \quad (30)$$

$$\Gamma(x_{p+1}, \dots, x_{m-1}, x_m) = \Gamma(x_{p+1}, \dots, x_{m-1}) \Gamma(x_{N_1}). \quad (30')$$

При этом из разложения вектора x_{N_1} следует:

$$\Gamma(x_N) \leq \Gamma(x_{N_1}), \quad (31)$$

причем здесь знак $=$ имеет место, лишь когда $x_{N_1} = x_N$.

Используя теперь соотношения (30), (30'), (31) и предположение индукции, получим:

$$\begin{aligned} \Gamma(x_1, \dots, x_m) &= \Gamma(x_1, \dots, x_{m-1}) \Gamma(x_N) \leq \\ &\leq \Gamma(x_1, \dots, x_{m-1}) \Gamma(x_{N_1}) \leq \Gamma(x_1, \dots, x_p) \Gamma(x_{p+1}, \dots, x_{m-1}) \Gamma(x_{N_1}) = \\ &= \Gamma(x_1, \dots, x_p) \Gamma(x_{p+1}, \dots, x_m). \end{aligned} \quad (32)$$

Мы получили неравенство (29). Переходя к выяснению, когда в этом неравенстве имеет место знак $=$, примем, что $\Gamma(x_1, \dots, x_p) \neq 0$ и $\Gamma(x_{p+1}, \dots, x_m) \neq 0$. Тогда согласно (30') также $\Gamma(x_{p+1}, \dots, x_{m-1}) \neq 0$ и $\Gamma(x_{N_1}) \neq 0$. Коль скоро в соотношениях (32) всюду имеет место знак равенства, то $x_{N_1} = x_N$ и, кроме того, по предположению индукции, каждый из векторов x_{p+1}, \dots, x_{m-1} ортогонален к каждому из векторов x_1, \dots, x_p . Этим свойством обладает, очевидно, и вектор

$$x_m = x_{S_1} + x_{N_1} = x_{S_1} + x_N.$$

Таким образом, обобщенное неравенство Адамара установлено полностью.

4. Обобщенному неравенству Адамара (29) можно придать и аналитическую форму.

Пусть $\sum_{i, k=1}^n h_{ik}x_i\bar{x}_k$ — произвольная положительно определенная эрмитова форма. Рассматривая x_1, x_2, \dots, x_n как координаты вектора x в n -мерном пространстве R при базисе e_1, e_2, \dots, e_n , примем форму $\sum_{i, k=1}^n h_{ik}x_i\bar{x}_k$ за основную метрическую форму в R (см. стр. 224). Тогда R станет унитарным пространством. Применим обобщенное неравенство Адамара к базисным векторам e_1, e_2, \dots, e_n :

$$\Gamma(e_1, e_2, \dots, e_n) \leq \Gamma(e_1, \dots, e_p) \Gamma(e_{p+1}, \dots, e_n).$$

Полагая $H = \|h_{ik}\|_1^n$ и замечая, что $(e_i e_k) = h_{ik}$ ($i, k = 1, 2, \dots, n$), мы последнее неравенство сможем записать так:

$$H \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} \leq H \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} H \begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} \quad (p < n); \quad (33)$$

при этом знак равенства имеет место в том и только в том случае, когда $h_{ik} = h_{ki} = 0$ ($i = 1, 2, \dots, p$; $k = p+1, \dots, n$).

Неравенство (33) имеет место для матрицы коэффициентов $H = \|h_{ik}\|_1^n$ произвольной положительно определенной эрмитовой формы. В частности, неравенство (33) имеет место, если H — вещественная матрица коэффициентов положительно определенной квадратичной формы $\sum_{i, k=1}^n h_{ik}x_i\bar{x}_k$ ¹.

5. Обратим внимание читателя на неравенство Буняковского. Для произвольных векторов $x, y \in R$

$$|(xy)|^2 \leq (xx)(yy), \quad (34)$$

причем знак равенства имеет место лишь тогда, когда векторы x и y отличаются скалярным множителем.

Справедливость неравенства Буняковского сразу вытекает из установленного уже неравенства

$$\Gamma(x, y) = \begin{vmatrix} (xx) & (xy) \\ (yx) & (yy) \end{vmatrix} \geq 0.$$

По аналогии со скалярным произведением векторов в трехмерном евклидовом пространстве в n -мерном унитарном пространстве можно ввести «угол» θ между векторами x и y , определив его из соотношения²)

$$\cos^2 \theta = \frac{|(xy)|^2}{(xx)(yy)}.$$

Из неравенства Буняковского следует, что θ имеет вещественное значение.

1) Аналитический вывод обобщенного неравенства Адамара приведен в книге [7], § 8.

2) В случае евклидова пространства угол θ между векторами x и y определяется из формулы

$$\cos \theta = \frac{(xy)}{|x||y|}.$$

§ 6. Ортогонализация ряда векторов

1. Наименьшее подпространство, содержащее векторы x_1, x_2, \dots, x_p , будем обозначать через $[x_1, x_2, \dots, x_p]$. Это подпространство состоит из всевозможных линейных комбинаций $c_1x_1 + c_2x_2 + \dots + c_px_p$ векторов x_1, x_2, \dots, x_p (c_1, c_2, \dots, c_p — комплексные числа)¹⁾. Если векторы x_1, x_2, \dots, x_p линейно независимы, то они образуют базис подпространства $[x_1, x_2, \dots, x_p]$. В этом случае это подпространство имеет p измерений.

Два ряда векторов

$$\mathbf{X}: x_1, x_2, \dots,$$

$$\mathbf{Y}: y_1, y_2, \dots,$$

содержащих одинаковое конечное или оба бесконечное число векторов, назовем *эквивалентными*, если для всех возможных p

$$[x_1, x_2, \dots, x_p] \equiv [y_1, y_2, \dots, y_p] \quad (p = 1, 2, \dots).$$

Ряд векторов

$$\mathbf{X}: x_1, x_2, \dots$$

назовем *невырожденным*, если при любом возможном p векторы x_1, x_2, \dots, x_p линейно независимы.

Ряд векторов называется *ортогональным*, если любые два вектора этого ряда взаимно ортогональны.

Под *ортогонализацией* ряда векторов будем понимать замену этого ряда эквивалентным ортогональным рядом.

Теорема 2. *Всякий невырожденный ряд векторов можно проортогонализировать. Процесс ортогонализации приводит к векторам, определенным однозначно с точностью до скалярных множителей.*

Доказательство. 1. Докажем сначала вторую часть этой теоремы. Пусть два ортогональных ряда $\mathbf{Y}: y_1, y_2, \dots$ и $\mathbf{Z}: z_1, z_2, \dots$ эквивалентны одному и тому же невырожденному ряду $\mathbf{X}: x_1, x_2, \dots$. Тогда ряды \mathbf{Y} и \mathbf{Z} эквивалентны между собой. Поэтому при любом p существуют числа $c_{p1}, c_{p2}, \dots, c_{pp}$ такие, что

$$z_p = c_{p1}y_1 + c_{p2}y_2 + \dots + c_{p, p-1}y_{p-1} + c_{pp}y_p \quad (p = 1, 2, \dots).$$

Умножая последовательно обе части этого равенства скалярно на y_1, y_2, \dots, y_{p-1} и учитывая ортогональность ряда \mathbf{Y} и соотношения

$$z_p \perp [z_1, z_2, \dots, z_{p-1}] \equiv [y_1, y_2, \dots, y_{p-1}],$$

получим: $c_{p1} = c_{p2} = \dots = c_{p, p-1} = 0$ и, следовательно,

$$z_p = c_{pp}y_p \quad (p = 1, 2, \dots).$$

2. Конкретное осуществление процесса ортогонализации произвольного невырожденного ряда векторов $\mathbf{X}: x_1, x_2, \dots$ дается следующим построением.

