

наоборот, последние $n-r$ строки матрицы C заполнить нулями, а последние $n-r$ столбцов матрицы B взять произвольными.

Доказательство. Возможность представления матрицы, удовлетворяющей условию (34), в виде произведения (35) была доказана выше [см. (33')].

Пусть теперь B и C — произвольные нижняя и верхняя треугольные матрицы, произведение которых равно A . Пользуясь формулой для миноров произведения двух матриц, найдем:

$$A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} = \sum_{a_1 < a_2 < \dots < a_k} B \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ a_1 & a_2 & \dots & a_{k-1} & a_k \end{pmatrix} C \begin{pmatrix} a_1 & a_2 & \dots & a_k \\ 1 & 2 & \dots & k \end{pmatrix} \quad (38)$$

$$(g = k, k+1, \dots, n; k = 1, 2, \dots, r).$$

Поскольку C — верхняя треугольная матрица, то первые k столбцов матрицы C содержат только один отличный от нуля минор k -го порядка $C \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}$. Поэтому равенство (38) может быть записано так:

$$A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} = B \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} C \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} = \\ = b_{11}b_{22} \dots b_{k-1, k-1}b_{gg}c_{11}c_{22} \dots c_{kk} \quad (39)$$

$$(g = k, k+1, \dots, n; k = 1, 2, \dots, r).$$

Положим сначала здесь $g = k$. Тогда получим:

$$b_{11}b_{22} \dots b_{kk}c_{11}c_{22} \dots c_{kk} = D_k \quad (k = 1, 2, \dots, r), \quad (40)$$

откуда уже вытекают соотношения (36).

Не нарушая равенства (35), мы можем в нем умножить матрицу B справа на произвольную неособенную диагональную матрицу $M = \|\mu_i \delta_{ik}\|_1^n$, одновременно умножая матрицу C слева на $M^{-1} = \|\mu_i^{-1} \delta_{ik}\|_1^n$. Это равносильно умножению столбцов матрицы B соответственно на $\mu_1, \mu_2, \dots, \mu_n$ и строк матрицы C на $\mu_1^{-1}, \mu_2^{-1}, \dots, \mu_n^{-1}$. Поэтому диагональным элементам $b_{11}, \dots, b_{rr}, c_{11}, \dots, c_{rr}$ можно придать любые значения, удовлетворяющие условиям (36).

Далее, из (39) и (40) находим:

$$b_{gg} = b_{kk} \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}} \quad (g = k, k+1, \dots, n; k = 1, 2, \dots, r),$$

т. е. первые формулы (37). Совершенно аналогично устанавливаются вторые формулы (37) для элементов матрицы C .

Обратим внимание на то, что при перемножении матриц B и C элементы b_{kg} последних $n-r$ столбцов матрицы B и элементы c_{gh} последних $n-r$ строк матрицы C перемножаются только между собой. Мы видели, что все элементы последних $n-r$ строк матрицы C можно выбрать равными нулю¹⁾. Тогда элементы последних $n-r$ столбцов матрицы B можно

¹⁾ Это следует из представления (33'). При этом диагональным элементам $b_{11}, \dots, b_{rr}, c_{11}, \dots, c_{rr}$, как было уже показано, можно придать любые значения, удовлетворяющие условиям (36) за счет надлежащих множителей $\mu_1, \mu_2, \dots, \mu_r$.

выбрать произвольными. Ясно, что произведение матриц B и C не изменится, если мы последние $n-r$ столбцов матрицы B возьмем нулевыми, а элементы последних $n-r$ строк матрицы C произвольными.

Теорема доказана.

Из доказанной теоремы вытекает ряд интересных следствий.

Следствие 1. Элементы первых r столбцов матрицы B и первых r строк матрицы C связаны с элементами матрицы A рекуррентными соотношениями:

$$\left. \begin{aligned} a_{ik} - \sum_{j=1}^{k-1} b_{ij} c_{jk} \\ c_{ik} - \sum_{j=1}^{i-1} b_{ij} c_{jk} \end{aligned} \right\} \quad \begin{aligned} (i \geq k; i = 1, 2, \dots, n; k = 1, 2, \dots, r), \\ (i \leq k; i = 1, 2, \dots, r; k = 1, 2, \dots, n). \end{aligned} \quad (41)$$

Соотношения (41) непосредственно следуют из матричного равенства (35) и ими удобно пользоваться для фактического вычисления элементов матриц B и C .

Следствие 2. Если $A = \|a_{ik}\|_1^n$ — неособенная матрица ($r=n$), удовлетворяющая условию (34), то в представлении (35) матрицы B и C определяются однозначно, как только диагональные элементы этих матриц выбраны в соответствии с условиями (36).

Следствие 3. Если $S = \|s_{ik}\|_1^n$ — симметрическая матрица ранга r и

$$D_k = S \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r),$$

то

$$S = BB',$$

где $B = \|b_{ik}\|_1^n$ — нижняя треугольная матрица, в которой

$$b_{ik} = \begin{cases} \frac{1}{\sqrt{D_k D_{k-1}}} A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} & (g = k, k+1, \dots, n; k = 1, 2, \dots, r), \\ 0 & (g = k, k+1, \dots, n; k = r+1, \dots, n). \end{cases} \quad (42)$$

2. Пусть в представлении (35) у матрицы B элементы последних $n-r$ столбцов равны нулю. Тогда можно положить:

$$B = F \cdot \begin{vmatrix} b_{11} & & 0 \\ \ddots & b_{rr} & \\ & 0 & \ddots & 0 \end{vmatrix}, \quad C = \begin{vmatrix} c_{11} & & 0 \\ \ddots & c_{rr} & \\ & 0 & \ddots & 0 \end{vmatrix} \cdot L, \quad (43)$$

где F — нижняя, а L — верхняя треугольная матрица; при этом первые r диагональных элементов у матриц F и L равны 1, а элементы последних $n-r$ столбцов матрицы F и последних $n-r$ строк матрицы L выбраны совершенно произвольно. Подставляя в (35) выражения (43) для B и C и используя равенства (36), придем к следующей теореме:

Теорема 2. Всякая матрица $A = \{a_{ik}\}_{1}^n$ ранга r , у которой

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r),$$

представима в виде произведения нижней треугольной матрицы F , диагональной D и верхней треугольной L :

$$A = FDL = \left\| \begin{array}{cccc} 1 & 0 & \dots & 0 \\ f_{21} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & 1 \end{array} \right\| \cdot \left\| \begin{array}{c} D_1 \\ \frac{D_2}{D_1} \\ \vdots \\ \frac{D_r}{D_{r-1}} \\ 0 \end{array} \right\| \cdot \left\| \begin{array}{cccc} 1 & l_{12} & \dots & l_{1n} \\ 0 & 1 & \dots & l_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{array} \right\|, \quad (44)$$

где

$$f_{gk} = \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}}, \quad l_{kg} = \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & g \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}} \quad (45)$$

$$(g = k+1, \dots, n; k = 1, 2, \dots, r),$$

а f_{gk}, l_{kg} произвольны при $g = k+1, \dots, n; k = r+1, \dots, n$.

3. Метод исключения Гаусса, будучи применен к матрице $A = \{a_{ik}\}_{1}^n$ ранга r , для которой $D_k \neq 0$ ($k = 1, 2, \dots, r$), дает нам две матрицы: нижнюю треугольную матрицу W с диагональными элементами 1 и верхнюю треугольную матрицу G , у которой первые r диагональных элементов равны $D_1, \frac{D_2}{D_1}, \dots, \frac{D_r}{D_{r-1}}$, а последние $n-r$ строки заполнены нулями. G — гауссова форма матрицы A , W — преобразующая матрица.

Для конкретного вычисления элементов матрицы W можно рекомендовать следующий прием.

Мы получим матрицу W , если к единичной матрице E применим все те преобразования (задаваемые матрицами W_1, \dots, W_N), которые мы в алгоритме Гаусса делали над матрицей A (в этом случае вместо произведения WA , равного G , мы будем иметь произведение WE , равное W). Поэтому к матрице A приписываем справа единичную матрицу E :

$$\left\| \begin{array}{cccccc} a_{11} & \dots & a_{1n} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} & 0 & \dots & 1 \end{array} \right\|. \quad (46)$$

Применяя к этой прямоугольной матрице все преобразования алгоритма Гаусса, получим прямоугольную матрицу, состоящую из двух квадратных матриц G и W :

$$(G, W).$$

Таким образом, применение алгоритма Гаусса к матрице (46) дает одновременно и матрицу G и матрицу W .

Если A — неособенная матрица, т. е. $|A| \neq 0$, то и $|G| \neq 0$. В этом случае из (33) следует $A^{-1} = G^{-1}W$. Поскольку матрицы G и W определены при помощи алгоритма Гаусса, то нахождение обратной матрицы A^{-1} сводится к определению G^{-1} и умножению G^{-1} на W .

Хотя нахождение обратной матрицы G^{-1} , после того как определена матрица G , не представляет затруднений, поскольку G — треугольная матрица, тем не менее можно избежать этой операции. Для этого наряду с матрицами G и W введем аналогичные матрицы G_1 и W_1 для транспонированной матрицы A' . Тогда $A' = W_1^{-1}G_1$, т. е.

$$A = G_1'W_1^{-1}. \quad (47)$$

Сопоставим между собой равенства (33') и (44):

$$A = W^{-1}G, \quad A = FDL.$$

Эти равенства могут быть рассматриваемы как два различных разложения вида (35); при этом мы произведение DL рассматриваем как второй множитель C . Поскольку первые r диагональных элементов у первых множителей одинаковы (они равны 1), то первые r столбцов у них совпадают. Тогда, поскольку последние $n - r$ столбцов матрицы F могут быть выбраны произвольными, то выберем их так, чтобы

$$F = W^{-1}. \quad (48)$$

С другой стороны, сопоставление равенств (47) и (44):

$$A = G_1'W_1^{-1}, \quad A = FDL$$

показывает, что можно так подобрать произвольные элементы в L , чтобы

$$L = W_1^{-1}. \quad (49)$$

Подставляя в (44) вместо F и L их выражения из (48) и (49), получим:

$$A = W^{-1}DW_1^{-1}. \quad (50)$$

Сопоставляя это равенство с равенствами (33') и (47), мы найдем:

$$G = DW_1^{-1}, \quad G_1 = W^{-1}D. \quad (51)$$

Введем в рассмотрение диагональную матрицу

$$\widehat{D} = \left\{ \frac{1}{D_1}, \frac{D_1}{D_2}, \dots, \frac{D_{r-1}}{D_r}, 0, \dots, 0 \right\}. \quad (52)$$

Тогда, поскольку

$$D = D\widehat{D}D,$$

то из (50) и (51) следует:

$$A = G_1'\widehat{D}G. \quad (53)$$

Формула (53) показывает, что разложение матрицы A на треугольные множители может быть получено применением алгоритма Гаусса к матрицам A и A' .

Пусть теперь A — неособенная матрица ($r = n$). Тогда $|D| \neq 0$, $\widehat{D} = D^{-1}$. Поэтому из (50) следует:

$$A^{-1} = W_1'\widehat{D}W. \quad (54)$$

Эта формула дает возможность эффективного вычисления обратной матрицы A^{-1} путем применения алгоритма Гаусса к прямоугольным матрицам

$$(A, E) (A', E).$$

В частном случае, когда вместо матрицы A возьмем симметрическую матрицу S , матрица G_1 совпадает с G , а матрица W_1 — с матрицей W , и потому формулы (53) и (54) принимают вид

$$S = G' \widehat{D} G, \quad (55)$$

$$S^{-1} = W' \widehat{D} W. \quad (56)$$

§ 5. Разбиение матрицы на блоки. Техника оперирования с блочными матрицами. Обобщенный алгоритм Гаусса

Часто приходится пользоваться матрицами, разбитыми на прямоугольные части — «клетки» или «блоки». Рассмотрению таких «блочных» матриц мы посвящаем настоящий параграф.

1. Пусть дана прямоугольная матрица

$$A = \|a_{ik}\| \quad (i=1, 2, \dots, m; k=1, 2, \dots, n). \quad (57)$$

При помощи горизонтальных и вертикальных линий рассечем матрицу A на прямоугольные блоки:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{st} \end{pmatrix} \left. \begin{array}{l} n_1 \\ n_2 \\ \vdots \\ n_t \end{array} \right\} m_1 \left. \begin{array}{l} \\ \\ \vdots \\ \end{array} \right\} m_2 \left. \begin{array}{l} \\ \\ \vdots \\ \end{array} \right\} m_s \quad (58)$$

Про матрицу (58) будем говорить, что она разбита на st блоков $A_{\alpha\beta}$ размером $m_\alpha \times n_\beta$ ($\alpha=1, 2, \dots, s; \beta=1, 2, \dots, t$) или что она представлена в виде блочной матрицы. Вместо (58) будем сокращенно писать:

$$A = (A_{\alpha\beta}) \quad (\alpha=1, 2, \dots, s; \beta=1, 2, \dots, t). \quad (59)$$

В случае $s=t$ будем пользоваться и такой записью:

$$A = (A_{\alpha\beta})_1^s. \quad (60)$$

Действия над блочными матрицами производятся по тем же формальным правилам, как и в случае, когда вместо блоков имеем числовые элементы. Пусть, например, даны две прямоугольные матрицы одинаковых размеров и с одинаковым разбиением на блоки:

$$A = (A_{\alpha\beta}), B = (B_{\alpha\beta}) \quad (\alpha=1, 2, \dots, s; \beta=1, 2, \dots, t). \quad (61)$$

Легко усмотреть, что

$$A + B = (A_{\alpha\beta} + B_{\alpha\beta}) \quad (\alpha=1, 2, \dots, s; \beta=1, 2, \dots, t). \quad (62)$$

Подробнее остановимся на умножении блочных матриц. Известно (см. гл. I, стр. 17), что при умножении двух прямоугольных матриц A и B длина строк в первом сомножителе A должна совпадать с высотой столбцов во втором сомножителе B . Для возможности «блочного»

умножения этих матриц мы дополнительного потребуем, чтобы при разбиении на блоки все горизонтальные размеры в первом сомножителе совпадали с соответствующими вертикальными размерами во втором:

$$A = \begin{pmatrix} n_1 & n_2 & \dots & n_t \\ \widehat{A}_{11} & \widehat{A}_{12} & \dots & \widehat{A}_{1t} \\ A_{21} & A_{22} & \dots & A_{2t} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{st} \end{pmatrix} \begin{array}{l} \\ \} m_1 \\ \} m_2 \\ \vdots \\ \} m_s \end{array}, \quad B = \begin{pmatrix} p_1 & p_2 & \dots & p_u \\ \widehat{B}_{11} & \widehat{B}_{12} & \dots & \widehat{B}_{1u} \\ B_{21} & B_{22} & \dots & B_{2u} \\ \dots & \dots & \dots & \dots \\ B_{t1} & B_{t2} & \dots & B_{tu} \end{pmatrix} \begin{array}{l} \\ \} n_1 \\ \} n_2 \\ \vdots \\ \} n_t \end{array}. \quad (63)$$

Тогда легко проверить, что

$$AB = C = (C_{\alpha\beta}), \quad \text{где } C_{\alpha\beta} = \sum_{\delta=1}^t A_{\alpha\delta} B_{\delta\beta} \quad \begin{cases} \alpha = 1, 2, \dots, s \\ \beta = 1, 2, \dots, u \end{cases}. \quad (64)$$

Отдельно отметим тот частный случай, когда одним из сомножителей является квазidiагональная матрица. Пусть A — квазidiагональная матрица, т. е. $s=t$ и $A_{\alpha\beta}=0$ при $\alpha \neq \beta$. В этом случае формула (64) нам дает:

$$C_{\alpha\beta} = A_{\alpha\alpha} B_{\alpha\beta} \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, u). \quad (65)$$

При умножении блочной матрицы слева на квазidiагональную матрицу строки блочной матрицы умножаются слева на соответствующие диагональные клетки квазidiагональной матрицы.