Пусть $S_p \equiv [x_1, x_2, \dots, x_p]$, $\Gamma_p = \Gamma(x_1, x_2, \dots, x_p)$ ($p = 1, 2, \dots$). Спроектируем ортогонально вектор x_p на подпространство S_{p-1} ($p = 1, 2, \dots$)²⁾:

$$x_p = x_{ps_{p-1}} + x_{pN}, \quad x_{ps_{p-1}} \in S_{p-1}, \quad x_{pN} \perp S_{p-1} \quad (p = 1, 2, \dots).$$

¹⁾ В случае евклидова пространства эти числа вещественны.

²⁾ При $p = 1$ мы полагаем: $x_{1S_0} = 0$, $x_{1N} = x_1$.

Положим

$$\mathbf{y}_p = \lambda_p \mathbf{x}_{pN} \quad (p = 1, 2, \dots; \mathbf{x}_{1N} = \mathbf{x}_1),$$

где λ_p ($p = 1, 2, \dots$) — произвольные отличные от нуля числа.
Тогда (как легко видеть)

$$\mathbf{X}: \mathbf{y}_1, \mathbf{y}_2, \dots$$

— ортогональный ряд, эквивалентный ряду \mathbf{X} . Теорема 2 доказана.
Согласно (24)

$$\mathbf{x}_{pN} = \frac{\begin{vmatrix} x_1 & & & \\ \vdots & & & \\ \Gamma & & & \\ \vdots & & & \\ x_{p-1} & & & \\ \hline (x_p x_1) \dots (x_p x_{p-1}) & x_p \end{vmatrix}}{\Gamma_{p-1}} \quad (p = 1, 2, \dots; \Gamma_0 = 1).$$

Полагая $\lambda_p = \Gamma_{p-1}$ ($p = 1, 2, \dots; \Gamma_0 = 1$), получим для векторов проортогонализированного ряда следующие формулы:

$$\mathbf{y}_1 = \mathbf{x}_1, \mathbf{y}_2 = \begin{vmatrix} (x_1 x_1) & x_1 \\ (x_2 x_1) & x_2 \end{vmatrix}, \dots, \mathbf{y}_p = \begin{vmatrix} (x_1 x_1) & \dots & (x_1 x_{p-1}) & x_1 \\ \dots & \dots & \dots & \dots \\ (x_{p-1} x_1) & \dots & (x_{p-1} x_{p-1}) & x_{p-1} \\ (x_p x_1) & \dots & (x_p x_{p-1}) & x_p \end{vmatrix}, \dots \quad (35)$$

В силу (22)

$$(\mathbf{y}_p \mathbf{y}_p) = \Gamma_{p-1}^2 |\mathbf{x}_{pN}| = \Gamma_{p-1}^2 \cdot \frac{\Gamma_p}{\Gamma_{p-1}} = \Gamma_{p-1} \Gamma_p \quad (p = 1, 2, \dots; \Gamma_0 = 1). \quad (36)$$

Поэтому, полагая

$$z_p = \frac{\mathbf{y}_p}{\sqrt{\Gamma_{p-1} \Gamma_p}} \quad (p = 1, 2, \dots), \quad (37)$$

получим ортонормированный ряд Z , эквивалентный данному ряду \mathbf{X} .

П р и м е р. Определим скалярное произведение в пространстве кусочно непрерывных в интервале $[-1, +1]$ вещественных функций равенством

$$(f, g) = \int_{-1}^{+1} f(x) g(x) dx.$$

Рассмотрим невырожденный ряд «векторов»

$$1, x, x^2, x^3, \dots$$

Проортогонализируем его по формулам (35):

$$y_0 \equiv 1, y_m = \begin{vmatrix} \frac{1}{1} & 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \dots & 1 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \dots & x \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \dots & x^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & x^m \end{vmatrix} \quad (m = 1, 2, \dots)$$

Эти ортогональные между собой многочлены с точностью до постоянных множителей совпадают с известными многочленами Лежандра¹⁾:

$$P_0(x) = 1, \quad P_m(x) = \frac{1}{2^m m!} \frac{d^m (x^2 - 1)^m}{dx^m} \quad (m = 1, 2, \dots).$$

Тот же ряд степеней 1, x , x^2, \dots при другой метрике

$$(f, g) = \int_a^b f(x) g(x) \tau(x) dx \quad [\tau(x) \geq 0 \text{ при } a \leq x \leq b]$$

даст другой ряд ортогональных многочленов.

Так, например, при $a = -1$, $b = 1$ и $\tau(x) = \frac{1}{\sqrt{1-x^2}}$ получаются многочлены Чебышева:

$$T_n(x) = \frac{1}{2^{n-1}} \cos(n \arccos x).$$

При $a = -\infty$, $b = +\infty$ и $\tau(x) = e^{-x^2}$ получаются многочлены Чебышева — Эрмита и т. д.²⁾.

2. Отметим еще так называемое неравенство Бесселя для ортонормированного ряда векторов Z : z_1, z_2, \dots . Пусть дан произвольный вектор x . Обозначим через ξ_p проекцию этого вектора на орт z_p :

$$\xi_p = (xz_p) \quad (p = 1, 2, \dots).$$

Тогда проекция вектора x на подпространство $S_p = [z_1, z_2, \dots, z_p]$ представится в виде [см. (20)]

$$x_{S_p} = \xi_1 z_1 + \xi_2 z_2 + \dots + \xi_p z_p \quad (p = 1, 2, \dots).$$

Но $|x_{S_p}|^2 = |\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_p|^2 \leq |x|^2$. Поэтому для произвольного p

$$|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_p|^2 \leq |x|^2. \quad (38)$$

Это — неравенство Бесселя.

В случае конечномерного пространства n измерений это неравенство имеет совершенно очевидный геометрический смысл. При $p = n$ оно переходит в равенство Пифагора

$$|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2 = |x|^2.$$

В случае бесконечномерного пространства и бесконечного ряда Z из (38) следуют сходимость ряда $\sum_{k=1}^{\infty} |\xi_k|^2$ и неравенство

$$\sum_{k=1}^{\infty} |\xi_k|^2 \leq |x|^2.$$

Составим ряд

$$\sum_{k=1}^{\infty} \xi_k z_k,$$

p -й отрезок этого ряда (при любом p)

$$\xi_1 z_1 + \xi_2 z_2 + \dots + \xi_p z_p$$

¹⁾ См. [19], стр. 77 и далее.

²⁾ Более подробно об этом см. [19], гл. II, § 9, а также [9].

равен проекции x_{S_p} вектора x на подпространство $S_p = [z_1, z_2, \dots, z_p]$ и потому является наилучшим приближением для вектора x в этом подпространстве:

$$\left| (x - \sum_{k=1}^p \xi_k z_k) \right| \leq \left| (x - \sum_{k=1}^p c_k z_k) \right|,$$

где c_1, c_2, \dots, c_p — произвольные комплексные числа. Вычислим соответствующее квадратичное отклонение δ_p :

$$\delta_p^2 = \left| (x - \sum_{k=1}^p \xi_k z_k) \right|^2 = (x - \sum_{k=1}^p \xi_k z_k, x - \sum_{k=1}^p \xi_k z_k) = |x|^2 - \sum_{k=1}^p |\xi_k|^2.$$

Отсюда

$$\lim_{p \rightarrow \infty} \delta_p^2 = |x|^2 - \sum_{k=1}^{\infty} |\xi_k|^2.$$

Если

$$\lim_{p \rightarrow \infty} \delta_p = 0,$$

то говорят, что ряд $\sum_{k=1}^{\infty} \xi_k z_k$ сходится в среднем (сходится по норме) к вектору x .

В этом случае для вектора x из R имеет место равенство (теорема Пифагора в бесконечномерном пространстве!)

$$(xx) = |x|^2 = \sum_{k=1}^{\infty} |\xi_k|^2. \quad (39)$$

Если для любого вектора x из R ряд $\sum_{k=1}^p \xi_k z_k$ в среднем сходится к вектору x , то ортонормированный ряд векторов z_1, z_2, \dots называется полным. В этом случае, заменяя в (39) x на $x+y$ и используя равенство (39) трижды, для векторов $x+y, x$ и y мы легко получим:

$$(xy) = \sum_{k=1}^{\infty} \xi_k \bar{\eta}_k \quad [\xi_k = (xz_k), \eta_k = (yz_k); k = 1, 2, \dots]. \quad (40)$$

П р и м е р. Рассмотрим пространство всех комплексных функций $f(t)$ (t — вещественный аргумент), кусочно непрерывных в замкнутом интервале $[0, 2\pi]$. Скалярное произведение двух функций $f(t)$ и $g(t)$ определим формулой

$$(f, g) = \int_0^{2\pi} f(t) \overline{g(t)} dt.$$

В частности,

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

Возьмем бесконечную последовательность функций

$$\frac{1}{\sqrt{2\pi}} e^{ikt} \quad (k = 0, \pm 1, \pm 2, \dots).$$

Эти функции образуют ортонормированный ряд, так как

$$\int_0^{2\pi} e^{i\mu t} e^{-ivt} dt = \int_0^{2\pi} e^{i(\mu-v)t} dt = \begin{cases} 0 & \text{при } \mu \neq v, \\ 2\pi & \text{при } \mu = v. \end{cases}$$

Ряд

$$\sum_{k=-\infty}^{\infty} f_k e^{ikt} \quad \left(f_k = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt; \quad k=0, \pm 1, \pm 2, \dots \right)$$

сходится в среднем к функции $f(t)$ в интервале $[0, 2\pi]$. Этот ряд называется *рядом Фурье* для функции $f(t)$, а коэффициенты $f_k (k=0, \pm 1, \pm 2, \dots)$ — коэффициентами Фурье для $f(t)$.

В теории рядов Фурье доказывается, что система функций $e^{ikt} (k=0, \pm 1, \pm 2, \dots)$ является полной¹⁾.

Условие полноты дает *равенство Парсеваля* [см. равенство (40)]

$$\int_0^{2\pi} f(t) \overline{g(t)} dt = \sum_{k=-\infty}^{+\infty} \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt \int_0^{2\pi} \overline{g(t)} e^{ikt} dt.$$

Если $f(t)$ — вещественная функция, то f_0 вещественно, а f_k и f_{-k} — комплексно сопряженные числа ($k=1, 2, \dots$). Полагая

$$f_k = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt = \frac{1}{2} (a_k - ib_k),$$

где

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt dt, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt dt \quad (k=0, 1, 2, \dots),$$

будем иметь:

$$f_k e^{ikt} + f_{-k} e^{-ikt} = a_k \cos kt + b_k \sin kt \quad (k=1, 2, \dots).$$

Поэтому для вещественной функции $f(t)$ ряд Фурье принимает вид

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt) \left(\begin{array}{l} a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt dt, \\ b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt dt, \\ k=0, 1, 2, \dots \end{array} \right).$$

§ 7. Ортонормированный базис

Базис любого конечномерного подпространства S в унитарном или евклидовом пространстве R является невырожденным рядом векторов и потому согласно теореме 2 предыдущего параграфа может быть проортогонализирован и пронормирован. Таким образом, в любом конечномерном подпространстве S (и, в частности, во всем пространстве R , если оно конечномерно) существует ортонормированный базис.

Пусть e_1, e_2, \dots, e_n — ортонормированный базис пространства R . Обозначим через x_1, x_2, \dots, x_n координаты произвольного вектора x в этом базисе:

$$x = \sum_{k=1}^n x_k e_k.$$

¹⁾ См., например, [19], гл. II.

Умножая обе части этого равенства справа на e_k и учитывая ортонормированность базиса, легко найдем:

$$x_k = (xe_k) \quad (k=1, 2, \dots, n),$$

т. е. в ортонормированном базисе координата вектора равна скалярному произведению его на соответствующий базисный орт

$$x = \sum_{k=1}^n (xe_k) e_k. \quad (41)$$

Пусть x_1, x_2, \dots, x_n и x'_1, x'_2, \dots, x'_n суть соответственно координаты одного и того же вектора x в двух различных ортонормированных базисах e_1, e_2, \dots, e_n и e'_1, e'_2, \dots, e'_n унитарного пространства R . Формулы преобразования координат имеют вид

$$x_i = \sum_{k=1}^n u_{ik} x'_k \quad (i=1, 2, \dots, n). \quad (42)$$

При этом коэффициенты $u_{1k}, u_{2k}, \dots, u_{nk}$, образующие k -й столбец матрицы $U = \|u_{ik}\|_1^n$, являются, как нетрудно видеть, координатами вектора e'_k в базисе e_1, e_2, \dots, e_n . Поэтому, записывая в координатах [см. (10)] условия ортонормированности базиса e'_1, e'_2, \dots, e'_n , получим соотношения

$$\sum_{i=1}^n u_{ik} \bar{u}_{il} = \delta_{kl} = \begin{cases} 1 & k=l, \\ 0 & k \neq l. \end{cases} \quad (43)$$

Преобразование (42), у которого коэффициенты удовлетворяют условию (43), называется *унитарным*, а соответствующая матрица U — *унитарной матрицей*. Таким образом, в n -мерном унитарном пространстве переход от одного ортонормированного базиса к другому ортонормированному осуществляется при помощи унитарного преобразования координат.

Пусть дано n -мерное евклидово пространство R . Переход от одного ортонормированного базиса в R к другому осуществляется при помощи преобразования координат

$$x_i = \sum_{k=1}^n v_{ik} x'_k \quad (i=1, 2, \dots, n), \quad (44)$$

коэффициенты которого связаны между собой соотношениями

$$\sum_{i=1}^n v_{ik} v_{il} = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (45)$$

Такое преобразование координат называется *ортогональным*, а соответствующая матрица V — *ортогональной матрицей*.