Пусть теперь B — квазidiагональная матрица, т. е. $t=u$ и $B_{\alpha\beta}=0$ при $\alpha \neq \beta$. Тогда из (64) получаем:

$$C_{\alpha\beta} = A_{\alpha\beta} B_{\beta\beta} \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, u). \quad (66)$$

При умножении блочной матрицы справа на квазidiагональную все столбцы блочной матрицы умножаются справа на соответствующие диагональные клетки квазidiагональной матрицы.

Заметим, что умножение квадратных блочных матриц одного и того же порядка всегда выполнимо, когда сомножители разбиты на одинаковые квадратные схемы блоков и в каждом из сомножителей на диагональных местах стоят квадратные матрицы.

Блочная матрица (58) называется *верхней (нижней) квазитреугольной*, если $s=t$ и все $A_{\alpha\beta}=0$ при $\alpha > \beta$ (соответственно все $A_{\alpha\beta}=0$ при $\alpha < \beta$). Частным случаем квазитреугольной матрицы является квазidiагональная матрица.

Из формулы (64) легко усмотреть, что

Произведение двух верхних (нижних) квазитреугольных матриц является снова верхней (нижней) квазитреугольной матрицей¹⁾; при этом диагональные блоки произведения получаются путем перемножения соответствующих диагональных блоков сомножителей.

Действительно, полагая в (64) $s=t$ и

$$A_{\alpha\beta} = 0, \quad B_{\alpha\beta} = 0 \quad \text{при } \alpha < \beta,$$

найдем:

$$\left. \begin{array}{l} C_{\alpha\beta} = 0 \quad \text{при } \alpha < \beta \\ C_{\alpha\alpha} = A_{\alpha\alpha} B_{\alpha\alpha} \end{array} \right\} \quad (\alpha, \beta = 1, 2, \dots, s).$$

¹⁾ При этом предполагается, что блочное умножение выполнимо.

Аналогично разбирается случай нижних квазитреугольных матриц.

Отметим правило вычисления определителя квазитреугольной матрицы. Это правило можно получить, исходя из разложения Лапласа.

Если A — квазитреугольная (в частности, квазидиагональная) матрица с квадратными диагональными блоками, то определитель этой матрицы равен произведению определителей диагональных блоков:

$$|A| = |A_{11}| |A_{22}| \dots |A_{ss}|. \quad (67)$$

2. Пусть дана блочная матрица

$$A = \begin{pmatrix} n_1 & n_2 & \dots & n_t \\ \widehat{A}_{11} & \widehat{A}_{12} & \dots & \widehat{A}_{1t} \\ A_{21} & A_{22} & \dots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \dots & A_{st} \end{pmatrix} \left. \begin{array}{l} \\ \\ \} m_1 \\ \\ \} m_2 \\ \vdots \\ \} m_s \end{array} \right\}. \quad (68)$$

Прибавим к α -й блочной строке β -ю строку, предварительно помноженную слева на прямоугольную матрицу X размером $m_\alpha \times m_\beta$. Получим блочную матрицу

$$B = \begin{pmatrix} A_{11} & \dots & A_{1t} \\ \vdots & \ddots & \vdots \\ A_{\alpha 1} + XA_{\beta 1} & \dots & A_{\alpha t} + XA_{\beta t} \\ \vdots & \ddots & \vdots \\ A_{\beta 1} & \dots & A_{\beta t} \\ \vdots & \ddots & \vdots \\ A_{s1} & \dots & A_{st} \end{pmatrix}. \quad (69)$$

Введем вспомогательную квадратную матрицу V , представленную в виде следующей квадратной схемы блоков:

$$V = \begin{pmatrix} m_1 \dots m_\alpha \dots m_\beta \dots m_s \\ \widehat{E} \dots \widehat{0} \dots \widehat{0} \dots \widehat{0} \\ \vdots & \vdots & \vdots & \vdots \\ 0 \dots E \dots X \dots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 \dots 0 \dots E \dots 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 \dots 0 \dots 0 \dots E \end{pmatrix} \left. \begin{array}{l} \\ \\ \} m_1 \\ \\ \} m_\alpha \\ \\ \} m_\beta \\ \vdots \\ \} m_s \end{array} \right\}. \quad (70)$$

В диагональных клетках матрицы V стоят единичные матрицы, порядки которых равны соответственно m_1, m_2, \dots, m_s ; все недиагональные блоки матрицы V равны нулю, за исключением блока X , стоящего на пересечении α -й блочной строки с β -м блочным столбцом.

Нетрудно видеть, что

$$VA = B. \quad (71)$$

Отсюда, поскольку V — неособенная матрица, для рангов матриц A и B имеем¹⁾:

$$r_A = r_B. \quad (72)$$

¹⁾ См. стр. 22.

В частном случае, когда A — квадратная матрица, из (71) имеем:

$$|V| |A| = |B|. \quad (73)$$

Но определитель квазитреугольной матрицы V равен 1:

$$|V| = 1. \quad (74)$$

Следовательно,

$$|A| = |B|. \quad (75)$$

К таким же выводам можно прийти, если к какому-либо столбцу матрицы (67) прибавлять другой столбец, предварительно помноженный справа на прямоугольную матрицу X надлежащих размеров.

Полученные результаты могут быть сформулированы в виде следующей теоремы.

Теорема 3. Если в блочной матрице A к α -й блочной строке (столбцу) прибавить β -ю блочную строку (столбец), предварительно помноженную слева (справа) на прямоугольную матрицу X соответствующих размеров, то при этом преобразование не изменяется ранг матрицы A , а также в случае, когда A — квадратная матрица, и определитель матрицы A .

3. Рассмотрим теперь тот частный случай, когда в матрице A диагональный блок A_{11} — квадратная и притом неособенная матрица ($|A_{11}| \neq 0$).

К α -й строке матрицы A прибавим первую строку, помноженную слева на $-A_{\alpha 1} A_{11}^{-1}$ ($\alpha = 2, \dots, s$). Тогда получим матрицу

$$B_1 = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1t} \\ 0 & A_{22}^{(1)} & \dots & A_{2t}^{(1)} \\ \dots & \dots & \dots & \dots \\ 0 & A_{s2}^{(1)} & \dots & A_{st}^{(1)} \end{pmatrix}, \quad (76)$$

где

$$A_{\alpha\beta}^{(1)} = -A_{\alpha 1} A_{11}^{-1} A_{1\beta} + A_{\alpha\beta} \quad (\alpha = 2, \dots, s; \beta = 2, \dots, t). \quad (77)$$

Если $A_{22}^{(1)}$ — квадратная неособенная матрица, то этот процесс можно продолжить. Мы приходим, таким образом, к обобщенному алгоритму Гаусса.

Пусть A — квадратная матрица. Тогда

$$|A| = |B_1| = |A_{11}| \begin{vmatrix} A_{22}^{(1)} & \dots & A_{2t}^{(1)} \\ \dots & \dots & \dots \\ A_{s2}^{(1)} & \dots & A_{st}^{(1)} \end{vmatrix}. \quad (78)$$

Формула (78) сводит вычисление определителя $|A|$, состоящего из st блоков, к вычислению определителя меньшего порядка, состоящего из $(s-1)(t-1)$ блоков¹⁾.

Рассмотрим определитель Δ , разбитый на четыре блока:

$$\Delta = \begin{vmatrix} A & B \\ C & D \end{vmatrix}, \quad (79)$$

где A и D — квадратные матрицы.

¹⁾ Если $A_{22}^{(1)}$ — квадратная матрица и $|A_{22}^{(1)}| \neq 0$, то к полученному определителю из $(s-1)(t-1)$ блоков мы можем снова применить такое же преобразование и т. д.

Пусть $|A| \neq 0$. Тогда вычтем из второй строки первую, предварительно помноженную слева на $-CA^{-1}$. Получим:

$$\Delta = \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix} = |A| |D - CA^{-1}B|. \quad (\text{I})$$

Точно так же, если $|D| \neq 0$, то мы вычтем в Δ из первой строки вторую, предварительно помноженную слева на $-BD^{-1}$. Получим:

$$\Delta = \begin{vmatrix} A - BD^{-1}C & 0 \\ C & D \end{vmatrix} = |A - BD^{-1}C| |D|. \quad (\text{II})$$

В частном случае, когда все четыре матрицы A, B, C, D — квадратные (одного и того же порядка n), из (I) и (II) следуют формулы Шура, сводящие вычисление определителя $2n$ -го порядка к вычислению определителя n -го порядка:

$$\Delta = |AD - ACA^{-1}B| \quad (|A| \neq 0), \quad (\text{Ia})$$

$$\Delta = |AD - BD^{-1}CD| \quad (|D| \neq 0). \quad (\text{IIa})$$

Если матрицы A и C перестановочны между собой, то из (Ia) следует:

$$\Delta = |AD - CB| \quad (\text{при условии } AC = CA). \quad (\text{Ib})$$

Точно так же, если C и D перестановочны между собой, то

$$\Delta = |AD - BC| \quad (\text{при условии } CD = DC). \quad (\text{IIb})$$

Формула (Ib) была получена в предположении $|A| \neq 0$, а формула (IIb) при условии $|D| \neq 0$. Однако, исходя из соображений непрерывности, эти ограничения можно отбросить.

Из формул (I)–(IIb) можно получить еще шесть формул, поменяв местами в правых частях A и D и одновременно B и C .

Пример.

$$\Delta = \begin{vmatrix} 1 & 0 & b_1 & b_2 \\ 0 & 1 & b_3 & b_4 \\ c_1 & c_2 & d_1 & d_2 \\ c_3 & c_4 & d_3 & d_4 \end{vmatrix}.$$

По формуле (Ib)

$$\Delta = \begin{vmatrix} d_1 - c_1b_1 - c_2b_3 & d_2 - c_1b_2 - c_2b_4 \\ d_3 - c_3b_1 - c_4b_3 & d_4 - c_3b_2 - c_4b_4 \end{vmatrix}.$$

4. Установим формулу Фробениуса¹⁾ для обращения блочной матрицы. Пусть неособенная квадратная матрица M ($|M| \neq 0$) разбита на блоки

$$M = \begin{matrix} n & q \\ q & \{ \end{matrix} \left(\begin{array}{cc} \tilde{A} & \tilde{B} \\ C & D \end{array} \right), \quad (80)$$

и пусть A — также неособенная квадратная матрица ($|A| \neq 0$). Требуется определить M^{-1} .

¹⁾ См., например, [44].

Применим к матрице M обобщенный алгоритм Гаусса. Из второй блочной строки вычтем первую, предварительно помноженную слева на $-CA^{-1}$. Эта операция равносильна умножению матрицы M слева на матрицу¹⁾ $\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}$, где $X = -CA^{-1}$. Поэтому

$$\begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix} M = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}. \quad (81)$$

Введем обозначение

$$H = D - CA^{-1}B$$

и заметим, что из равенства (81) следует:

$$|M| = |A| |H|. \quad (82)$$

Поэтому, поскольку $|M| \neq 0$, то и $|H| \neq 0$ ²⁾. Переходя к обратным матрицам в равенстве (81), получим

$$M^{-1} \begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}^{-1} = \begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1}. \quad (83)$$

Обратную матрицу для матрицы $\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}$ будем искать в виде $\begin{pmatrix} A^{-1} & U \\ 0 & H^{-1} \end{pmatrix}$.

Тогда из равенства

$$\begin{pmatrix} A & B \\ 0 & H \end{pmatrix} \begin{pmatrix} A^{-1} & U \\ 0 & H^{-1} \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}$$

находим, что $U = -A^{-1}BH^{-1}$. Таким образом,

$$\begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BH^{-1} \\ 0 & H^{-1} \end{pmatrix}. \quad (84)$$

Но тогда из равенства (83) находим

$$M^{-1} = \begin{pmatrix} A & B \\ 0 & H \end{pmatrix}^{-1} \begin{pmatrix} E & 0 \\ -CA^{-1} & B \end{pmatrix} = \begin{pmatrix} A^{-1} & -A^{-1}BH^{-1} \\ 0 & H^{-1} \end{pmatrix} \begin{pmatrix} E & 0 \\ -CA^{-1} & E \end{pmatrix}. \quad (85)$$

Выполнив умножение блочных матриц в правой части равенства (85), приходим к формуле Фробениуса

$$M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BH^{-1}CA^{-1} & -A^{-1}BH^{-1} \\ -H^{-1}CA^{-1} & H^{-1} \end{pmatrix}, \quad (86)$$

где

$$H = D - CA^{-1}B. \quad (87)$$

Формула Фробениуса (86) сводит обращение матрицы порядка $n+q$ к обращению двух матриц порядка n и q и к операциям сложения и умножения матриц с размерами $n \times n$, $q \times q$, $n \times q$ и $q \times n$.

1) В первой блочной строке буква E обозначает единичную матрицу n -го порядка, а во второй—единичную матрицу порядка q .

2) Нам нет необходимости заранее устанавливать неособенность матрицы M , так как это свойство матрицы M следует из того, что $|A| \neq 0$ и $|H| \neq 0$. Если бы оказалось, что $|H|=0$, то тогда и $|M|=0$, и в этом случае не существует обратной матрицы M^{-1} .

Если предположить, что $|D| \neq 0$ (вместо $|A| \neq 0$) и поменять ролями матрицы A и D , то можно получить другой вид формулы Фробениуса:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} K^{-1} & -K^{-1}BD^{-1} \\ -D^{-1}CK^{-1} & D^{-1} + D^{-1}CK^{-1}BD^{-1} \end{pmatrix}, \quad (88)$$

где

$$K = A - BD^{-1}C. \quad (89)$$

Пример. Требуется найти элементы обратной матрицы для матрицы

$$M = \begin{vmatrix} 1 & -1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & 1 & 2 \end{vmatrix}.$$

Полагаем

$$A = \begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix}, \quad B = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}, \quad C = \begin{vmatrix} 0 & 2 \\ 2 & 0 \end{vmatrix}, \quad D = \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix}.$$

Находим последовательно

$$A^{-1} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix}, \quad CA^{-1} = \begin{vmatrix} 0 & 2 \\ 2 & 0 \end{vmatrix} \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 2 \\ 4 & 2 \end{vmatrix},$$

$$H = D - CA^{-1}B = D - \begin{vmatrix} 2 & 2 \\ 4 & 2 \end{vmatrix} \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} - \begin{vmatrix} -2 & 2 \\ -2 & 4 \end{vmatrix} = \begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix},$$

$$H^{-1} = \begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix}, \quad A^{-1}B = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = \begin{vmatrix} -1 & 2 \\ -1 & 1 \end{vmatrix},$$

$$A^{-1}BH^{-1} = \begin{vmatrix} -1 & 2 \\ -1 & 1 \end{vmatrix} \begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix} = \begin{vmatrix} 4 & -3 \\ 1 & -1 \end{vmatrix},$$

$$A^{-1}BH^{-1}CA^{-1} = \begin{vmatrix} 4 & -3 \\ 1 & -1 \end{vmatrix} \begin{vmatrix} 2 & 2 \\ 4 & 2 \end{vmatrix} = \begin{vmatrix} -4 & 2 \\ -2 & 0 \end{vmatrix},$$

$$A^{-1} + A^{-1}BH^{-1}CA^{-1} = \begin{vmatrix} -2 & 3 \\ -1 & 1 \end{vmatrix},$$

$$H^{-1}CA^{-1} = \begin{vmatrix} 2 & -1 \\ 3 & -2 \end{vmatrix} \begin{vmatrix} 2 & 2 \\ 4 & 2 \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ -2 & 2 \end{vmatrix}.$$

Поэтому по формуле (86) находим:

$$M^{-1} = \begin{vmatrix} -2 & 3 & -4 & 3 \\ -1 & 1 & -1 & 1 \\ 0 & -2 & 2 & -1 \\ 2 & -2 & 3 & -2 \end{vmatrix}.$$

5. Из теоремы 3 вытекает также

Теорема 4. Если прямоугольная матрица R представлена в блочном виде

$$R = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (90)$$

где A — квадратная неособенная матрица порядка n ($|A| \neq 0$), то ранг

матрицы R равен n в том и только в том случае, когда

$$D = CA^{-1}B. \quad (91)$$

Доказательство. Вычтем из второй блочной строки матрицы R первую, предварительно помноженную слева на CA^{-1} . Тогда получим матрицу

$$T = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}. \quad (92)$$

Матрицы R и T согласно теореме 3 имеют один и тот же ранг. Ранг же матрицы T совпадает с рангом матрицы A (т. е. с n) тогда и только тогда, когда $D - CA^{-1}B = 0$, т. е. когда имеет место (91). Теорема доказана.