Отметим интересную матричную запись процесса ортогонализации. Пусть $A = \|a_{ik}\|_1^n$ — произвольная неособенная матрица ($|A| \neq 0$) с комплексными элементами. Рассмотрим унитарное пространство R с ортонормированным базисом e_1, e_2, \dots, e_n и определим линейно независимые векторы a_1, a_2, \dots, a_n равенством

$$a_k = \sum_{i=1}^n a_{ik} e_i \quad (k=1, 2, \dots, n).$$

Подвергнем векторы a_1, a_2, \dots, a_n процессу ортогонализации. Полученный ортонормированный базис в R обозначим через u_1, u_2, \dots, u_n . Пусть при этом

$$u_k = \sum_{i=1}^n u_{ik} e_i \quad (k=1, 2, \dots, n).$$

Тогда

$$[a_1 a_2 \dots a_p] = [u_1 u_2 \dots u_p] \quad (p=1, 2, \dots, n),$$

т. е.

$$\begin{aligned} a_1 &= c_{11} u_1, \\ a_2 &= c_{12} u_1 + c_{22} u_2, \end{aligned}$$

$$\dots \dots \dots$$

$$a_n = c_{1n} u_1 + c_{2n} u_2 + \dots + c_{nn} u_n,$$

где c_{ik} ($i, k = 1, 2, \dots, n$; $i \leq k$) — некоторые комплексные числа.
Полагая $c_{ik}=0$ при $i > k$ ($i, k = 1, 2, \dots, n$), будем иметь:

$$a_k = \sum_{p=1}^n c_{pk} u_p \quad (k=1, 2, \dots, n).$$

Переходя здесь к координатам и вводя верхнюю треугольную матрицу $C = \|c_{ik}\|_1^n$ и унитарную матрицу $U = \|u_{ik}\|_1^n$, получим:

$$a_{ik} = \sum_{p=1}^n u_{ip} c_{pk} \quad (i, k = 1, 2, \dots, n),$$

или

$$A = UC. \quad (*)$$

Согласно этой формуле произвольная неособенная матрица $A = \|a_{ik}\|_1^n$ представима в виде произведения унитарной матрицы U на верхнюю треугольную C .

Так как процесс ортогонализации однозначно определяет векторы u_1, u_2, \dots, u_n с точностью до скалярных множителей $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ ($|\varepsilon_i|=1; i=1, 2, \dots, n$), то в формуле (*) множители U и C определяются однозначно с точностью до диагонального множителя $M = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$:

$$U = U_1 M, \quad C = M^{-1} C_1.$$

В этом можно убедиться и непосредственно.

Замечание 1. Если A — вещественная матрица, то в формуле (*) множители U и C можно выбрать вещественными. В этом случае U — ортогональная матрица.

Замечание 2. Формула (*) сохраняет свою силу и для особенной матрицы A ($|A|=0$). В этом можно убедиться, полагая $A = \lim_{m \rightarrow \infty} A_m$, где $|A_m| \neq 0$ ($m=1, 2, \dots$).

Тогда $A_m = U_m C_m$ ($m=1, 2, \dots$). Выделяя из последовательности U_m сходящуюся подпоследовательность U_{m_p} ($\lim_{p \rightarrow \infty} U_{m_p} = U$) и переходя к пределу, из равенства

$A_m = U_{m_p} C_{m_p}$ при $p \rightarrow \infty$, получим искомое разложение $A = UC$. Однако в случае $|A|=0$ множители U и C уже не определяются однозначно с точностью до диагонального множителя M .

Замечание 3. Вместо (*) можно получить формулу

$$A = DW, \quad (**)$$

где D — нижняя треугольная, а W — унитарная матрица. Действительно, применяя установленную ранее формулу (*) к транспонированной матрице A'

$$A' = UC$$

и полагая $W = U'$, $D = C'$, получим (**)¹).

§ 8. Сопряженный оператор

Пусть в n -мерном унитарном пространстве R задан произвольный линейный оператор.

Определение 4. Линейный оператор A^* называется *сопряженным* по отношению к оператору A в том и только в том случае, если

1) Из унитарности матрицы U следует унитарность и матрицы U' , так как условия унитарности (43), записанные в матричном виде: $U' \bar{U} = E$, влечут: $U \bar{U}' = E$.

для любых двух векторов \mathbf{x} , \mathbf{y} из \mathbf{R} выполняется равенство

$$(Ax, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y}). \quad (46)$$

Мы докажем, что для каждого линейного оператора A существует сопряженный оператор A^* и притом только один. Для доказательства выберем в \mathbf{R} некоторый ортонормированный базис e_1, e_2, \dots, e_n . Тогда [см. (41)] для искомого оператора A^* и произвольного вектора \mathbf{y} из \mathbf{R} должно выполняться равенство

$$A^*\mathbf{y} = \sum_{k=1}^n (A^*\mathbf{y}, e_k) e_k.$$

В силу (46) это равенство может быть переписано так:

$$A^*\mathbf{y} = \sum_{k=1}^n (\mathbf{y}, Ae_k) e_k. \quad (47)$$

Примем теперь равенство (47) за определение оператора A^* .

Легко проверить, что определенный таким образом оператор A^* является линейным и удовлетворяет равенству (46) при произвольных векторах \mathbf{x} и \mathbf{y} из \mathbf{R} . Кроме того, равенство (47) однозначно определяет оператор A^* . Таким образом, устанавливаются существование и единственность сопряженного оператора A^* .

Пусть A — линейный оператор в унитарном пространстве, а $A = \|a_{ik}\|_1^n$ — матрица, отвечающая этому оператору в ортонормированном базисе e_1, e_2, \dots, e_n . Тогда, применяя формулу (41) к вектору $Ae_k = \sum_{i=1}^n a_{ik}e_i$, получим:

$$a_{ik} = (Ae_k, e_i) \quad (i, k = 1, 2, \dots, n). \quad (48)$$

Пусть теперь сопряженному оператору A^* в этом же базисе отвечает матрица $A^* = \|a_{ik}^*\|_1^n$. Тогда по формуле (48).

$$a_{ik}^* = (A^*e_k, e_i) \quad (i, k = 1, 2, \dots, n). \quad (49)$$

Из (48) и (49) в силу (46) следует:

$$a_{ik}^* = \bar{a}_{ki} \quad (i, k = 1, 2, \dots, n),$$

т. е.

$$A^* = \bar{A}'.$$

Матрица A^* является транспонированной и комплексно сопряженной для A . Такую матрицу принято называть (см. главу I) *сопряженной* по отношению к A .

Таким образом, в ортонормированном базисе *сопряженным* операторам отвечают *сопряженные* матрицы.

Из определения сопряженного оператора вытекают следующие его свойства:

$$1^\circ \quad (A^*)^* = A,$$

$$2^\circ \quad (A + B)^* = A^* + B^*,$$

$$3^\circ \quad (\alpha A)^* = \alpha A^* \quad (\alpha — скаляр),$$

$$4^\circ \quad (AB)^* = B^*A^*.$$

Введем теперь одно важное понятие. Пусть S — произвольное подпространство в R . Обозначим через T совокупность всех векторов y из R , ортогональных к S . Легко видеть, что T есть тоже подпространство в R и что каждый вектор x из R однозначно представляется в виде суммы $x = x_S + x_T$, где $x_S \in S$, $x_T \in T$, т. е. имеет место расщепление

$$R = S + T, \quad S \perp T.$$

Это расщепление получаем, применяя к произвольному вектору x из R разложение (15) предыдущего параграфа. T называется *ортогональным дополнением* к S . Очевидно, S будет ортогональным дополнением к T . Мы пишем $S \perp T$, понимая под этим то, что любой вектор из S ортогонален любому вектору из T .

Теперь мы сможем сформулировать фундаментальное свойство сопряженного оператора:

5° *Если некоторое подпространство S инвариантно относительно A , то ортогональное дополнение T этого подпространства будет инвариантно относительно A^* .*

Действительно, пусть $x \in S$, $y \in T$. Тогда из $Ax \in S$ следует $(Ax, y) = 0$ и отсюда в силу (46) $(x, A^*y) = 0$. Так как x — произвольный вектор из S , то $A^*y \in T$, что и требовалось доказать.

Введем следующее определение:

Определение 5. Две системы векторов x_1, x_2, \dots, x_m и y_1, y_2, \dots, y_m назовем *биортонормированными*, если

$$(x_i y_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, m), \quad (50)$$

где δ_{ik} — символ Кронекера.

Теперь докажем следующее предложение:

6° *Если A — линейный оператор простой структуры, то сопряженный оператор A^* также имеет простую структуру, причем можно так выбрать полные системы собственных векторов x_1, x_2, \dots, x_n и y_1, y_2, \dots, y_n операторов A и A^* , чтобы они были биортонормированы:*

$$Ax_i = \lambda_i x_i, \quad A^*y_i = \mu_i y_i, \quad (x_i y_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

Действительно, пусть x_1, x_2, \dots, x_n — полная система собственных векторов оператора A . Введем обозначение

$$S_k = [x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n] \quad (k = 1, 2, \dots, n).$$

Рассмотрим одномерное ортогональное дополнение $T_k = [y_k]$ к $(n-1)$ -мерному подпространству S_k ($k = 1, 2, \dots, n$). Тогда T_k инвариантно относительно A^* :

$$A^*y_k = \mu_k y_k, \quad y_k \neq 0 \quad (k = 1, 2, \dots, n).$$

Из $S_k \perp y_k$ следует: $(x_k y_k) \neq 0$, так как в противном случае вектор y_k должен был бы равняться нулю. Помножая x_k, y_k ($k = 1, 2, \dots, n$) на надлежащие числовые множители, получим:

$$(x_i y_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

Из биортонормированности систем векторов x_1, x_2, \dots, x_n и y_1, y_2, \dots, y_n следует, что векторы каждой из этих систем линейно независимы.

Отметим еще такое предложение:

7° *Если операторы A и A^* имеют общий собственный вектор, то характеристические числа этих операторов, отвечающие общему собственному вектору, комплексно сопряжены.*

В самом деле, пусть $Ax = \lambda x$, $A^*x = \bar{\mu}x$ ($x \neq 0$). Тогда, полагая в (46) $y = x$, будем иметь $\lambda(x, x) = \bar{\mu}(x, x)$, откуда $\lambda = \bar{\mu}$.

8° Пусть y — собственный вектор оператора A^* , и пусть $S^{(n-1)}$ — ортогональное дополнение к одномерному подпространству $T = [y]$. Поскольку $A = (A^*)^*$, то согласно утверждению 5° подпространство $S^{(n-1)}$ инвариантно относительно оператора A . Таким образом, *у всякого линейного оператора в n -мерном унитарном пространстве существует $(n-1)$ -мерное инвариантное подпространство*.

Рассматривая далее оператор A в подпространстве $S^{(n-1)}$, мы сможем указать на основании установленного предложения *$(n-2)$ -мерное инвариантное подпространство $S^{(n-2)}$ оператора A , принадлежащее $S^{(n-1)}$.* Повторяя рассуждение, мы построим цепочку из n последовательно вложенных инвариантных подпространств оператора A (индекс наверху указывает размерность):

$$S^{(1)} \subset S^{(2)} \subset \dots \subset S^{(n-1)} \subset S^{(n)} = R.$$

Пусть теперь e_1 — нормированный вектор, принадлежащий $S^{(1)}$. Выберем в $S^{(2)}$ нормированный вектор e_2 такой, что $(e_1, e_2) = 0$. В $S^{(3)}$ найдем нормированный вектор e_3 такой, что $(e_1, e_3) = 0$ и $(e_2, e_3) = 0$. Продолжая этот процесс, мы построим ортонормированный базис векторов

$$e_1, e_2, \dots, e_n,$$

обладающий тем свойством, что каждое подпространство, натянутое на первые k базисных векторов

$$S^{(k)} = [e_1, e_2, \dots, e_k] \quad (k = 1, 2, \dots, n),$$

инвариантно относительно оператора A .

Пусть теперь $\|a_{ij}\|_1^n$ — матрица оператора A в построенном базисе. Мы имеем $Ae_j = \sum_{i=1}^n a_{ij}e_i$, где $a_{ij} = (Ae_j, e_i)$. Поскольку Ae_j принадлежит $S^{(j)}$, то при $i > j$ $a_{ij} = (Ae_j, e_i) = 0$ и, следовательно, матрица оператора является верхней треугольной. Мы пришли к следующей теореме:

Для любого линейного оператора A в n -мерном унитарном пространстве можно построить ортонормированный базис, в котором матрица этого оператора является треугольной.

Это предложение принято называть теоремой Шура. Разумеется, привлекая общую теорему о приведении матрицы оператора к жордановой форме, легко доказать теорему Шура последовательной ортогонализацией жорданова базиса. Приведенное доказательство по существу использует лишь существование у линейного оператора, действующего в n -мерном унитарном пространстве, собственного вектора.

§ 9. Нормальные операторы в унитарном пространстве

Определение 6. Линейный оператор A называется *нормальным*, если он перестановочен со своим сопряженным:

$$AA^* = A^*A. \quad (51)$$

Определение 7. Линейный оператор H называется *эрмитовым*, если он равен своему сопряженному:

$$H^* = H. \quad (52)$$

Определение 8. Линейный оператор U называется *унитарным*, если он обратен своему сопряженному:

$$UU^* = E. \quad (53)$$

Заметим, что унитарный оператор можно определить как изометрический оператор в эрмитовом пространстве, т. е. как оператор, сохраняющий метрику.

Действительно, пусть при произвольных векторах x и y из R

$$(Ux, Uy) = (x, y). \quad (54)$$

Тогда согласно (46) $(U^*Ux, y) = (x, y)$ и, следовательно, в силу произвольности вектора y $U^*Ux = x$, т. е. $U^*U = E$ или $U^* = U^{-1}$. Обратно, из (53) следует (54).

Из (53) или (54) вытекает, что 1° произведение двух унитарных операторов есть снова унитарный оператор, 2° единичный оператор E является унитарным и 3° обратный оператор для унитарного есть также унитарный оператор. Поэтому совокупность всех унитарных операторов является группой¹⁾. Эту группу называют *унитарной группой*.

Эрмитов оператор и унитарный оператор являются частными видами нормального оператора.

Теорема 3. *Произвольный линейный оператор A всегда можно представить в виде*

$$A = H_1 + iH_2, \quad (55)$$

где H_1 и H_2 — эрмитовы операторы («эрмитовы компоненты» оператора A). Эрмитовы компоненты однозначно определяются заданием оператора A . Оператор A нормален тогда и только тогда, когда его эрмитовы компоненты H_1 и H_2 перестановочны между собой.

Доказательство. Пусть имеет место (55). Тогда

$$A^* = H_1 - iH_2. \quad (56)$$

Из (55) и (56) находим:

$$H_1 = \frac{1}{2} (A + A^*), \quad H_2 = \frac{1}{2i} (A - A^*). \quad (57)$$

Обратно, формулы (57) определяют эрмитовы операторы H_1 и H_2 , связанные с A равенством (55).