Из теоремы 4 вытекает алгоритм построения обратной матрицы A^{-1} и вообще произведения $CA^{-1}B$, где B , C —прямоугольные матрицы размером $n \times p$, $q \times n$ ¹⁾.

Приведем при помощи алгоритма Гаусса²⁾ матрицу

$$\begin{pmatrix} A & B \\ -C & 0 \end{pmatrix} \quad (|A| \neq 0) \quad (93)$$

к виду

$$\begin{pmatrix} G & B_1 \\ 0 & X \end{pmatrix}. \quad (94)$$

Докажем, что

$$X = CA^{-1}B. \quad (95)$$

В самом деле, то же преобразование, которое было применено к матрице (93), приведет матрицу

$$\begin{pmatrix} A & B \\ -C & -CA^{-1}B \end{pmatrix} \quad (96)$$

к виду

$$\begin{pmatrix} G & B_1 \\ 0 & X - CA^{-1}B \end{pmatrix}. \quad (97)$$

Согласно теореме 4 матрица (96) имеет ранг n (n —порядок матрицы A). Но тогда и матрица (97) должна иметь ранг n . Отсюда $X - CA^{-1}B = 0$, т. е. имеет место (95).

В частности, если $B = y$, где y —столбцевая матрица, и $C = E$, то

$$X = A^{-1}y.$$

Следовательно, применяя алгоритм Гаусса к матрице

$$\begin{pmatrix} A & y \\ -E & 0 \end{pmatrix},$$

1) См. [88].

2) Здесь мы применяем к матрице (93) не весь алгоритм Гаусса, а только его первые n этапов, где n —порядок матрицы A . Это можно сделать, если выполняются условия (15) при $p=n$. Если же эти условия не выполнены, то, поскольку $|A| \neq 0$, мы можем так перенумеровать первые n строк (или первые n столбцов) матрицы (93), чтобы n этапов алгоритма Гаусса оказались выполнимыми. Такой видоизмененный алгоритм Гаусса иногда применяют и при выполнении условий (15) для $p=n$.

мы получаем решение системы уравнений

$$Ax = y.$$

Далее, если в (93) положить $B = C = E$, то после применения алгоритма Гаусса к матрице

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

получим:

$$\begin{pmatrix} G, & W \\ 0, & X \end{pmatrix},$$

где

$$X = A^{-1}.$$

Проиллюстрируем этот способ нахождения A^{-1} на следующем примере.

Пример. Пусть

$$A = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 2 \end{vmatrix}.$$

Требуется вычислить A^{-1} .

Применяем несколько видоизмененный метод исключения¹⁾ к матрице

$$\left| \begin{array}{cccccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{array} \right|.$$

Ко всем строкам прибавляем вторую строку с некоторым множителем и добиваемся того, чтобы все элементы первого столбца, кроме второго элемента, равнялись нулю. После этого ко всем строкам, кроме второй, прибавляем третью строку с некоторым множителем и достигаем того, чтобы во втором столбце все элементы, кроме второго и третьего, были равны нулю. После этого к последним трем строкам прибавляем первую строку с некоторым множителем и получаем матрицу вида

$$\left| \begin{array}{cccccc} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ 0 & 0 & 0 & -2 & -1 & 2 \\ 0 & 0 & 0 & 4 & 1 & -3 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{array} \right|.$$

Поэтому

$$A^{-1} = \begin{vmatrix} -2 & -1 & 2 \\ 4 & 1 & -3 \\ 1 & 1 & -1 \end{vmatrix}.$$

6. Разбиение матрицы на блоки может быть использовано также для нахождения псевдообратной матрицы (см. гл. I, § 5). Пусть снова

¹⁾ См. сноску²⁾ на стр. 62.

прямоугольная матрица R представлена в виде

$$R = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (98)$$

где A — неособенная квадратная матрица ($|A| \neq 0$) и $r_A = r_R$. Тогда согласно теореме 4 справедливо равенство $D = CA^{-1}B$, и потому

$$R = \begin{pmatrix} A \\ C \end{pmatrix} A^{-1}(AB). \quad (99)$$

Так как эта формула является результатом двух последовательных скелетных разложений (см. гл. I, § 5):

$$R = \begin{pmatrix} A \\ C \end{pmatrix} (E \ A^{-1}B), \quad (E \ A^{-1}B) = A^{-1}(AB),$$

то

$$R^+ = (AB)^+ (A^{-1})^+ \begin{pmatrix} A \\ C \end{pmatrix}^+ = (AB)^+ A \begin{pmatrix} A \\ C \end{pmatrix}^+. \quad (100)$$

Применяя формулы (43) и (44) на стр. 34, окончательно найдем:

$$R^+ = \begin{pmatrix} A^* \\ B^* \end{pmatrix} (AA^* + BB^*)^{-1} A (A^*A + C^*C)^{-1} (A^*C^*). \quad (101)$$

Формула (101) дает явное выражение для псевдообратной матрицы R^+ через блоки A , B , C .

П р и м е р.

$$R = \left\| \begin{array}{cc|cc} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ \hline 0 & 1 & -1 & 1 \end{array} \right\| = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Здесь $r_R = 2$ и

$$A = \left\| \begin{array}{cc} 1 & -1 \\ -1 & 2 \end{array} \right\|, \quad B = \left\| \begin{array}{cc} 2 & 0 \\ -3 & 1 \end{array} \right\|, \quad C = \left\| \begin{array}{cc} 0 & 1 \end{array} \right\|, \quad D = \left\| \begin{array}{cc} -1 & 1 \end{array} \right\|,$$

$$(AA^* + BB^*)^{-1} = \left(3 \left\| \begin{array}{cc} 2 & -3 \\ -3 & 5 \end{array} \right\| \right)^{-1} = \frac{1}{3} \left\| \begin{array}{cc} 5 & 3 \\ 3 & 2 \end{array} \right\|,$$

$$(A^*A + C^*C)^{-1} = \left\| \begin{array}{cc} 2 & -3 \\ -3 & 6 \end{array} \right\|^{-1} = \frac{1}{3} \left\| \begin{array}{cc} 6 & 3 \\ 3 & 2 \end{array} \right\|,$$

$$(AA^* + BB^*)^{-1} A (A^*A + C^*C)^{-1} = \frac{1}{9} \left\| \begin{array}{cc} 15 & 8 \\ 9 & 5 \end{array} \right\|.$$

Тогда по формуле (101)

$$R^+ = \frac{1}{9} \left\| \begin{array}{cc|cc} 1 & -1 & 15 & 8 \\ -1 & 2 & 9 & 5 \\ 2 & -3 & 15 & 8 \\ 0 & 1 & 9 & 5 \end{array} \right\| \left\| \begin{array}{ccc} 1 & 0 & \frac{1}{3} \\ \frac{1}{9} & \frac{1}{9} & \frac{2}{9} \\ \frac{2}{9} & -\frac{1}{9} & \frac{1}{9} \\ \frac{4}{9} & \frac{1}{9} & \frac{5}{9} \end{array} \right\|.$$

ГЛАВА III

ЛИНЕЙНЫЕ ОПЕРАТОРЫ В n -МЕРНОМ ВЕКТОРНОМ ПРОСТРАНСТВЕ

Матрицы составляют основной аналитический аппарат для изучения линейных операций в n -мерном пространстве. В свою очередь изучение этих операций дает возможность разбить все матрицы на классы и выявить важные свойства, присущие всем матрицам одного и того же класса.

В настоящей главе излагаются наиболее простые свойства линейных операторов в n -мерном пространстве. Дальнейшее исследование линейных операторов в n -мерном пространстве будет продолжено в главах VII и IX.

§ 1. Векторное пространство

1. Пусть дана некоторая совокупность R произвольных элементов x, y, z, \dots , в которой определены две операции: операция «сложения» и операция «умножения на число из поля K »¹⁾. Допустим, что эти операции всегда выполнимы и однозначны в R и для любых элементов x, y, z из R и чисел α, β из K :

$$1^\circ \quad x + y = y + x.$$

$$2^\circ \quad (x + y) + z = x + (y + z).$$

3° Существует такой элемент 0 в R , что произведение числа 0 на любой элемент x из R равно элементу 0 :

$$0 \cdot x = 0.$$

$$4^\circ \quad 1 \cdot x = x.$$

$$5^\circ \quad \alpha(\beta x) = (\alpha\beta)x.$$

$$6^\circ \quad (\alpha + \beta)x = \alpha x + \beta x.$$

$$7^\circ \quad \alpha(x + y) = \alpha x + \alpha y.$$

Определение 1. Совокупность элементов R , в которой всегда выполнимы и однозначны две операции: «сложение» элементов и «умножение элемента из R на число из K », причем эти операторы удовлетворяют постулатам $1^\circ - 7^\circ$, мы будем называть *векторным пространством* (над полем K), а сами элементы — *векторами*²⁾.

1) Эти операции будем отмечать обычными знаками «+» и «·», причем последний знак иногда не ставится, а только подразумевается.

2) Нетрудно видеть, что из свойств $1^\circ - 7^\circ$ следуют все обычные свойства операций сложения и умножения на число. Так, например, при любом x из R : $x + 0 = 1 \cdot x + 0 \cdot x = (1 + 0)x = 1 \cdot x = x$, $x + (-x) = 0$, где $-x = (-1) \cdot x$ и т. п.

Определение 2. Векторы x, y, \dots, u из R называются линейно зависимыми, если существуют такие числа $\alpha, \beta, \dots, \delta$ из K , не равные одновременно нулю, что

$$\alpha x + \beta y + \dots + \delta u = 0. \quad (1)$$

В случае, если не существует подобной линейной зависимости, векторы x, y, \dots, u называются линейно независимыми.

Если векторы x, y, \dots, u линейно зависимы, то один из векторов может быть представлен в виде линейной комбинации остальных с коэффициентами из поля K . Так, например, если в (1) $\alpha \neq 0$, то

$$x = -\frac{\beta}{\alpha} y - \dots - \frac{\delta}{\alpha} u.$$

Определение 3. Пространство R называется конечномерным, а число n — числом измерений этого пространства, если в R существует n линейно независимых векторов, в то время как любые $n+1$ векторов из R линейно зависимы. Если же в пространстве можно найти линейно независимую систему из любого числа векторов, то пространство называется бесконечномерным.

В настоящей книге в основном изучаются конечномерные пространства.

Определение 4. Система из n линейно независимых заданных в определенном порядке векторов e_1, e_2, \dots, e_n в n -мерном пространстве называется базисом этого пространства.

Пример 1. Совокупность обычных векторов (направленных геометрических отрезков) является трехмерным векторным пространством. Часть этого пространства, состоящая из векторов, параллельных некоторой плоскости, является двумерным пространством, а все векторы, параллельные некоторой прямой, образуют одномерное векторное пространство.

Пример 2. Столбец из n чисел поля K : $x = (x_1, x_2, \dots, x_n)$ назовем вектором (n — фиксированное число). Основные операции определим как операции над столбцевыми матрицами:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \\ a(x_1, x_2, \dots, x_n) = (ax_1, ax_2, \dots, ax_n).$$

Элементом нуль будет столбец $(0, 0, \dots, 0)$. Легко проверить, что все постулаты 1° — 7° выполняются. Эти векторы образуют n -мерное пространство. В качестве базиса этого пространства можно, например, взять столбцы единичной матрицы n -го порядка:

$$(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1).$$

Пространство, рассмотренное в этом примере, часто называют численным n -мерным пространством.

Пример 3. Совокупность бесконечных последовательностей $(x_1, x_2, \dots, x_n, \dots)$, в которой операции определены естественным образом, т. е.

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n, \dots) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n, \dots), \\ a(x_1, x_2, \dots, x_n, \dots) = (ax_1, ax_2, \dots, ax_n, \dots),$$

представляет собой бесконечномерное пространство.

Пример 4. Совокупность многочленов $a_0 + a_1 t + \dots + a_{n-1} t^{n-1}$ степени $< n$ с коэффициентами из K представляет собой n -мерное векторное пространство¹⁾. Базисом такого пространства является, например, система степеней $t^0, t, t^2, \dots, t^{n-1}$.

Все же такие многочлены (без ограничения степени) образуют бесконечно-мерное пространство.

Пример 5. Все функции, определенные в замкнутом интервале $[a, b]$, образуют бесконечномерное пространство.

¹⁾ В качестве основных операций берутся обычное сложение многочленов и умножение многочлена на число.

3. Пусть векторы e_1, e_2, \dots, e_n образуют базис n -мерного векторного пространства R , а x — произвольный вектор этого пространства. Тогда векторы x, e_1, e_2, \dots, e_n линейно зависимы (ибо число их равно $n+1$):

$$a_0x + a_1e_1 + a_2e_2 + \dots + a_ne_n = 0,$$

где по крайней мере одно из чисел a_0, a_1, \dots, a_n отлично от нуля. Однако в данном случае $a_0 \neq 0$, так как векторы e_1, e_2, \dots, e_n не могут быть связаны линейной зависимостью. Поэтому

$$x = x_1e_1 + x_2e_2 + \dots + x_ne_n, \quad (2)$$

где $x_i = -\frac{a_i}{a_0}$ ($i = 1, 2, \dots, n$).

Заметим, что числа x_1, x_2, \dots, x_n однозначно определяются заданием вектора x и базиса e_1, e_2, \dots, e_n . В самом деле, если наряду с (2) имеется другое разложение для вектора x :

$$x = x'_1e_1 + x'_2e_2 + \dots + x'_ne_n, \quad (2')$$

то, вычитая почленно (2) из (2'), получим:

$$(x'_1 - x_1)e_1 + (x'_2 - x_2)e_2 + \dots + (x'_n - x_n)e_n = 0,$$

откуда в силу линейной независимости векторов базиса следует

$$x'_1 - x_1 = x'_2 - x_2 = \dots = x'_n - x_n = 0,$$

т. е.

$$x'_1 = x_1, x'_2 = x_2, \dots, x'_n = x_n. \quad (3)$$

Числа x_1, x_2, \dots, x_n называются координатами вектора x в базисе e_1, e_2, \dots, e_n .

Если

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

то

$$x + y = \sum_{i=1}^n (x_i + y_i) e_i, \quad ax = \sum_{i=1}^n a x_i e_i,$$

т. е. координаты суммы векторов получаются почленным сложением соответствующих координат слагаемых векторов и при умножении вектора на число a все координаты вектора умножаются на это число.

4. Пусть векторы

$$x_k = \sum_{i=1}^n x_{ik} e_i \quad (k = 1, 2, \dots, m)$$

линейно зависимы, т. е.

$$\sum_{k=1}^m c_k x_k = 0, \quad (4)$$

где по крайней мере одно из чисел c_1, c_2, \dots, c_m не равно нулю.