Пусть теперь A — нормальный оператор: $AA^* = A^*A$. Тогда из (57) следует: $H_1H_2 = H_2H_1$. Обратно, из $H_1H_2 = H_2H_1$ в силу (55) и (56) следует: $AA^* = A^*A$. Теорема доказана.

Представление произвольного линейного оператора A в виде (55) является аналогом представления произвольного комплексного числа z в виде $x_1 + ix_2$, где x_1 и x_2 — вещественные числа.

¹⁾ См. сноску ⁵⁾ к стр. 28.

Пусть в некотором ортонормированном базисе операторам A , H и U отвечают соответственно матрицы A , H , U . Тогда операторным равенствам

$$AA^* = A^*A, \quad H^* = H, \quad UU^* = E \quad (58)$$

будут соответствовать матричные равенства

$$AA^* = A^*A, \quad H^* = H, \quad UU^* = E. \quad (59)$$

Поэтому мы и определяем *нормальную матрицу* как матрицу, перестановочную со своей сопряженной, *эрмитову* как равную своей сопряженной и, наконец, *унитарную* как обратную своей сопряженной.

Тогда в ортонормированном базисе нормальному (эрмитову, унитарному) оператору отвечает соответственно нормальная (эрмитова, унитарная) матрица.

Эрмитова матрица $H = \|h_{ik}\|_1^n$ в силу (59) характеризуется следующими соотношениями между элементами:

$$h_{ki} = \bar{h}_{ik} \quad (i, k = 1, 2, \dots, n),$$

т. е. эрмитова матрица всегда является матрицей коэффициентов некоторой эрмитовой формы (см. § 1).

Унитарная матрица $U = \|u_{ik}\|_1^n$ в силу (59) характеризуется следующими соотношениями между элементами:

$$\sum_{j=1}^n u_{ij} \bar{u}_{kj} = \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (60)$$

Так как из $UU^* = E$ следует $U^*U = E$, то из (60) следуют эквивалентные соотношения

$$\sum_{j=1}^n u_{ji} \bar{u}_{jh} = \delta_{ih} \quad (i, h = 1, 2, \dots, n). \quad (61)$$

Равенства (60) выражают собой «ортонормированность» строк, а равенства (61) — ортонормированность столбцов в матрице $U = \|u_{ik}\|_1^{n-1}$.

Унитарная матрица является матрицей коэффициентов некоторого унитарного преобразования (см. § 7).

Оператор P , осуществляющий ортогональное проектирование векторов унитарного пространства R на заданное подпространство S , является эрмитовым проекционным оператором.

Действительно, этот оператор является проекционным, т. е. $P^2 = P$ (см. гл. III, § 6). Далее, из ортогональности векторов $x_s = Px$ и $y - y_s = = (E - P)y$ ($x, y \in R$) следует:

$$0 = (Px, (E - P)y) = ((E - P^*)Px, y).$$

Отсюда в силу произвольности векторов x , y

$$(E - P^*)P = 0,$$

т. е. $P = P^*P$. Из этого равенства следует, что P — эрмитов оператор, так как $(P^*P)^* = P^*P$.

¹⁾ Таким образом, ортонормированность столбцов в матрице U является следствием ортонормированности строк и наоборот.

§ 10. Спектр нормальных, эрмитовых, унитарных операторов

Установим предварительно одно свойство перестановочных операторов, сформулировав его в виде леммы.

Лемма 1. *Перестановочные операторы A и B ($AB = BA$) всегда имеет общий собственный вектор.*

Доказательство. Пусть x есть собственный вектор оператора A : $Ax = \lambda x$, $x \neq 0$. Тогда в силу перестановочности операторов A и B

$$AB^k x = \lambda B^k x \quad (k = 0, 1, 2, \dots). \quad (62)$$

Пусть в ряду векторов

$$x, Bx, B^2x, \dots$$

первые p векторов линейно независимы, в то время как $(p+1)$ -й вектор $B^p x$ является уже линейной комбинацией предыдущих. Тогда подпространство $S \equiv [x, Bx, \dots, B^{p-1}x]$ будет инвариантно относительно B и потому в этом подпространстве S будет существовать собственный вектор y оператора B : $By = \mu y$, $y \neq 0$. С другой стороны, равенства (62) показывают, что векторы $x, Bx, \dots, B^{p-1}x$ являются собственными векторами оператора A , отвечающими одному и тому же характеристическому числу λ . Поэтому и любая линейная комбинация этих векторов, в частности вектор y , будет собственным вектором оператора A , отвечающим характеристическому числу λ . Таким образом, доказано существование общего собственного вектора операторов A и B .

Пусть A — произвольный нормальный оператор в n -мерном эрмитовом пространстве R . В этом случае операторы A и A^* перестановочны между собой и потому имеют общий собственный вектор x_1 . Тогда (см. § 8, 7°)

$$Ax_1 = \lambda_1 x_1, \quad A^* x_1 = \bar{\lambda}_1 x_1 \quad (x_1 \neq 0).$$

Обозначим через S_1 одномерное подпространство, содержащее вектор x_1 ($S_1 = [x_1]$), а через T_1 — ортогональное дополнение для S_1 в R :

$$R = S_1 + T_1, \quad S_1 \perp T_1.$$

Так как S_1 инвариантно относительно A и A^* , то (см. § 8, 5°) T_1 также инвариантно относительно этих операторов. Поэтому перестановочные операторы A и A^* имеют согласно лемме 1 общий собственный вектор x_2 в T_1 :

$$Ax_2 = \lambda_2 x_2, \quad A^* x_2 = \bar{\lambda}_2 x_2 \quad (x_2 \neq 0).$$

Очевидно, $x_1 \perp x_2$. Полагая $S_2 = [x_1, x_2]$ и

$$R = S_2 + T_2, \quad S_2 \perp T_2,$$

мы аналогичными соображениями установим существование в T_2 общего собственного вектора x_3 операторов A и A^* . Очевидно, $x_1 \perp x_3$ и $x_2 \perp x_3$. Продолжая этот процесс далее, мы получим n попарно ортогональных общих собственных векторов x_1, x_2, \dots, x_n операторов A и A^* :

$$\left. \begin{aligned} Ax_k &= \lambda_k x_k, \quad A^* x_k = \bar{\lambda}_k x_k \quad (x_k \neq 0), \\ (x_i x_k) &= 0 \quad \text{при } i \neq k \end{aligned} \right\} \quad (i, k = 1, 2, \dots, n). \quad (63)$$

Векторы x_1, x_2, \dots, x_n можно пронормировать. При этом равенства (63) сохраняются.

Таким образом, мы доказали, что нормальный оператор всегда имеет полную ортонормированную¹⁾ систему собственных векторов.

Так как из $\lambda_k = \lambda_l$ всегда следует $\bar{\lambda}_k = \bar{\lambda}_l$, то из равенств (63) вытекает:

1° Если оператор A нормален, то каждый собственный вектор оператора A является собственным вектором сопряженного оператора A^* , т. е. если оператор A нормален, то операторы A и A^* имеют одни и те же собственные векторы.