Если вектор равен нулю, то все его координаты равны нулю. Поэтому векторное равенство (4) эквивалентно следующей системе скалярных равенств:

$$\left. \begin{aligned} c_1x_{11} + c_2x_{12} + \dots + c_mx_{1m} &= 0, \\ c_1x_{21} + c_2x_{22} + \dots + c_mx_{2m} &= 0, \\ \dots & \\ c_1x_{n1} + c_2x_{n2} + \dots + c_mx_{nm} &= 0. \end{aligned} \right\} \quad (4')$$

Эта система однородных линейных уравнений относительно c_1, c_2, \dots, c_m , как известно, имеет ненулевое решение тогда и только тогда, когда ранг матрицы коэффициентов меньше числа неизвестных, т. е. меньше m . Поэтому равенство этого ранга числу m является необходимым и достаточным условием для линейной независимости векторов x_1, x_2, \dots, x_m .

Таким образом, имеет место следующая

Теорема 1. Для того чтобы векторы x_1, x_2, \dots, x_m были линейно независимы, необходимо и достаточно, чтобы ранг r матрицы, составленной из координат этих векторов в произвольном базисе,

$$X = \begin{vmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{vmatrix}$$

был равен m , т. е. числу векторов.

Замечание. Линейная независимость векторов x_1, x_2, \dots, x_m означает линейную независимость столбцов матрицы X , поскольку в k -м столбце стоят координаты вектора x_k ($k = 1, 2, \dots, m$). Поэтому, согласно теореме, если в прямоугольной матрице столбцы линейно независимы, то ранг матрицы равен числу столбцов. Отсюда следует, что в произвольной прямоугольной матрице максимальное число линейно независимых столбцов равно рангу матрицы. Кроме того, если мы транспонируем матрицу, т. е. строки делаем столбцами (и наоборот), то ранг матрицы при этом, очевидно, не меняется. Поэтому в прямоугольной матрице всегда число линейно независимых столбцов равно числу линейно независимых строк и равно рангу матрицы.

5. Если в n -мерном пространстве выбран базис e_1, e_2, \dots, e_n , то каждому вектору x однозначно отвечает столбец $x = (x_1, x_2, \dots, x_n)$, где x_1, x_2, \dots, x_n — координаты вектора x в данном базисе. Таким образом, задание базиса устанавливает взаимно однозначное соответствие между векторами произвольного n -мерного векторного пространства R и векторами n -мерного численного пространства R' , рассмотренного в примере 2. При этом сумме векторов из R отвечает сумма соответствующих векторов из R' . Аналогичное имеет место и для произведения вектора на число a из K . Другими словами, произвольное n -мерное векторное пространство изоморфно численному n -мерному пространству и, следовательно, все векторные пространства одного и того же числа измерений n над одним и тем же числовым полем K изоморфны между собой. Это означает, что с точностью до изоморфизма существует только одно n -мерное векторное пространство при заданном числовом поле.

Может возникнуть вопрос, зачем мы ввели «абстрактное» n -мерное пространство, если с точностью до изоморфизма оно совпадает с n -мерным численным пространством. Действительно, можно было бы определить вектор как систему n чисел, заданных в определенном порядке, и установить операцию над этими векторами, как это было сделано в примере 2. Но при этом смешались бы воедино свойства вектора, не зависящие от выбора базиса, со свойствами специального базиса. Например, равенство нулю всех координат вектора есть свойство самого вектора; оно не зависит от выбора базиса. Равенство между собой всех координат вектора не есть свойство самого вектора, ибо при изменении базиса оно исчезает. Аксиоматическое определение векторного пространства непосредственно выделяет свойства векторов, не зависящие от выбора базиса.

6. Если некоторая совокупность векторов R' , составляющая часть R , обладает тем свойством, что сумма любых двух векторов из R' и произведение любого вектора из R' на число $a \in K$ всегда принадлежат R' , то такое многообразие R' само является векторным пространством, подпространством в R .

- Если даны два подпространства R' и R'' в R и известно, что
 1° R' и R'' не имеют общих векторов, кроме нуля, и
 2° любой вектор x из R представляется в виде суммы

$$x = x' + x'' \quad (x' \in R', \quad x'' \in R''), \quad (5)$$

то мы будем говорить, что пространство R расщепляется на два подпространства R' и R'' и будем писать:

$$R = R' + R''. \quad (6)$$

Заметим, что условие 1° означает единственность представления (5). В самом деле, если бы для некоторого вектора x мы имели два разных представления в виде суммы слагаемых из R' и R'' , представление (5) и представление

$$x = \tilde{x}' + \tilde{x}'' \quad (\tilde{x}' \in R', \quad \tilde{x}'' \in R''), \quad (7)$$

то, вычитая почленно (7) из (5), мы получили бы:

$$x' - \tilde{x}' = \tilde{x}'' - x'',$$

т. е. равенство между отличными от нуля векторами $x' - \tilde{x}' \in R'$ и $\tilde{x}'' - x'' \in R''$, что невозможно в силу 1°.

Таким образом, условие 1° можно заменить требованием единственности представления (5). В таком виде определение расщепления непосредственно распространяется на любое число слагаемых подпространств.

Пусть

$$R = R' + R''$$

и $e'_1, e'_2, \dots, e'_{n'}$ и $e''_1, e''_2, \dots, e''_{n''}$ — базисы соответственно в R' и R'' . Тогда читатель без труда докажет, что все эти $n' + n''$ векторов линейно независимы и образуют базис в R , т. е. что из базисов слагаемых подпространств составляется базис всего пространства. В частности, отсюда будет следовать, что $n = n' + n''$.

Пример 1. Пусть в пространстве трех измерений даны три непараллельных одной и той же плоскости направления. Так как любой вектор в пространстве можно разложить на составляющие по этим трем направлениям, притом единственным образом, то

$$R = R' + R'' + R''',$$

где R — совокупность всех векторов нашего пространства, R' — совокупность всех векторов, параллельных первому направлению, R'' — второму, R''' — третьему. В данном случае $n=3$, $n'=n''=n'''=1$.

Пример 2. Пусть в пространстве трех измерений даны плоскость и пересекающая ее прямая. Тогда

$$R = R' + R'',$$

где R — совокупность всех векторов нашего пространства, R' — совокупность всех векторов, параллельных заданной плоскости, и R'' — совокупность всех векторов, параллельных заданной прямой. В этом примере $n=3$, $n'=2$, $n''=1$.

Задание базиса e_1, e_2, \dots, e_n в пространстве R по существу означает некоторое расщепление всего пространства R на n одномерных подпространств.

§ 2. Линейный оператор, отображающий n -мерное пространство в m -мерное

1. Рассмотрим линейное преобразование

$$\left. \begin{array}{l} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ \dots \dots \dots \dots \dots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n, \end{array} \right\} \quad (8)$$

коэффициенты которого принадлежат числовому полю K , и два векторных пространства над этим полем: n -мерное R и m -мерное S . Выберем в R некоторый базис e_1, e_2, \dots, e_n и в S некоторый базис g_1, g_2, \dots, g_m .

Тогда преобразование (8) относит каждому вектору $x = \sum_{i=1}^n x_i e_i$ из R некоторый вектор $y = \sum_{k=1}^m y_k g_k$ из S , т. е. преобразование (8) определяет некоторый оператор A , относящий вектору x вектор y : $y = Ax$. Нетрудно видеть, что этот оператор A обладает свойством линейности, которое мы сформулируем так:

Определение 5. Оператор A , отображающий R в S , т. е. относящий каждому вектору x из R некоторый вектор $y = Ax$ из S , называется *линейным*, если для любых x, x_1 из R и a из K

$$A(x + x_1) = Ax + Ax_1, \quad A(ax) = aAx. \quad (9)$$

Таким образом, преобразование (8) при заданных базисах в R и S определяет некоторый линейный оператор, отображающий R в S .

Покажем теперь обратное, т. е. что для произвольного линейного оператора A , отображающего R в S , и произвольных базисов e_1, e_2, \dots, e_n в R и g_1, g_2, \dots, g_m в S существует такая прямоугольная матрица с элементами из поля K

$$\left| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right|, \quad (10)$$

что составленное при помощи этой матрицы линейное преобразование (1) выражает координаты преобразованного вектора $y = Ax$ через координаты исходного вектора x .

Действительно, применим оператор A к базисному вектору e_k и координаты полученного вектора Ae_k в базисе g_1, g_2, \dots, g_m обозначим через $a_{1k}, a_{2k}, \dots, a_{mk}$ ($k = 1, 2, \dots, n$):

$$Ae_k = \sum_{i=1}^m a_{ik} g_i \quad (k = 1, 2, \dots, n). \quad (11)$$

Помножая обе части равенства (11) на x_k и суммируя в пределах от 1 до n , получим:

$$\sum_{k=1}^n x_k Ae_k = \sum_{i=1}^m \left(\sum_{k=1}^n a_{ik} x_k \right) g_i,$$

откуда

$$\mathbf{y} = \mathbf{Ax} = \mathbf{A} \left(\sum_{k=1}^n x_k \mathbf{e}_k \right) = \sum_{k=1}^n x_k \mathbf{Ae}_k = \sum_{i=1}^m y_i \mathbf{g}_i,$$

где

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i = 1, 2, \dots, m),$$

что и требовалось установить.

Таким образом, при заданных базисах в \mathbf{R} и \mathbf{S} каждому линейному оператору \mathbf{A} , отображающему \mathbf{R} в \mathbf{S} , отвечает некоторая прямоугольная матрица (10) с размерами $m \times n$ и, наоборот, каждой такой матрице отвечает некоторый линейный оператор, отображающий \mathbf{R} в \mathbf{S} .

При этом в матрице A , отвечающей оператору \mathbf{A} , k -й столбец состоит из последовательных координат вектора \mathbf{Ae}_k ($k = 1, 2, \dots, n$).

Обозначим через $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ столбцы координат векторов \mathbf{x} и \mathbf{y} . Тогда векторному равенству

$$\mathbf{y} = \mathbf{Ax}$$

соответствует матричное равенство

$$y = Ax,$$

которое является матричной записью преобразования (8).

Пример.

Рассмотрим совокупность всех многочленов от t степени $\leq n-1$ с коэффициентами из числового поля K . Эта совокупность представляет собой некоторое n -мерное векторное пространство \mathbf{R}_n (см. пример 4 на стр. 66). Точно так же многочлены от t степени $\leq n-2$ с коэффициентами из K образуют пространство \mathbf{R}_{n-1} . Оператор дифференцирования $\frac{d}{dt}$ относит каждому многочлену из \mathbf{R}_n некоторый многочлен из \mathbf{R}_{n-1} . Таким образом, этот оператор отображает \mathbf{R}_n в \mathbf{R}_{n-1} . Оператор дифференцирования является линейным оператором, так как

$$\frac{d}{dt} [\Phi(t) + \Psi(t)] = \frac{d\Phi(t)}{dt} + \frac{d\Psi(t)}{dt}, \quad \frac{d}{dt} [a\Phi(t)] = a \frac{d\Phi(t)}{dt}.$$

В пространствах \mathbf{R}_n и \mathbf{R}_{n-1} выберем базисы из степеней t :

$$t^0 = 1, t, \dots, t^{n-1} \text{ и } t^0 = 1, t, \dots, t^{n-2}.$$

Пользуясь формулой (11), построим прямоугольную матрицу размером $(n-1) \times n$, соответствующую оператору дифференцирования $\frac{d}{dt}$ в этих базисах:

$$\begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-1 \end{vmatrix}.$$

§ 3. Сложение и умножение линейных операторов

1. Пусть даны два линейных оператора \mathbf{A} и \mathbf{B} , отображающие \mathbf{R} в \mathbf{S} , и соответствующие им матрицы

$$A = \|a_{ik}\|, \quad B = \|b_{ik}\| \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n).$$

Определение 6. Суммой операторов A и B называется оператор C , определяемый равенством

$$Cx = Ax + Bx \quad (x \in R) \text{¹⁾.} \quad (12)$$

На основе этого определения легко проверяется, что сумма $C = A + B$ линейных операторов A и B есть также линейный оператор.

Далее,

$$Ce_k = Ae_k + Be_k = \sum_{k=1}^n (a_{ik} + b_{ik}) e_k.$$

Отсюда следует, что оператору C отвечает матрица $C = \|c_{ik}\|$, где $c_{ik} = a_{ik} + b_{ik}$ ($i = 1, 2, \dots, m$; $k = 1, 2, \dots, n$), т. е. оператору C отвечает матрица

$$C = A + B. \quad (13)$$

К этому же выводу можно прийти из рассмотрения матричного равенства

$$Cx = Ax + Bx \quad (14)$$

(x — столбец координат вектора x), соответствующего векторному равенству (12). Поскольку x — произвольный столбец, то из (14) следует (13).

2. Пусть даны три векторных пространства R , S и T соответственно q , n и m измерений и два линейных оператора A и B , из которых B отображает R в S , а A отображает S в T ; в символической записи:

$$T \xleftarrow{A} S \xleftarrow{B} R.$$

Определение 7. Произведением операторов A и B называется оператор C , для которого при любом x из R

$$Cx = A(Bx) \quad (x \in R). \quad (15)$$

Оператор C отображает R в T :

$$T \xleftarrow{C=A B} R.$$

Из линейности операторов A и B вытекает линейность оператора C . Выберем в пространствах R , S , T произвольные базисы и обозначим через A , B , C матрицы, соответствующие операторам A , B , C при этом выборе базисов. Тогда векторным равенствам

$$z = Ay, \quad y = Bx, \quad z = Cx \quad (16)$$

будут соответствовать матричные равенства

$$z = Ay, \quad y = Bx, \quad z = Cx,$$

где x , y , z — столбцы координат векторов x , y , z . Отсюда находим:

$$Cx = A(Bx) = (AB)x,$$

и в силу произвольности столбца x

$$C = AB. \quad (17)$$

Таким образом, произведению $C = AB$ операторов A и B отвечает матрица $C = \|c_{ij}\|$ ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, q$), равная произведению матриц A и B .

¹⁾ $x \in R$ означает, что элемент x принадлежит совокупности R . Предполагается, что равенство (12) имеет место при любом x из R .

Предоставляем читателю самому доказать, что оператору

$$C = aA \quad (a \in K)$$

отвечает матрица

$$C = aA.$$

Таким образом, мы видим, что в главе I действия над матрицами были определены так, что сумме линейных операторов $A + B$, произведениям AB и aA отвечают соответственно матрицы $A + B$, AB и aA , где A и B — матрицы, соответствующие операторам A и B , а a — число из K .

§ 4. Преобразование координат

Рассмотрим в n -мерном векторном пространстве два базиса: e_1, e_2, \dots, e_n («старый» базис) и $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$ («новый» базис).

Взаимное расположение векторов базиса определится, если задать координаты векторов одного базиса относительно другого.

Мы положим:

$$\left. \begin{aligned} \tilde{e}_1 &= t_{11} e_1 + t_{21} e_2 + \dots + t_{n1} e_n, \\ \tilde{e}_2 &= t_{12} e_1 + t_{22} e_2 + \dots + t_{n2} e_n, \\ &\dots \dots \dots \dots \dots \\ \tilde{e}_n &= t_{1n} e_1 + t_{2n} e_2 + \dots + t_{nn} e_n, \end{aligned} \right\} \quad (18)$$

или в сокращенной записи:

$$\tilde{e}_k = \sum_{i=1}^n t_{ik} e_i \quad (k = 1, 2, \dots, n). \quad (18')$$

Установим связь между координатами одного и того же вектора в различных базисах.

Пусть x_1, x_2, \dots, x_n и $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ — координаты вектора x соответственно в «старом» и «новом» базисах:

$$x = \sum_{i=1}^n x_i e_i = \sum_{k=1}^n \tilde{x}_k \tilde{e}_k. \quad (19)$$

Подставим в (19) вместо векторов \tilde{e}_k их выражения из (18). Получим:

$$x = \sum_{k=1}^n \tilde{x}_k \sum_{i=1}^n t_{ik} e_i = \sum_{i=1}^n \left(\sum_{k=1}^n t_{ik} \tilde{x}_k \right) e_i.$$

Сопоставляя это равенство с (19) и учитывая, что координаты вектора однозначно определяются заданием вектора и базиса, находим:

$$x_i = \sum_{k=1}^n t_{ik} \tilde{x}_k \quad (i = 1, 2, \dots, n), \quad (20)$$

¹⁾ То есть оператору, для которого $Cx = aAx$ ($x \in R$).

или в подробной записи:

$$\left. \begin{array}{l} x_1 = t_{11}\tilde{x}_1 + t_{12}\tilde{x}_2 + \dots + t_{1n}\tilde{x}_n, \\ x_2 = t_{21}\tilde{x}_1 + t_{22}\tilde{x}_2 + \dots + t_{2n}\tilde{x}_n, \\ \dots \dots \dots \dots \dots \dots \\ x_n = t_{n1}\tilde{x}_1 + t_{n2}\tilde{x}_2 + \dots + t_{nn}\tilde{x}_n. \end{array} \right\} \quad (21)$$

Формулы (21) определяют преобразование координат вектора при переходе от одного базиса к другому. Они выражают «старые» координаты через «новые». Матрица

$$T = \|t_{ik}\|_1^n \quad (22)$$

называется *матрицей преобразования координат* или *преобразующей матрицей*. В ней k -й столбец состоит из «старых» координат k -го «нового» базисного вектора. В этом можно убедиться из формулы (18) или непосредственно из формул (21), положив в последних $\tilde{x}_k=1$, $\tilde{x}_i=0$ при $i \neq k$.

Заметим, что матрица T неособенная, т. е.

$$|T| \neq 0. \quad (23)$$

Действительно, положив в (21) $x_1=x_2=\dots=x_n=0$, получим систему n линейных однородных уравнений с n неизвестными $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ и с определителем $|T|$. Эта система может иметь только нулевое решение $\tilde{x}_1=0, \tilde{x}_2=0, \dots, \tilde{x}_n=0$, так как в противном случае из (19) следовала бы линейная зависимость между векторами $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$. Поэтому $|T| \neq 0^1$.

Введем в рассмотрение столбцевые матрицы $x=(x_1, x_2, \dots, x_n)$ и $\tilde{x}=(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$.

Тогда формулы преобразования координат (21) могут быть записаны в виде следующего матричного равенства:

$$x = T\tilde{x}. \quad (24)$$

Помножая снова обе части этого равенства на T^{-1} , найдем выражение для обратного преобразования

$$\tilde{x} = T^{-1}x. \quad (25)$$

§ 5. Эквивалентные матрицы. Ранг оператора. Неравенства Сильвестра

1. Пусть даны два векторных пространства R и S , соответственно n и m измерений над числовым полем K , и линейный оператор A , отображающий R в S . В настоящем параграфе мы выясним, как меняется матрица A , соответствующая данному линейному оператору A , при изменении базисов в R и S .

Выберем в R и S произвольные базисы e_1, e_2, \dots, e_n и g_1, g_2, \dots, g_m . В этих базисах оператору A будет соответствовать матрица $A = \|a_{ik}\|$ ($i=1, 2, \dots, m; k=1, 2, \dots, n$). Векторному равенству

$$y = Ax \quad (26)$$

¹⁾ Неравенство (23) вытекает и из теоремы 1 (стр. 68), поскольку элементами матрицы T являются «старые» координаты линейно независимых векторов $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$.

соответствует матричное равенство

$$y = Ax, \quad (27)$$

где x и y — координатные столбцы для векторов \mathbf{x} и \mathbf{y} в базисах e_1, e_2, \dots, e_n и g_1, g_2, \dots, g_m .

Выберем теперь в R и S другие базисы $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$ и $\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_m$. В новых базисах вместо x, y, A будем иметь: $\tilde{x}, \tilde{y}, \tilde{A}$. При этом

$$\tilde{y} = \tilde{A} \tilde{x}. \quad (28)$$

Обозначим через Q и N неособенные квадратные матрицы соответственно порядков n и m , осуществляющие преобразование координат в пространствах R и S при переходе от старых базисов к новым (см. § 4):

$$x = Q\tilde{x}, \quad y = N\tilde{y}. \quad (29)$$

Тогда из (27) и (29) получаем:

$$\tilde{y} = N^{-1}y = N^{-1}Ax = N^{-1}AQ\tilde{x}. \quad (30)$$

Полагая $P = N^{-1}$, мы из (28) и (30) находим:

$$\tilde{A} = PAQ. \quad (31)$$

Определение 8. Две прямоугольные матрицы A и B одинаковых размеров называются *эквивалентными*, если существуют две неособенные квадратные матрицы P и Q такие¹⁾, что

$$B = PAQ. \quad (32)$$

Из (31) следует, что две матрицы, соответствующие одному и тому же линейному оператору A при различном выборе базисов в R и S , всегда эквивалентны между собой. Нетрудно видеть, что и обратно, если матрица A отвечает оператору A при некоторых базисах в R и S , а матрица B эквивалентна матрице A , то она отвечает тому же линейному оператору при некоторых других базисах в R и S .

Таким образом, каждому линейному оператору, отображающему R в S , соответствует класс эквивалентных между собой матриц с элементами из поля K .

2. Следующая теорема устанавливает критерий эквивалентности двух матриц:

Теорема 2. Для того чтобы две прямоугольные матрицы одинаковых размеров были эквивалентны, необходимо и достаточно, чтобы эти матрицы имели один и тот же ранг.

Доказательство. Условие необходимо. При умножении прямоугольной матрицы на какую-либо неособенную квадратную матрицу (слева или справа) ранг исходной прямоугольной матрицы не может измениться (см. гл. I, стр. 27). Поэтому из (32) следует

$$r_A = r_B.$$

Условие достаточно. Пусть A — прямоугольная матрица размера $m \times n$. Она определяет линейный оператор A , отображающий пространство R с базисом e_1, e_2, \dots, e_n в пространство S с базисом

¹⁾ Если матрицы A и B имеют размеры $m \times n$, то в (32) квадратная матрица P имеет порядок m , а квадратная матрица Q — порядок n . Если элементы эквивалентных матриц A и B принадлежат некоторому числовому полю, то матрицы P и Q могут быть выбраны так, чтобы их элементы принадлежали тому же числовому полю.

g_1, g_2, \dots, g_m . Обозначим через r число линейно независимых векторов среди векторов Ae_1, Ae_2, \dots, Ae_n . Не нарушая общности, можем считать, что линейно независимыми являются векторы Ae_1, Ae_2, \dots, Ae_r ¹⁾, а остальные, Ae_{r+1}, \dots, Ae_n , выражаются линейно через них:

$$Ae_k = \sum_{j=1}^r c_{kj} Ae_j \quad (k = r+1, \dots, n). \quad (33)$$

Определим новый базис в R следующим образом:

$$\tilde{e}_i = \begin{cases} e_i & (i = 1, 2, \dots, r), \\ e_i - \sum_{j=1}^r c_{ij} e_j & (i = r+1, \dots, n). \end{cases} \quad (34)$$

Тогда в силу (33)

$$A\tilde{e}_k = 0 \quad (k = r+1, \dots, n). \quad (35)$$

Далее положим:

$$A\tilde{e}_j = \tilde{g}_j \quad (j = 1, 2, \dots, r). \quad (36)$$

Векторы $\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_r$ линейно независимы. Дополним их некоторыми векторами $\tilde{g}_{r+1}, \dots, \tilde{g}_m$ до базиса $\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_m$ в S .

Тогда матрица, отвечающая тому же оператору A в новых базисах $\tilde{e}_1, \dots, \tilde{e}_n; \tilde{g}_1, \dots, \tilde{g}_m$, согласно (35) и (36) будет иметь вид

$$I_r = \left[\begin{array}{cccccc} & & & & & & r \\ & \overbrace{1 \ 0 \ \dots \ 0} & & 0 \dots 0 & & & \\ & 0 & 1 \dots 0 & 0 \dots 0 & & & \\ & 0 & 0 \dots 1 & 0 \dots 0 & & & \\ & 0 & 0 \dots 0 & 0 \dots 0 & & & \\ & \vdots & \ddots & \ddots & \ddots & & \\ & 0 & 0 \dots 0 & 0 \dots 0 & & & \end{array} \right]. \quad (37)$$

В матрице I_r вдоль главной диагонали сверху вниз идут r единиц; все остальные элементы матрицы I_r равны нулю. Так как матрицы A и I_r соответствуют одному и тому же оператору A , то они эквивалентны между собой. По доказанному эквивалентные матрицы имеют один и тот же ранг. Поэтому ранг исходной матрицы A равен r .

Мы показали, что произвольная прямоугольная матрица ранга r эквивалентна «канонической» матрице I_r . Но матрица I_r полностью определяется заданием размеров $m \times n$ и числа r . Поэтому все прямоугольные матрицы данных размеров $m \times n$ и данного ранга r эквивалентны одной и той же матрице I_r и, следовательно, эквивалентны между собой. Теорема доказана.

3. Пусть дан линейный оператор A , отображающий n -мерное пространство R в m -мерное S . Совокупность векторов вида Ax , где $x \in R$, образует векторное пространство²⁾. Это пространство мы будем обозначать

1) Этого можно достичь надлежащей нумерацией базисных векторов e_1, e_2, \dots, e_n .

2) Совокупность векторов вида Ax ($x \in R$) удовлетворяет постулатам 1°—7° § 1, поскольку сумма двух векторов вида Ax ($x \in R$) и произведение такого вектора на число снова дают вектор такого вида.

через $\mathbf{A}\mathbf{R}$; оно составляет часть пространства \mathbf{S} или, как говорят, является подпространством в пространстве \mathbf{S} .

Наряду с подпространством $\mathbf{A}\mathbf{R}$ в \mathbf{S} рассмотрим совокупность всех векторов $x \in \mathbf{R}$, удовлетворяющих уравнению

$$\mathbf{A}x = 0. \quad (38)$$

Эти векторы также образуют подпространство в \mathbf{R} ; это подпространство мы обозначим через N_A .

Определение 9. Если линейный оператор A отображает \mathbf{R} в \mathbf{S} , то число измерений r пространства $\mathbf{A}\mathbf{R}$ называется рангом оператора A ¹⁾, а число измерений d пространства N_A , состоящего из всех векторов $x \in \mathbf{R}$, удовлетворяющих условию (38), — дефектом оператора A .

Среди всех эквивалентных прямоугольных матриц, задающих данный оператор A в различных базисах, имеется каноническая матрица I_r [см. (37)]. Обозначим через $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$ и $\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_m$ соответствующие ей базисы в \mathbf{R} и \mathbf{S} . Тогда

$$\mathbf{A}\tilde{e}_1 = \tilde{g}_1, \dots, \mathbf{A}\tilde{e}_r = \tilde{g}_r, \quad \mathbf{A}\tilde{e}_{r+1} = \dots = \mathbf{A}\tilde{e}_n = 0.$$

Из определения $\mathbf{A}\mathbf{R}$ и N_A следует, что векторы $\tilde{g}_1, \dots, \tilde{g}_r$ образуют базис в $\mathbf{A}\mathbf{R}$, а векторы $\tilde{e}_{r+1}, \dots, \tilde{e}_n$ составляют базис в N_A . Отсюда вытекает, что r — ранг оператора A и

$$d = n - r. \quad (39)$$

Если A — произвольная матрица, соответствующая оператору A , то она эквивалентна I_r и, следовательно, имеет тот же ранг r . Таким образом, ранг оператора A совпадает с рангом прямоугольной матрицы A

$$A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix},$$

определенной оператором A в некоторых базисах $e_1, e_2, \dots, e_n \in \mathbf{R}$ и $g_1, g_2, \dots, g_m \in \mathbf{S}$.

В столбцах матрицы A стоят координаты векторов $\mathbf{A}e_1, \mathbf{A}e_2, \dots, \mathbf{A}e_n$. Так как из $x = \sum_{i=1}^n x_i e_i$ следует $\mathbf{A}x = \sum_{i=1}^n x_i \mathbf{A}e_i$, то ранг оператора A , т. е. число измерений $\mathbf{A}\mathbf{R}$, равняется максимальному числу линейно независимых векторов среди $\mathbf{A}e_1, \mathbf{A}e_2, \dots, \mathbf{A}e_n$. Таким образом, ранг матрицы совпадает с числом линейно независимых столбцов матрицы. Поскольку при транспонировании строки матрицы делаются столбцами, а ранг не меняется, то число линейно независимых строк матрицы также равно рангу матрицы²⁾.

1) Число измерений пространства $\mathbf{A}\mathbf{R}$ всегда \leq числа измерений пространства \mathbf{R} , т. е. $r \leq n$. Это следует из того, что равенство $x = \sum_{i=1}^n x_i e_i$ (e_1, e_2, \dots, e_n — базис в \mathbf{R})

влечет равенство $\mathbf{A}x = \sum_{i=1}^n x_i \mathbf{A}e_i$.

2) К этим же выводам мы пришли в § 1 из других соображений (см. стр. 68).

4. Пусть даны два линейных оператора A , B и их произведение $C = AB$.

Пусть оператор B отображает R в S , а оператор A отображает S в T . Тогда оператор C отображает R в T :

$$T \xleftarrow{A} S \xleftarrow{B} R, \quad T \xleftarrow{C} R.$$

Введем матрицы A , B , C , соответствующие операторам A , B , C при некотором выборе базисов в R , S и T . Тогда операторному равенству $C = AB$ будет соответствовать матричное равенство $C = AB$.

Обозначим через r_A , r_B , r_C ранги операторов A , B , C , или, что то же, ранги матриц A , B , C . Эти числа определяют число измерений подпространств AS , BR , $A(BR)$. Поскольку $BR \subset S$, то $A(BR) \subset AS^1$. Кроме того, число измерений $A(BR)$ не может превосходить числа измерений BR^2 . Поэтому

$$r_C \leq r_A, \quad r_C \leq r_B.$$

Эти неравенства были нами получены в главе I, § 2 из формулы для миноров произведения двух матриц.

Рассмотрим оператор A как оператор, отображающий BR в T . Тогда ранг этого оператора будет равен числу измерений пространства $A(BR)$, т. е. r_C . Поэтому, применив формулу (39), получим:

$$r_C = r_B - d_1, \quad (40)$$

где d_1 — максимальное число линейно независимых векторов из BR , удовлетворяющих уравнению

$$Ax = 0. \quad (41)$$

Но все решения этого уравнения, принадлежащие S , образуют подпространство d измерений, где

$$d = n - r_A \quad (42)$$

— дефект оператора A , отображающего S в T . Поскольку $BR \subset S$, то $d_1 \leq d$. (43)

Из (40), (42) и (43) находим:

$$r_A + r_B - n \leq r_C.$$

Таким образом, мы получили следующие неравенства Сильвестра для ранга произведения двух прямоугольных матриц A и B с размерами $m \times n$ и $n \times q$:

$$r_A + r_B - n \leq r_{AB} \leq r_A, r_B. \quad (44)$$

Если матричное уравнение $AXB = C$, где размеры прямоугольных матриц A , X , B — $m \times n$, $n \times p$, $p \times q$, имеет решение X (см. стр. 23), то из неравенств Сильвестра легко следует:

$$r_C \leq r_x \leq n + p - r_A - r_B.$$

Можно доказать, что если уравнение $AXB = C$ имеет какое-либо решение, то оно имеет решение любого ранга r , заключенного между числами r_C и $n + p - r_A - r_B$.

¹⁾ $R \subset S$ означает, что совокупность R составляет часть совокупности S .

²⁾ См. сноску ²⁾ на стр. 22.