Пусть теперь, обратно, дано, что линейный оператор A имеет полную ортонормированную систему собственных векторов:

$$Ax_k = \lambda_k x_k, \quad (x_i x_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

Докажем, что в этом случае A является нормальным оператором. Действительно, положим:

$$y_l = A^* x_l - \bar{\lambda}_l x_l.$$

Тогда

$$\begin{aligned} (x_k y_l) &= (x_k, A^* x_l) - \lambda_l (x_k x_l) = (Ax_k, x_l) - \lambda_l (x_k x_l) = \\ &= (\lambda_k - \lambda_l) \delta_{kl} = 0 \quad (k, l = 1, 2, \dots, n). \end{aligned}$$

Отсюда следует:

$$y_l = A^* x_l - \bar{\lambda}_l x_l = 0 \quad (l = 1, 2, \dots, n),$$

т. е. имеют место все равенства (63).

Но тогда

$$AA^* x_k = \lambda_k \bar{\lambda}_k x_k \quad \text{и} \quad A^* Ax_k = \lambda_k \bar{\lambda}_k x_k \quad (k = 1, 2, \dots, n),$$

откуда

$$AA^* = A^* A.$$

Таким образом, мы получили следующую «внутреннюю» (спектральную) характеристику нормального оператора A (наряду с «внешней»: $AA^* = A^* A$):

Теорема 4. Линейный оператор тогда и только тогда является нормальным, когда этот оператор имеет полную ортонормированную систему собственных векторов.

В частности, нами доказано, что нормальный оператор всегда является оператором простой структуры.

Пусть A — нормальный оператор с характеристическими числами $\lambda_1, \lambda_2, \dots, \lambda_n$. По интерполяционной формуле Лагранжа определим два многочлена $p(\lambda)$ и $q(\lambda)$ из условий

$$p(\lambda_k) = \bar{\lambda}_k, \quad q(\bar{\lambda}_k) = \lambda_k \quad (k = 1, 2, \dots, n).$$

Тогда в силу (63)

$$A^* = p(A), \quad A = q(A^*), \tag{64}$$

т. е.

2° Для нормального оператора A каждый из операторов A и A^* представим в виде многочлена от другого из операторов; при этом эти два многочлена определяются заданием характеристических чисел оператора A .

¹⁾ Под полной ортонормированной системой векторов мы здесь и в дальнейшем понимаем ортонормированную систему из n векторов, где n — число измерений пространства.

Пусть S — инвариантное подпространство в R для нормального оператора A и $R = S + T$, $S \perp T$. Тогда согласно § 8, 5° (стр. 241) подпространство T инвариантно относительно A^* . Но $A = q(A^*)$, где $q(\lambda)$ — многочлен. Поэтому T инвариантно и относительно данного оператора A . Таким образом,

3° *Если S — инвариантное подпространство относительно нормального оператора T , а T — ортогональное дополнение к S , то и T является инвариантным подпространством для A .*

Остановимся теперь на спектре эрмитова оператора. Так как эрмитов оператор H является частным видом нормального оператора, то по доказанному он имеет полную ортонормированную систему собственных векторов:

$$Hx_k = \lambda_k x_k, \quad (x_k x_l) = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (65)$$

Из $H^* = H$ следует:

$$\bar{\lambda}_k = \lambda_k \quad (k = 1, 2, \dots, n), \quad (66)$$

т. е. все характеристические числа эрмитова оператора H вещественны.

Нетрудно видеть, что и обратно, нормальный оператор с вещественными характеристическими числами всегда эрмитов. В самом деле, из (65), (66) и

$$H^* x_k = \lambda_k x_k \quad (k = 1, 2, \dots, n)$$

следует:

$$H^* x_k = H x_k \quad (k = 1, 2, \dots, n),$$

т. е.

$$H^* = H.$$

Таким образом, мы получили следующую «внутреннюю» характеристику эрмитова оператора (наряду с «внешней»: $H^* = H$):

Теорема 5. *Линейный оператор H является эрмитовым тогда и только тогда, когда он имеет полную ортонормированную систему собственных векторов с вещественными характеристическими числами.*

Остановимся теперь на спектре унитарного оператора. Поскольку унитарный оператор U является нормальным, то он имеет полную ортонормированную систему собственных векторов

$$Ux_k = \lambda_k x_k, \quad (x_k x_l) = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (67)$$

При этом

$$U^* x_k = \bar{\lambda}_k x_k \quad (k = 1, 2, \dots, n). \quad (68)$$

Из $UU^* = E$ находим:

$$\lambda_k \bar{\lambda}_k = 1. \quad (69)$$

Обратно, из (67), (68), (69) следует: $UU^* = E$. Таким образом, среди нормальных операторов унитарный оператор выделяется тем, что у него все характеристические числа по модулю равны единице.

Мы получили следующую «внутреннюю» характеристику унитарного оператора (наряду с «внешней»: $UU^* = E$):

Теорема 6. *Линейный оператор тогда и только тогда является унитарным, когда он имеет полную ортонормированную систему собственных векторов с характеристическими числами, по модулю равными единице.*

Так как в ортонормированном базисе нормальная (эрмитова, унитарная) матрица соответственно определяет нормальный (эрмитов, унитарный) оператор, то получаем следующие предложения:

Теорема 4'. Матрица A является нормальной тогда и только тогда, когда она унитарно-подобна диагональной матрице:

$$A = U \|\lambda_i \delta_{ik}\|_1^n U^{-1} \quad (U^* = U^{-1}). \quad (70)$$

Теорема 5'. Матрица H является эрмитовой тогда и только тогда, когда она унитарно-подобна диагональной матрице с вещественными числами на диагонали:

$$H = U \|\lambda_i \delta_{ik}\|_1^n U^{-1} \quad (U^* = U^{-1}; \lambda_i = \bar{\lambda}_i; i = 1, 2, \dots, n). \quad (71)$$

Теорема 6'. Матрица U является унитарной тогда и только тогда, если она унитарно-подобна диагональной матрице с диагональными элементами, по модулю равными единице:

$$U = U_1 \|\lambda_i \delta_{ik}\|_1^n U_1^{-1} \quad (U_1^* = U_1^{-1}; |\lambda_i| = 1; i = 1, 2, \dots, n). \quad (72)$$

§ 11. Неотрицательные и положительно определенные эрмитовы операторы

Введем следующее определение.

Определение 9. Эрмитов оператор H называется *неотрицательным*, если для любого вектора x из R

$$(Hx, x) \geq 0,$$

и *положительно определенным*, если для любого вектора $x \neq 0$ из R

$$(Hx, x) > 0.$$

Если задать вектор x его координатами x_1, x_2, \dots, x_n в произвольном ортонормированном базисе, то (Hx, x) , как легко видеть, представится в виде эрмитовой формы от переменных x_1, x_2, \dots, x_n , причем неотрицательному (соответственно положительно определенному) оператору будет отвечать неотрицательная (соответственно положительно определенная) эрмитова форма (см. § 1).

Выберем ортонормированный базис x_1, x_2, \dots, x_n из собственных векторов оператора H :

$$Hx_k = \lambda_k x_k, \quad (x_k x_l) = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (73)$$

Тогда, полагая $x = \sum_{k=1}^n \xi_k x_k$, будем иметь:

$$(Hx, x) = \sum_{k=1}^n \lambda_k |\xi_k|^2 \quad (k = 1, 2, \dots, n).$$

Отсюда сразу следует «внутренняя» характеристика неотрицательного и положительно определенного оператора:

Теорема 7. Эрмитов оператор тогда и только тогда является *неотрицательным* (соответственно *положительно определенным*), если все его характеристические числа *неотрицательны* (соответственно *положительны*).