§ 6. Линейные операторы, отображающие n -мерное пространство само в себя

1. Линейный оператор, отображающий n -мерное векторное пространство R само в себя (в данном случае $R \equiv S$, $n = m$), мы будем просто называть *линейным оператором в R* .

Сумма двух линейных операторов в R , а также произведение такого оператора на число — снова линейные операторы в R . Умножение двух таких линейных операторов всегда выполнимо, и произведение их есть снова линейный оператор в R . Таким образом, линейные операторы в R образуют кольцо¹⁾. В этом кольце имеется единичный оператор, т. е. оператор E , для которого

$$Ex = x \quad (x \in R).$$

При этом для произвольного оператора A в R

$$EA = AE = A.$$

Если A — линейный оператор в R , то имеет смысл $A^2 = AA$, $A^3 = AAA$, ... и вообще $A^m = \underbrace{AA \dots A}_{m \text{ раз}}$. Кроме того, полагаем $A^0 = E$.

Тогда, как легко видеть, при любых целых неотрицательных p и q

$$A^p A^q = A^{p+q}.$$

Пусть $f(t) = a_0 t^m + a_1 t^{m-1} + \dots + a_{m-1} t + a_m$ — многочлен относительно скалярного аргумента t с коэффициентами из поля K . Тогда полагаем:

$$f(A) = a_0 A^m + a_1 A^{m-1} + \dots + a_{m-1} A + a_m E.$$

При этом $f(A)g(A) = g(A)f(A)$ для любых двух многочленов $f(t)$ и $g(t)$.

Пусть

$$y = Ax \quad (x, y \in R). \quad (45)$$

Обозначим через x_1, x_2, \dots, x_n координаты вектора x в произвольном базисе e_1, e_2, \dots, e_n , а через y_1, y_2, \dots, y_n — координаты вектора y в том же базисе. Тогда

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i = 1, 2, \dots, n). \quad (46)$$

В базисе e_1, e_2, \dots, e_n линейному оператору A отвечает квадратная матрица $A = [a_{ik}]$ ²⁾. Напомним читателю (см. стр. 71), что в k -м столбце этой матрицы стоят координаты вектора Ae_k ($k = 1, 2, \dots, n$), т. е.

$$Ae_k = \sum_{i=1}^n a_{ik} e_i \quad (k = 1, \dots, n). \quad (47)$$

Введя координатные столбцы $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$, мы преобразование (46) можем записать в матричной форме

$$y = Ax. \quad (48)$$

¹⁾ Это кольцо является алгеброй. Ср. гл. I, стр. 28.

²⁾ См. стр. 70—71. В данном случае пространства R и S совпадают; точно также отождествлены базисы e_1, e_2, \dots, e_n и g_1, g_2, \dots, g_m в этих пространствах.

Сумме и произведению двух операторов A и B отвечают сумма и произведение соответствующих квадратных матриц $A = \|a_{ik}\|_1^n$ и $B = \|b_{ik}\|_1^n$. Произведению aA соответствует матрица aA . Единичному оператору E отвечает квадратная единичная матрица $E = \|\delta_{ik}\|_1^n$. Таким образом, выбор базиса устанавливает изоморфное соответствие между кольцом линейных операторов в R и кольцом квадратных матриц n -го порядка с элементами из K . При этом соответствию многочлену $f(A)$ соответствует матрица $f(A)$.

2. Рассмотрим наряду с базисом e_1, e_2, \dots, e_n другой базис $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$ в R . Тогда аналогично (48)

$$\tilde{y} = \tilde{A} \tilde{x}, \quad (49)$$

где \tilde{x} , \tilde{y} — столбцевые матрицы, составленные из координат векторов x , y в базисе $\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n$, а $\tilde{A} = \|\tilde{a}_{ik}\|_1^n$ — квадратная матрица, соответствующая оператору A в этом базисе. Запишем в матричной форме формулы преобразования координат

$$x = T\tilde{x}, \quad y = T\tilde{y}. \quad (50)$$

Тогда из (48) и (50) находим:

$$\tilde{y} = T^{-1}AT\tilde{x},$$

что в сопоставлении с (49) дает:

$$\tilde{A} = T^{-1}AT. \quad (51)$$

Формула (51) представляет собой специальный частный случай формулы (31) на стр. 75 (в данном случае $P = T^{-1}$, $Q = T$).

Определение 10. Две матрицы A и B , связанные соотношением

$$B = T^{-1}AT, \quad (51')$$

где T — некоторая неособенная матрица, называются подобными¹⁾.

Таким образом, мы показали, что две матрицы, соответствующие одному и тому же линейному оператору в R при различных базисах, подобны между собой, причем матрица T , связывающая эти матрицы, совпадает с матрицей преобразования координат при переходе от первого базиса ко второму (см. (50)).

Другими словами, линейному оператору в R отвечает целый класс подобных между собой матриц; эти матрицы представляют данный оператор в различных базисах.

Изучая свойства линейного оператора в R , мы тем самым изучаем свойства матриц, присущие одновременно всему классу подобных матриц, т. е. изучаем свойства матриц, остающиеся неизменными инвариантными при переходе от данной матрицы к матрице, ей подобной.

Заметим еще, что две подобные матрицы имеют всегда равные определители. Действительно, из (51') следует, что

$$|B| = |T|^{-1}|A||T| = |A|. \quad (52)$$

Равенство $|B| = |A|$ является необходимым, но не достаточным условием для подобия матриц A и B .

1) Матрицу T всегда можно выбрать так, чтобы ее элементы принадлежали основному числовому полю K , которому принадлежат элементы матриц A и B . Легко проверяются три свойства подобия матриц: рефлексивность (матрица A всегда подобна самой себе), симметричность (если A подобна B , то и B подобна A) и транзитивность (если A подобна B , B подобна C , то A подобна C).

В главе VI будет установлен критерий подобия двух матриц, т. е. будут даны необходимые и достаточные условия для того, чтобы две квадратные матрицы n -го порядка были подобны между собой.

Согласно равенству (52) мы можем под определителем линейного оператора A в $R(|A|)$ понимать определитель любой матрицы, соответствующей данному оператору.

Если $|A|=0$ ($\neq 0$), то оператор A называется *особенным* (соответственно *неособенным*). Согласно этому определению в любом базисе особенному (неособенному) оператору отвечает особенная (соответственно неособенная) матрица. Для особенного оператора:

- 1) всегда существует вектор $x \neq 0$ такой, что $Ax=0$,
- 2) AR составляет правильную часть R . Для неособенного оператора:
- 1) из $Ax=0$ следует $x=0$;
- 2) $AR \equiv R$, т. е. векторы вида Ax ($x \in R$) заполняют все пространство R .

Другими словами, линейный оператор в R является особенным или неособенным в зависимости от того, больше или равен нулю его дефект.

3. Если A — неособенный оператор, то в равенстве $y = Ax$ задание вектора $y \in R$ однозначно определяет вектор $x \in R$. Действительно, существование вектора x следует из того, что векторы вида Ax ($x \in R$) заполняют все пространство R . С другой стороны, из равенств $y = Ax'$ и $y = Ax''$ ($x', x'' \in R$) следует: $A(x' - x'') = Ax' - Ax'' = 0$ и отсюда: $x' - x'' = 0$, т. е. $x' = x''$. Поэтому, исходя из равенства $y = Ax$, можно определить обратный оператор A^{-1} равенством $x = A^{-1}y$. Легко видеть, что обратный оператор A^{-1} для линейного оператора A в R также является линейным оператором в R ; при этом

$$AA^{-1} = A^{-1}A = E,$$

где E — единичный оператор. Если в некотором базисе неособенному оператору A отвечает неособенная матрица A , то в этом базисе обратному оператору A^{-1} , соответствует матрица A^{-1} .

Рассмотрим некоторые частные типы линейных операторов в R .

1°. Оператор J в R называется *инволютивным*, если $J^2=E$. Инволютивный оператор неособенный и для него $J^{-1}=J$. Инволютивному оператору в любом базисе соответствует *инволютивная матрица* J , т. е. матрица J , для которой $J^2=E$.

2°. Оператор P в R называется *проекционным*, если $P^2=P$. Пусть дано произвольное расщепление пространства R на два подпространства S и T : $R=S+T$. Тогда для любого вектора $x \in R$ имеет место разложение $x=x_S+x_T$, где $x_S \in S$, $x_T \in T$. Вектор x_S называется *проекцией вектора* x на подпространство S параллельно подпространству T ¹⁾. Рассмотрим оператор P , осуществляющий проектирование пространства R на подпространство S параллельно подпространству T , т. е. оператор в R , определяемый равенством $Px=x_S$ для любого вектора $x \in R$. Очевидно, этот оператор является линейным, но он является и проективным, так как $Px=x_S$, $P^2x=Px=x_S$ и, следовательно, $(P^2-P)x=x_S-x_S=0$, т. е. $P^2=P$.

Легко проверяется и обратное утверждение. Произвольный проекционный оператор P в R осуществляет проектирование всего пространства R на подпространство $S=PR$ параллельно подпространству $T=(E-P)R$.

Любая натуральная степень проекционного оператора является проекционным оператором. Если P — проекционный оператор, то и $E-P$ — проекционный оператор, так как $(E-P)^2=E-2P+P^2=E-P$.

Квадратная матрица P называется *проекционной*, если $P^2=P$. Очевидно, в произвольном базисе проекционному оператору соответствует проекционная матрица.

¹⁾ Аналогично, вектор x_T — проекция вектора x на подпространство T параллельно подпространству S .

§ 7. Характеристические числа и собственные векторы линейного оператора

При исследовании структуры линейного оператора A в R большую роль играют векторы x , для которых

$$Ax = \lambda x \quad (\lambda \in K, x \neq 0). \quad (53)$$

Такие векторы называются *собственными векторами*, а соответствующие им числа λ — *характеристическими* или *собственными числами* оператора A (матрицы A).

Для нахождения характеристических чисел и собственных векторов оператора A выберем произвольно базис e_1, e_2, \dots, e_n в R . Пусть $x = \sum_{i=1}^n x_i e_i$ и $A = \|a_{ik}\|_1^n$ — матрица, отвечающая оператору A в базисе e_1, e_2, \dots, e_n . Тогда, приравнивая между собой соответственные координаты векторов, стоящих в левой и правой частях равенства (53), получим систему скалярных уравнений

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \lambda x_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = \lambda x_2, \\ \dots \dots \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = \lambda x_n, \end{array} \right\} \quad (54)$$

которую можно записать и так:

$$\left. \begin{array}{l} (a_{11} - \lambda)x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \dots + a_{2n}x_n = 0, \\ \dots \dots \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + (a_{nn} - \lambda)x_n = 0. \end{array} \right\} \quad (55)$$

Так как искомый вектор не должен быть равен нулю, то среди его координат x_1, x_2, \dots, x_n по крайней мере одна координата должна быть отлична от нуля.

Для того чтобы система линейных однородных уравнений (55) имела ненулевое решение, необходимо и достаточно, чтобы определитель этой системы был равен нулю:

$$\left| \begin{array}{cccc} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{array} \right| = 0. \quad (56)$$

Уравнение (56) представляет собой алгебраическое уравнение n -й степени относительно λ . Коэффициенты этого уравнения принадлежат тому же числовому полю, что и элементы матрицы $A = \|a_{ik}\|_1^n$, т. е. полю K .

Уравнение (56) часто встречается в различных проблемах геометрии, механики, астрономии, физики и носит название *характеристического уравнения* или *векового¹⁾ уравнения* матрицы $A = \|a_{ik}\|_1^n$ (левую часть этого уравнения называют *характеристическим многочленом*).

¹⁾ Такое название связано с тем, что это уравнение встречается при исследовании вековых возмущений планет.

Таким образом, каждое характеристическое число λ линейного оператора A является корнем характеристического уравнения (56). И наоборот, если некоторое число λ является корнем уравнения (56), то при этом значении λ система (55) и, следовательно, (54) имеет ненулевое решение x_1, x_2, \dots, x_n , т. е. этому числу λ отвечает собственный вектор $x = \sum x_i e_i$ оператора A .

Из сказанного следует, что любой линейный оператор A в R имеет не более чем n различных характеристических чисел.

Если K есть поле всех комплексных чисел, то любой линейный оператор в R всегда имеет по крайней мере один собственный вектор в R и соответствующее этому собственному вектору характеристическое число λ^1). Это следует из основной теоремы алгебры, согласно которой алгебраическое уравнение (56) в поле комплексных чисел всегда имеет по крайней мере один корень.

Напишем уравнение (56) в развернутом виде

$$|A - \lambda E| \equiv (-\lambda)^n + S_1(-\lambda)^{n-1} + S_2(-\lambda)^{n-2} + \dots + S_{n-1}(-\lambda) + S_n = 0; \quad (57)$$

здесь, как нетрудно видеть,

$$S_1 = \sum_{i=1}^n a_{ii}, \quad S_2 = \sum_{1 \leq i < k \leq n} A \begin{pmatrix} i & k \\ i & k \end{pmatrix}, \dots \quad (58)$$

и вообще S_p равно сумме главных миноров p -го порядка матрицы $A = \|a_{ik}\|_1^n$ ($p = 1, 2, \dots, n$)²). В частности, $S_n = |A|$.

Обозначим через \tilde{A} матрицу, соответствующую тому же оператору A в другом базисе. Матрица \tilde{A} подобна матрице A :

$$\tilde{A} = T^{-1}AT.$$

Отсюда

$$\tilde{A} - \lambda E = T^{-1}(A - \lambda E)T$$

¹⁾ Это положение справедливо и в более общем случае, когда K — произвольное алгебраически замкнутое поле, т. е. такое поле, которому принадлежат корни всех алгебраических уравнений с коэффициентами из данного поля.

²⁾ Степень $(-\lambda)^{n-p}$ имеется только в тех членах характеристического определителя (56), которые содержат какие-либо $n-p$ диагональных элементов

$$a_{j_1 j_1} - \lambda, a_{j_2 j_2} - \lambda, \dots, a_{j_{n-p} j_{n-p}} - \lambda.$$

Произведение этих диагональных элементов входит в состав определителя (56) со множителем, равным главному минору

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix},$$

где индексы i_1, i_2, \dots, i_p вместе с индексами j_1, j_2, \dots, j_{n-p} образуют полную систему индексов $1, 2, \dots, n$:

$$|A - \lambda E| = (a_{j_1 j_1} - \lambda)(a_{j_2 j_2} - \lambda) \dots (a_{j_{n-p} j_{n-p}} - \lambda) A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} + (\ast).$$

Здесь $(-\lambda)^{n-p}$ умножается на $A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix}$. Перебирая всевозможные сочетания j_1, j_2, \dots, j_{n-p} по $n-p$ из индексов $1, 2, \dots, n$, мы получим в качестве коэффициента S_p при $(-\lambda)^{n-p}$ сумму всех главных миноров p -го порядка матрицы A .

и, следовательно,

$$|\tilde{A} - \lambda E| = |A - \lambda E|. \quad (59)$$

Таким образом, подобные матрицы A и \tilde{A} имеют один и тот же характеристический многочлен. Этот многочлен иногда называют характеристическим многочленом оператора A и обозначают через $|A - \lambda E|$.

Если x, y, z, \dots — собственные векторы оператора A , соответствующие одному и тому же характеристическому числу λ , а a, β, γ, \dots — произвольные числа из K , то вектор $ax + \beta y + \gamma z + \dots$ либо равен нулю, либо также является собственным вектором оператора A при том же числе λ . Действительно, из

$$Ax = \lambda x, \quad Ay = \lambda y, \quad Az = \lambda z, \dots$$

следует:

$$A(ax + \beta y + \gamma z + \dots) = \lambda(ax + \beta y + \gamma z + \dots).$$

Поэтому линейно независимые собственные векторы, отвечающие одному и тому же характеристическому числу λ , образуют базис некоторого «собственного» подпространства, каждый вектор которого есть собственный вектор при том же λ . В частности, каждый собственный вектор порождает одномерное собственное подпространство, «собственное направление».