Из сказанного вытекает, что положительно определенный эрмитов оператор есть неособенный неотрицательный эрмитов оператор.

Пусть H — неотрицательный эрмитов оператор. Для него имеют место равенства (73) с $\lambda_k \geq 0$ ($k = 1, 2, \dots, n$). Положим $\varrho_k = \sqrt{\lambda_k} \geq 0$ ($k = 1, 2, \dots, n$).

$= 1, 2, \dots, n$) и определим линейный оператор \mathbf{F} равенствами

$$\mathbf{F}x_k = q_k x_k \quad (k = 1, 2, \dots, n). \quad (74)$$

Тогда \mathbf{F} будет также неотрицательным оператором, причем

$$\mathbf{F}^2 = \mathbf{H}. \quad (75)$$

Неотрицательный эрмитов оператор \mathbf{F} , связанный с \mathbf{H} равенством (75), будем называть арифметическим корнем квадратным из оператора \mathbf{H} и будем обозначать так:

$$\mathbf{F} = \sqrt{\mathbf{H}}.$$

Если \mathbf{H} — положительно определенный оператор, то и \mathbf{F} будет положительно определенным.

Определим интерполяционный многочлен Лагранжа $g(\lambda)$ равенствами

$$g(\lambda_k) = q_k (= \sqrt{\lambda_k}) \quad (k = 1, 2, \dots, n). \quad (76)$$

Тогда из (73), (74) и (76) следует:

$$\mathbf{F} = g(\mathbf{H}). \quad (77)$$

Последнее равенство показывает, что $\sqrt{\mathbf{H}}$ является многочленом от \mathbf{H} и однозначно определяется заданием неотрицательного эрмитова оператора \mathbf{H} (коэффициенты многочлена $g(\lambda)$ зависят от характеристических чисел оператора \mathbf{H}).

Примерами неотрицательных эрмитовых операторов являются операторы $\mathbf{A}\mathbf{A}^*$ и $\mathbf{A}^*\mathbf{A}$, где \mathbf{A} — произвольный линейный оператор в данном пространстве. Действительно, при произвольном векторе x

$$\begin{aligned} (\mathbf{A}\mathbf{A}^*x, x) &= (\mathbf{A}^*x, \mathbf{A}^*x) \geq 0, \\ (\mathbf{A}^*\mathbf{A}x, x) &= (\mathbf{A}x, \mathbf{A}x) \geq 0. \end{aligned}$$

Если оператор \mathbf{A} неособенный, то $\mathbf{A}\mathbf{A}^*$ и $\mathbf{A}^*\mathbf{A}$ — положительно определенные эрмитовы операторы.

Операторы $\sqrt{\mathbf{A}\mathbf{A}^*}$ и $\sqrt{\mathbf{A}^*\mathbf{A}}$ мы будем называть *левым* и *правым модулями* оператора \mathbf{A} .

У нормального оператора левый и правый модули равны между собой¹⁾.

§ 12. Полярное разложение линейного оператора в унитарном пространстве. Формулы Кэли

Докажем следующую теорему²⁾:

Теорема 8. Произвольный линейный оператор \mathbf{A} в унитарном пространстве всегда представим в виде

$$\mathbf{A} = \mathbf{H}\mathbf{U}, \quad (78)$$

$$\mathbf{A} = \mathbf{U}_1\mathbf{H}_1, \quad (79)$$

где \mathbf{H} , \mathbf{H}_1 — неотрицательные эрмитовы, а \mathbf{U} , \mathbf{U}_1 — унитарные операторы. Оператор \mathbf{A} нормален тогда и только тогда, когда

¹⁾ Относительно подробного исследования нормальных операторов см. [826]. В этой работе устанавливается необходимое и достаточное условие для того, чтобы произведение двух нормальных операторов было также нормальным оператором.

²⁾ См. [826], стр. 77.

в разложении (78) [или в (79)] множители \mathbf{H} и \mathbf{U} (соответственно \mathbf{H}_1 и \mathbf{U}_1) перестановочны между собой.

Доказательство. Из разложений (78) и (79) следует, что \mathbf{H} и \mathbf{H}_1 являются соответственно левым и правым модулями оператора A .

Действительно,

$$AA^* = \mathbf{H} \mathbf{U} \mathbf{U}^* \mathbf{H} = \mathbf{H}^2, \quad A^* A = \mathbf{H} \mathbf{U}_1^* \mathbf{U}_1 \mathbf{H}_1 = \mathbf{H}_1^2.$$

Заметим, что достаточно установить разложение (78), так как, применяя это разложение к оператору A^* , получим $A^* = \mathbf{H} \mathbf{U}$ и, следовательно,

$$A = \mathbf{U}^{-1} \mathbf{H},$$

т. е. разложение (79) для оператора A .

Установим сначала разложение (78) для частного случая, когда A — неособенный оператор ($|A| \neq 0$). Полагаем:

$$\mathbf{H} = \sqrt{AA^*} \quad (\text{при этом } |\mathbf{H}|^2 = |A|^2 \neq 0), \quad \mathbf{U} = \mathbf{H}^{-1} A$$

и проверяем унитарность оператора \mathbf{U} :

$$\mathbf{U} \mathbf{U}^* = \mathbf{H}^{-1} A A^* \mathbf{H}^{-1} = \mathbf{H}^{-1} \mathbf{H}^2 \mathbf{H} = \mathbf{E}.$$

Заметим, что в рассматриваемом случае в разложении (78) не только первый множитель \mathbf{H} , но и второй \mathbf{U} однозначно определяются заданием неособенного оператора A .

Рассмотрим теперь общий случай, когда A может быть и особенным оператором. Заметим прежде всего, что полная ортонормированная система собственных векторов оператора $A^* A$ всегда преобразуется оператором A снова в ортогональную же систему векторов. Действительно, пусть

$$A^* A x_k = q_k^2 x_k \quad [(x_k x_l) = \delta_{kl}, q_k \geq 0; k, l = 1, 2, \dots, n].$$

Тогда

$$(Ax_k, Ax_l) = (A^* A x_k, x_l) = q_k^2 \cdot (x_k x_l) = 0 \quad (k \neq l).$$

При этом

$$|Ax_k|^2 = (Ax_k, Ax_k) = q_k^2 \quad (k = 1, 2, \dots, n).$$

Поэтому существует такая ортонормированная система векторов $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$, что

$$Ax_k = q_k \mathbf{z}_k \quad [(z_k z_l) = \delta_{kl}; k, l = 1, 2, \dots, n]. \quad (80)$$

Определим линейные операторы \mathbf{H} и \mathbf{U} равенствами

$$\mathbf{U} x_k = z_k, \quad \mathbf{H} z_k = q_k z_k. \quad (81)$$

Из (80) и (81) находим:

$$A = \mathbf{H} \mathbf{U}.$$

При этом в силу (81) \mathbf{H} — неотрицательный эрмитов оператор, поскольку он имеет полную ортонормированную систему собственных векторов $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ и неотрицательные характеристические числа q_1, q_2, \dots, q_n , а \mathbf{U} — унитарный оператор, ибо он переводит ортонормированную систему векторов x_1, x_2, \dots, x_n снова в ортонормированную $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$.

Таким образом, можно считать доказанным, что для произвольного линейного оператора A имеют место разложения (78) и (79), причем эрмитовы множители \mathbf{H} и \mathbf{H}_1 всегда однозначно определяются заданием