Однако, если собственные векторы оператора A соответствуют различным характеристическим числам, то линейная комбинация этих собственных векторов, вообще говоря, не будет собственным вектором оператора A .

Значение собственных векторов и характеристических чисел при исследовании линейных операторов будет проиллюстрировано в следующем параграфе на примере операторов простой структуры.

§ 8. Линейные операторы простой структуры

Начнем со следующей леммы.

Л е м м а. Собственные векторы, соответствующие попарно различным характеристическим числам, всегда линейно независимы.

Д о к а з а т е л ь с т в о. Пусть

$$Ax_i = \lambda_i x_i \quad (x_i \neq 0; \lambda_i \neq \lambda_k \text{ при } i \neq k; i, k = 1, 2, \dots, m) \quad (60)$$

и пусть

$$\sum_{i=1}^m c_i x_i = 0. \quad (61)$$

Применяя к обеим частям этого равенства оператор A , получим:

$$\sum_{i=1}^m c_i \lambda_i x_i = 0. \quad (62)$$

Умножим обе части равенства (61) на λ_1 и вычтем почленно (61) из (62). Тогда получим:

$$\sum_{i=2}^m c_i (\lambda_i - \lambda_1) x_i = 0. \quad (63)$$

Можно сказать, что равенство (63) было получено из (61) путем почленного применения оператора $A - \lambda_1 E$. Применяя к (63) почленно операторы $A - \lambda_2 E, \dots, A - \lambda_{m-1} E$, мы придем к следующему равенству:

$$c_m (\lambda_m - \lambda_{m-1}) (\lambda_m - \lambda_{m-2}) \dots (\lambda_m - \lambda_1) x_m = 0,$$

откуда $c_m = 0$. Так как в (61) любое слагаемое может быть поставлено на последнее место, то в (61)

$$c_1 = c_2 = \dots = c_m = 0,$$

т. е. между векторами x_1, x_2, \dots, x_m нет линейной зависимости. Лемма доказана.

Если характеристическое уравнение оператора имеет n различных корней и эти корни принадлежат полю K , то на основании леммы собственные векторы, соответствующие этим корням, линейно независимы.

Определение 11. Линейный оператор A в R называется *оператором простой структуры*, если A имеет в R n линейно независимых собственных векторов, где n — число измерений.

Таким образом, линейный оператор в R имеет простую структуру, если все корни характеристического уравнения различны между собой и принадлежат полю K . Однако это условие не является необходимым. Существуют линейные операторы простой структуры, у которых характеристический многочлен имеет кратные корни.

Рассмотрим произвольный линейный оператор A простой структуры. Обозначим через g_1, g_2, \dots, g_n базис в R , состоящий из собственных векторов оператора, т. е.

$$Ag_k = \lambda_k g_k \quad (k = 1, 2, \dots, n).$$

Если

$$x = \sum_{k=1}^n x_k g_k, \text{ то } Ax = \sum_{k=1}^n x_k Ag_k = \sum_{k=1}^n \lambda_k x_k g_k.$$

Другими словами, воздействие оператора A простой структуры на вектор $x = \sum_{k=1}^n x_k g_k$ может быть описано следующим образом:

В n -мерном пространстве R существует n линейно независимых «направлений», вдоль которых оператор простой структуры A осуществляет «растяжение» с коэффициентами $\lambda_1, \lambda_2, \dots, \lambda_n$. Произвольный вектор x может быть разложен на компоненты, идущие вдоль этих собственных направлений. Эти компоненты подвергаются соответствующим «растяжениям», после чего они в сумме дают вектор Ax .

Нетрудно видеть, что оператору A в «собственном» базисе g_1, g_2, \dots, g_n соответствует диагональная матрица

$$\tilde{A} = \|\lambda_i \delta_{ik}\|_1^n.$$

Если мы через A обозначим матрицу, отвечающую оператору A в произвольном базисе e_1, e_2, \dots, e_n , то

$$A = T \|\lambda_i \delta_{ik}\|_1^n T^{-1}. \quad (64)$$

Матрицу, подобную диагональной, будем называть *матрицей простой структуры*. Таким образом, оператору простой структуры в любом базисе отвечает матрица простой структуры и наоборот.

Матрица T в равенстве (64) осуществляет переход от базиса e_1, e_2, \dots, e_n к базису g_1, g_2, \dots, g_n . В k -м столбце матрицы T стоят координаты (в базисе e_1, e_2, \dots, e_n) собственного вектора g_k , соответствующего характеристическому числу λ_k матрицы A ($k = 1, 2, \dots, n$). Матрица T называется *функциональной матрицей для матрицы A* .

Равенство (64) перепишем так:

$$A = TLT^{-1} \quad (L = \{\lambda_1, \lambda_2, \dots, \lambda_n\}). \quad (64')$$

Переходя к p -м ассоциированным матрицам ($1 \leq p \leq n$), получим (см. гл. I, § 4):

$$\mathfrak{A}_p = \mathfrak{L}_p \mathfrak{E}_p \mathfrak{L}_p^{-1}; \quad (65)$$

\mathfrak{L}_p — диагональная матрица N -го порядка ($N = C_n^p$), у которой на главной диагонали стоят всевозможные произведения по p из $\lambda_1, \lambda_2, \dots, \lambda_n$. Сопоставление (65) с (64') дает нам теорему:

Теорема 3. Если матрица $A = \|a_{ik}\|_1^n$ имеет простую структуру, то при любом $p \leq n$ ассоциированная матрица \mathfrak{A}_p также имеет простую структуру; при этом характеристическими числами матрицы \mathfrak{A}_p являются всевозможные произведения по p $\lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_p}$ ($1 \leq i_1 < i_2 < \dots < i_p \leq n$) из характеристических чисел $\lambda_1, \lambda_2, \dots, \lambda_n$ матрицы A , а фундаментальной матрицей матрицы \mathfrak{A}_p является ассоциированная \mathfrak{L}_p для фундаментальной матрицы T матрицы A .

Следствие. Если характеристическому числу λ_k матрицы простой структуры $A = \|a_{ik}\|_1^n$ отвечает собственный вектор с координатами $t_{1k}, t_{2k}, \dots, t_{nk}$ ($k = 1, 2, \dots, n$) и $T = \|t_{ik}\|_1^n$, то характеристическому числу $\lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_p}$ ($1 \leq k_1 < k_2 < \dots < k_p \leq n$) матрицы \mathfrak{A}_p отвечает собственный вектор с координатами

$$T \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n). \quad (66)$$

Произвольную матрицу $A = \|a_{ik}\|_1^n$ можно представить в виде предела последовательности матриц A_m ($m \rightarrow \infty$), каждая из которых не имеет кратных характеристических чисел и потому имеет простую структуру. Характеристические числа $\lambda_1^{(m)}, \lambda_2^{(m)}, \dots, \lambda_n^{(m)}$ матрицы A_m в пределе при $m \rightarrow \infty$ переходят в характеристические числа $\lambda_1, \lambda_2, \dots, \lambda_n$ матрицы A , т. е.

$$\lim_{m \rightarrow \infty} \lambda_k^{(m)} = \lambda_k \quad (k = 1, 2, \dots, n).$$

Отсюда

$$\lim_{m \rightarrow \infty} \lambda_{k_1}^{(m)} \lambda_{k_2}^{(m)} \dots \lambda_{k_p}^{(m)} = \lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_p} \quad (1 \leq k_1 < k_2 < \dots < k_p \leq n).$$

Так как, кроме того, $\lim_{m \rightarrow \infty} \mathfrak{A}_{(m)p} = \mathfrak{A}_p$, то из теоремы 3 вытекает

Теорема 4 (Кронекера). Если $\lambda_1, \lambda_2, \dots, \lambda_n$ — полная система характеристических чисел произвольной матрицы A , то полная система характеристических чисел ассоциированной матрицы \mathfrak{A}_p состоит из всевозможных произведений по p из чисел $\lambda_1, \lambda_2, \dots, \lambda_n$ ($p = 1, 2, \dots, n$).

В этом параграфе мы исследовали операторы и матрицы простой структуры. Изучение структуры операторов и матриц общего типа будет проведено в главах VI и VII.

ГЛАВА IV

ХАРАКТЕРИСТИЧЕСКИЙ И МИНИМАЛЬНЫЙ МНОГОЧЛЕНЫ МАТРИЦЫ

С каждой квадратной матрицей связаны два многочлена: характеристический и минимальный. Эти многочлены играют большую роль в различных вопросах теории матриц. Так, например, понятие о функции от матрицы, которое мы введем в следующей главе, будет целиком основываться на понятии о минимальном многочлене матрицы. В этой главе рассматриваются свойства характеристического и минимального многочлена. Этому исследованию предпосылаются основные сведения о многочленах с матричными коэффициентами и о действиях над ними.

§ 1. Сложение и умножение матричных многочленов

Рассмотрим квадратную *многочленную* матрицу $A(\lambda)$, т. е. квадратную матрицу, элементами которой являются многочлены относительно λ (с коэффициентами из данного числового поля K):

$$A(\lambda) = \| a_{ik}(\lambda) \|_1^n = \| a_{ik}^{(0)}\lambda^m + a_{ik}^{(1)}\lambda^{m-1} + \dots + a_{ik}^{(m)} \|_1^n. \quad (1)$$

Матрицу $A(\lambda)$ можно представить в виде многочлена с матричными коэффициентами, расположенного по степеням λ :

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \dots + A_m, \quad (2)$$

где

$$A_j = \| a_{ik}^{(j)} \|_1^n \quad (j = 0, 1, \dots, m). \quad (3)$$

Число m называется *степенью* многочлена, если $A_0 \neq 0$. Число n называется *порядком* многочлена. Многочлен (1) будем называть *регулярным*, если $|A_0| \neq 0$.

Многочлен с матричными коэффициентами мы будем иногда называть *матричным многочленом*. В отличие от матричного многочлена обычный многочлен со скалярными коэффициентами будем называть *скалярным многочленом*.

Рассмотрим основные действия над матричными многочленами. Пусть даны два матричных многочлена одного и того же порядка $A(\lambda)$ и $B(\lambda)$. Обозначим через m наибольшую из степеней этих многочленов. Эти многочлены можно записать в виде

$$\begin{aligned} A(\lambda) &= A_0\lambda^m + A_1\lambda^{m-1} + \dots + A_m, \\ B(\lambda) &= B_0\lambda^m + B_1\lambda^{m-1} + \dots + B_m. \end{aligned}$$

Тогда

$$A(\lambda) \pm B(\lambda) = (A_0 \pm B_0) \lambda^m + (A_1 \pm B_1) \lambda^{m-1} + \dots + A_m \pm B_m,$$

т. е. сумма (разность) двух матричных многочленов одного и того же порядка может быть представлена в виде многочлена, степень которого не превосходит наибольшей из степеней данных многочленов.

Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ степеней m и p одного и того же порядка n :

$$A(\lambda) = A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m \quad (A_0 \neq 0),$$

$$B(\lambda) = B_0 \lambda^p + B_1 \lambda^{p-1} + \dots + B_p \quad (B_0 \neq 0).$$

Тогда

$$A(\lambda) B(\lambda) = A_0 B_0 \lambda^{m+p} + (A_0 B_1 + A_1 B_0) \lambda^{m+p-1} + \dots + A_m B_p. \quad (4)$$

Если бы мы перемножили $B(\lambda)$ на $A(\lambda)$ (т. е. изменили бы порядок сомножителей), то мы получили бы, вообще говоря, другой многочлен.

Умножение матричных многочленов обладает еще одним специфическим свойством. В отличие от произведения скалярных многочленов произведение матричных многочленов (4) может иметь степень, меньшую $m+p$, т. е. меньшую суммы степеней сомножителей. Действительно, в (4) произведение матриц $A_0 B_0$ может равняться нулю при $A_0 \neq 0$ и $B_0 \neq 0$. Однако, если хотя бы одна из матриц A_0 и B_0 неособенная, то из $A_0 \neq 0$ и $B_0 \neq 0$ следует: $A_0 B_0 \neq 0$. Таким образом, произведение двух матричных многочленов равно многочлену, степень которого меньше или равна сумме степеней сомножителей. Если хотя бы один из двух сомножителей регулярный многочлен, то в этом случае степень произведения всегда равна сумме степеней сомножителей.

Матричный многочлен n -го порядка $A(\lambda)$ можно записать двояко:

$$A(\lambda) = A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m \quad (5)$$

и

$$A(\lambda) = \lambda^m A_0 + \lambda^{m-1} A_1 + \dots + A_m. \quad (5')$$

Обе записи при скалярном λ дают один и тот же результат. Однако если мы пожелаем вместо скалярного аргумента λ подставить квадратную матрицу n -го порядка Λ , то результаты подстановок в (5) и (5') будут, вообще говоря, различны, так как степени матрицы Λ могут не быть перестановочными с матричными коэффициентами A_0, A_1, \dots, A_m .

Положим

$$A(\Lambda) = A_0 \Lambda^m + A_1 \Lambda^{m-1} + \dots + A_m \quad (6)$$

и

$$\widehat{A}(\Lambda) = \Lambda^m A_0 + \Lambda^{m-1} A_1 + \dots + A_m \quad (6')$$

и будем называть $A(\Lambda)$ правым, а $\widehat{A}(\Lambda)$ левым значением матричного многочлена $A(\lambda)$ при подстановке вместо λ матрицы Λ ¹⁾.

Рассмотрим снова два матричных многочлена

$$A(\lambda) = \sum_{i=0}^m A_{m-i} \lambda^i, \quad B(\lambda) = \sum_{k=0}^p B_{p-k} \lambda^k$$

¹⁾ В «правом» значении $A(\lambda)$ степени матрицы Λ стоят справа от коэффициентов, а в «левом» — слева.

и их произведение

$$P(\lambda) = \sum_{i=0}^m \sum_{k=0}^p A_{m-i} \lambda^i B_{p-k} \lambda^k = \sum_{i=0}^m \sum_{k=0}^p A_{m-i} B_{p-k} \lambda^{i+k} = \sum_{j=0}^{m+p} \left(\sum_{i+k=j} A_{m-i} B_{p-k} \right) \lambda^j \quad (7')$$

и

$$P(\lambda) = \sum_{i=0}^m \sum_{k=0}^p \lambda^i A_{m-i} \lambda^k B_{p-k} = \sum_{i=0}^m \sum_{k=0}^p \lambda^{i+k} A_{m-i} B_{p-k} = \sum_{j=0}^{m+p} \lambda^j \sum_{i+k=j} A_{m-i} B_{p-k}. \quad (7'')$$

Преобразования в тождестве (7') сохраняют свою силу при замене λ матрицей n -го порядка Λ , если только матрица Λ перестановочна со всеми матричными коэффициентами B_{p-k} ¹⁾. Аналогично в тождестве (7'') можно заменить скаляр λ матрицей Λ , если матрица Λ перестановочна со всеми коэффициентами A_{m-i} . В первом случае получаем:

$$P(\Lambda) = A(\Lambda) B(\Lambda), \quad (8')$$

во втором

$$\widehat{P}(\Lambda) = \widehat{A}(\Lambda) \widehat{B}(\Lambda). \quad (8'')$$

Таким образом, *правое (левое) значение произведения двух матричных многочленов равно произведению правых (левых) значений сомножителей, если матрица-аргумент Λ перестановочна со всеми коэффициентами правого (левого) сомножителя.*

Если $S(\lambda)$ — сумма двух матричных многочленов n -го порядка $A(\lambda)$ и $B(\lambda)$, то при замене скаляра λ любой матрицей n -го порядка Λ всегда справедливы тождества

$$S(\Lambda) = A(\Lambda) + B(\Lambda), \quad \widehat{S}(\Lambda) = \widehat{A}(\Lambda) + \widehat{B}(\Lambda). \quad (9)$$

§ 2. Правое и левое деление матричных многочленов.

Обобщенная теорема Безу

Пусть даны два матричных многочлена $A(\lambda)$ и $B(\lambda)$ одного и того же порядка n , причем $B(\lambda)$ — регулярный многочлен:

$$A(\lambda) = A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m \quad (A_0 \neq 0),$$

$$B(\lambda) = B_0 \lambda^p + B_1 \lambda^{p-1} + \dots + B_p \quad (|B_0| \neq 0).$$

Мы будем говорить, что матричные многочлены $Q(\lambda)$ и $R(\lambda)$ являются соответственно *правым частным* и *правым остатком* при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda) = Q(\lambda) B(\lambda) + R(\lambda) \quad (10)$$

и степень $R(\lambda)$ меньше степени $B(\lambda)$.

Совершенно аналогично будем называть многочлены $\widehat{Q}(\lambda)$ и $\widehat{R}(\lambda)$ соответственно *левым частным* и *левым остатком* при делении $A(\lambda)$ на $B(\lambda)$, если

$$A(\lambda) = B(\lambda) \widehat{Q}(\lambda) + \widehat{R}(\lambda) \quad (11)$$

и степень $\widehat{R}(\lambda)$ меньше степени $B(\lambda)$.

¹⁾ В этом случае и любая степень матрицы Λ перестановочна со всеми коэффициентами B_{p-k} .

Обратим внимание читателя, что при «правом» делении (т. е. при нахождении правого частного и правого остатка) в (5) на «делитель» $B(\lambda)$ частное $Q(\lambda)$ умножается справа, а при «левом» делении в (6) на делитель $B(\lambda)$ частное $\widehat{Q}(\lambda)$ умножается слева. В общем случае многочлены $Q(\lambda)$ и $R(\lambda)$ не совпадают с $\widehat{Q}(\lambda)$ и $\widehat{R}(\lambda)$.

Покажем, что как правое, так и левое деление матричных многочленов одного и того же порядка всегда выполнимо и однозначно, если делитель — регулярный многочлен.

Рассмотрим правое деление $A(\lambda)$ на $B(\lambda)$. Если $m < p$, можно положить $Q(\lambda) = 0$, $R(\lambda) = A(\lambda)$. В случае $m \geq p$ для нахождения частного $Q(\lambda)$ и остатка $R(\lambda)$ применим обычную схему деления многочлена на многочлен. «Разделим» старший член делимого $A_0\lambda^m$ на старший член делителя $B_0\lambda^p$. Получим старший член искомого частного $A_0B_0^{-1}\lambda^{m-p}$. Умножим этот член справа на делитель $B(\lambda)$ и полученное произведение вычтем из $A(\lambda)$. Найдем «первый остаток» $A^{(1)}(\lambda)$:

$$A(\lambda) = A_0B_0^{-1}\lambda^{m-p}B(\lambda) + A^{(1)}(\lambda). \quad (12)$$

Степень $m^{(1)}$ многочлена $A^{(1)}(\lambda)$ меньше m :

$$A^{(1)}(\lambda) = A_0^{(1)}\lambda^{m^{(1)}} + \dots \quad (A_0^{(1)} \neq 0, m^{(1)} < m). \quad (13)$$

Если $m^{(1)} \geq p$, то, повторяя этот процесс, получаем:

$$\left. \begin{array}{l} A^{(1)}(\lambda) = A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p}B(\lambda) + A^{(2)}(\lambda), \\ A^{(2)}(\lambda) = A_0^{(2)}\lambda^{m^{(2)}} + \dots \quad (m^{(2)} < m^{(1)}) \end{array} \right\} \quad (14)$$

и т. д.

Так как степени многочленов $A(\lambda)$, $A^{(1)}(\lambda)$, $A^{(2)}(\lambda)$, ... убывают, то на некотором этапе мы придем к остатку $R(\lambda)$, степень которого меньше p . Тогда из (12) — (14) будет следовать:

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda),$$

где

$$Q(\lambda) = A_0B_0^{-1}\lambda^{m-p} + A_0^{(1)}B_0^{-1}\lambda^{m^{(1)}-p} + \dots$$

Докажем теперь однозначность правого деления. Пусть одновременно

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) \quad (15)$$

и

$$A(\lambda) = Q^*(\lambda)B(\lambda) + R^*(\lambda), \quad (15')$$

где степени многочленов $R(\lambda)$ и $R^*(\lambda)$ меньше степени $B(\lambda)$, т. е. меньше p . Вычитая почленно (15) из (15'), получим:

$$[Q(\lambda) - Q^*(\lambda)]B(\lambda) = R^*(\lambda) - R(\lambda). \quad (16)$$

Если бы $Q(\lambda) - Q^*(\lambda) \not\equiv 0$, то, поскольку $|B_0| \neq 0$, степень левой части равенства (16) равнялась бы сумме степеней $B(\lambda)$ и $Q(\lambda) - Q^*(\lambda)$ и потому была бы $\geq p$. Это невозможно, так как степень многочлена, стоящего в правой части равенства (16), меньше p . Таким образом, $Q(\lambda) - Q^*(\lambda) \equiv 0$, а тогда из (16) $R^*(\lambda) - R(\lambda) \equiv 0$, т. е.

$$Q(\lambda) = Q^*(\lambda), \quad R(\lambda) = R^*(\lambda).$$

Совершенно аналогично устанавливаются существование и единственность левого частного и левого остатка¹⁾.

Пример.

$$A(\lambda) = \begin{vmatrix} \lambda^3 + \lambda & 2\lambda^3 + \lambda^2 \\ -\lambda^3 - 2\lambda^2 + 1 & 3\lambda^3 + \lambda \end{vmatrix} =$$

A_0

$$= \underbrace{\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}}_{A_0} \lambda^3 + \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix} \lambda^2 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \lambda + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

B_0

$$B(\lambda) = \begin{vmatrix} 2\lambda^2 + 3 & -\lambda^2 + 1 \\ -\lambda^2 - 1 & \lambda^2 + 2 \end{vmatrix} = \underbrace{\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}}_{B_0} \lambda^2 + \begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix},$$

$$\|B_0\| = 1, \quad B_0^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad A_0 B_0^{-1} = \begin{pmatrix} 3 & 5 \\ 2 & 5 \end{pmatrix}, \quad A_0 B_0^{-1} B(\lambda) = \begin{pmatrix} \lambda^2 + 4 & 2\lambda^2 + 13 \\ -\lambda^2 + 1 & 3\lambda^2 + 12 \end{pmatrix},$$

$$A^{(1)}(\lambda) = \begin{vmatrix} \lambda^3 + \lambda & 2\lambda^3 + \lambda^2 \\ -\lambda^3 - 2\lambda^2 + 1 & 3\lambda^3 + \lambda \end{vmatrix} - \begin{vmatrix} \lambda^3 + 4\lambda & 2\lambda^3 + 13\lambda \\ -\lambda^3 + \lambda & 3\lambda^3 + 12\lambda \end{vmatrix} =$$

$$= \begin{vmatrix} -3\lambda & \lambda^2 - 13\lambda \\ -2\lambda^2 - \lambda + 1 & -11\lambda \end{vmatrix},$$

$$A^{(1)}(\lambda) = \begin{vmatrix} 0 & 1 \\ -2 & 0 \end{vmatrix} \lambda^2 + \begin{vmatrix} -3 & -13 \\ -1 & -11 \end{vmatrix} \lambda + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

$$A_0^{(1)} B_0^{-1} = \begin{vmatrix} 0 & 1 \\ -2 & 0 \end{vmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & -2 \end{pmatrix},$$

$$A_0^{(1)} B_0^{-1} B(\lambda) = \begin{pmatrix} 1 & 2 \\ -2 & -2 \end{pmatrix} \cdot \begin{vmatrix} 2\lambda^2 + 3 & -\lambda^2 + 1 \\ -\lambda^2 - 1 & \lambda^2 + 2 \end{vmatrix} = \begin{pmatrix} 1 & \lambda^2 + 5 \\ -2\lambda^2 - 4 & -6 \end{pmatrix},$$

$$R(\lambda) = A^{(1)}(\lambda) - A_0^{(1)} B_0^{-1} B(\lambda) =$$

$$= \begin{vmatrix} -3\lambda & \lambda^2 - 13\lambda \\ -2\lambda^2 - \lambda + 1 & -11\lambda \end{vmatrix} - \begin{vmatrix} 1 & \lambda^2 + 5 \\ -2\lambda^2 - 4 & -6 \end{vmatrix} = \begin{pmatrix} -3\lambda - 1 & -13\lambda - 5 \\ -\lambda + 5 & -11\lambda + 6 \end{pmatrix},$$

$$Q(\lambda) = A_0 B_0^{-1} \lambda + A_0^{(1)} B_0^{-1} = \begin{pmatrix} 3 & 5 \\ 2 & 5 \end{pmatrix} \lambda + \begin{pmatrix} 1 & 2 \\ -2 & -2 \end{pmatrix} = \begin{pmatrix} 3\lambda + 1 & 5\lambda + 2 \\ 2\lambda - 2 & 5\lambda - 2 \end{pmatrix}.$$

Предлагаем читателю в качестве упражнения непосредственно проверить, что

$$A(\lambda) = Q(\lambda) B(\lambda) + R(\lambda).$$

1) Заметим, что возможность и однозначность левого деления $A(\lambda)$ на $B(\lambda)$ следует из возможности и однозначности правого деления транспонированных матриц $A'(\lambda)$ и $B'(\lambda)$. (Из регулярности $B(\lambda)$ вытекает регулярность $B'(\lambda)$.) Действительно, из

$$A'(\lambda) = Q_1(\lambda) B'(\lambda) + R_1(\lambda)$$

следует (см. гл. I, стр. 29):

$$A(\lambda) = B(\lambda) Q'_1(\lambda) + R'_1(\lambda). \quad (11')$$

В силу этих же соображений левое деление $A(\lambda)$ на $B(\lambda)$ однозначно, так как из неоднозначности левого деления $A(\lambda)$ на $B(\lambda)$ следовала бы неоднозначность правого деления $A'(\lambda)$ на $B'(\lambda)$.

Сопоставление (11) и (11') дает:

$$\widehat{Q}(\lambda) = Q'_1(\lambda), \quad \widehat{R}(\lambda) = R'_1(\lambda).$$

Рассмотрим произвольный матричный многочлен n -го порядка

$$F(\lambda) = F_0\lambda^m + F_1\lambda^{m-1} + \dots + F_m \quad (F_0 \neq 0). \quad (17)$$

Разделим его на бином $\lambda E - A$ справа и слева:

$$F(\lambda) = Q(\lambda)(\lambda E - A) + R, \quad F(\lambda) = (\lambda E - A)\widehat{Q}(\lambda) + \widehat{R}. \quad (18)$$

В данном случае правый остаток R и левый остаток \widehat{R} не будут зависеть от λ . Для определения правого значения $F(A)$ и левого $\widehat{F}(A)$ можно соответственно в тождествах (18) заменить скаляр λ на матрицу A , поскольку матрица A перестановочна с матричными коэффициентами бинома $\lambda E - A$ (см. § 1):

$$F(A) = Q(A)(A - A) + R = R, \quad \widehat{F}(A) = (A - A)\widehat{Q}(A) + \widehat{R} = \widehat{R}. \quad (19)$$

Нами доказана

Теорема 1 (обобщенная теорема Безу). *При правом (левом) делении матричного многочлена $F(\lambda)$ на бином $\lambda E - A$ остаток от деления равен $F(A)$ (соответственно $\widehat{F}(A)$).*

Из доказанной теоремы следует, что многочлен $F(\lambda)$ делится без остатка справа (слева) на бином $\lambda E - A$ тогда и только тогда, когда $F(A) = 0$ (соответственно $\widehat{F}(A) = 0$).

Пример. Пусть $A = \|a_{ik}\|_1^n$ и $f(\lambda)$ — многочлен относительно λ . Тогда

$$F(\lambda) = f(\lambda)E - f(A)$$

делится (слева и справа) без остатка на $\lambda E - A$. Это следует непосредственно из обобщенной теоремы Безу, поскольку в данном случае $F(A) = \widehat{F}(A) = 0$.

§ 3. Характеристический многочлен матрицы. Присоединенная матрица

1. Рассмотрим матрицу $A = \|a_{ik}\|_1^n$. *Характеристической матрицей* для матрицы A называется матрица $\lambda E - A$. Определитель характеристической матрицы

$$\Delta(\lambda) = |\lambda E - A| = |\lambda\delta_{ik} - a_{ik}|_1^n$$

представляет собой скалярный многочлен относительно λ и называется *характеристическим многочленом* матрицы A (см. гл. III, § 7)¹⁾.

Матрицу $B(\lambda) = \|b_{ik}(\lambda)\|_1^n$, где $b_{ik}(\lambda)$ — алгебраическое дополнение элемента $\lambda\delta_{ik} - a_{ik}$ в определителе $\Delta(\lambda)$, мы будем называть *присоединенной матрицей* для матрицы A .

Так, например, для матрицы

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

¹⁾ Этот многочлен отличается множителем $(-1)^n$ от многочлена $\Delta(\lambda)$, введенного в гл. III, § 7.

будем иметь:

$$\lambda E - A = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \lambda^3 - (a_{11} + a_{22} + a_{33})\lambda^2 + \dots,$$

$$B(\lambda) = \begin{vmatrix} \lambda^2 - (a_{22} + a_{33})\lambda + a_{22}a_{33} - a_{23}a_{32} & * & * \\ a_{21}\lambda + a_{23}a_{31} - a_{21}a_{33} & * & * \\ a_{31}\lambda + a_{21}a_{32} - a_{22}a_{31} & * & * \end{vmatrix}.$$

Из приведенных определений следуют тождества относительно λ :

$$(\lambda E - A) B(\lambda) = \Delta(\lambda) E, \quad (20)$$

$$B(\lambda) (\lambda E - A) = \Delta(\lambda) E. \quad (20')$$

Правые части этих равенств мы можем рассматривать как многочлены с матричными коэффициентами (каждый из этих коэффициентов равен произведению скаляра на единичную матрицу E). Многочленную матрицу $B(\lambda)$ можно также представить в виде многочлена, расположенного по степеням λ . Равенства (20) и (20') показывают, что $\Delta(\lambda) E$ делится слева и справа на $\lambda E - A$ без остатка. Согласно обобщенной теореме Безу это возможно лишь тогда, когда остаток $\Delta(A) E = \Delta(A)$ равен нулю. Нами доказана

Теорема 2 (Гамильтона — Кэли). *Всякая квадратная матрица A удовлетворяет своему характеристическому уравнению, т. е.*

$$\Delta(A) = 0. \quad (21)$$

Пример.

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

$$\Delta(\lambda) = \begin{vmatrix} \lambda - 2 & -1 \\ 1 & \lambda - 3 \end{vmatrix} = \lambda^2 - 5\lambda + 7,$$

$$\Delta(A) = A^2 - 5A + 7E = \begin{vmatrix} 3 & 5 \\ -5 & 8 \end{vmatrix} - 5 \begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} + 7 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = 0.$$

2. Обозначим через $\lambda_1, \lambda_2, \dots, \lambda_n$ все характеристические числа матрицы A , т. е. все корни характеристического многочлена $\Delta(\lambda)$ (каждое из чисел λ_i повторяется в этом ряду столько раз, сколько его кратность как корня многочлена $\Delta(\lambda)$). Тогда

$$\Delta(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n). \quad (22)$$

Пусть дан произвольный скалярный многочлен $g(\mu)$. Найдем характеристические числа матрицы $g(A)$. Для этого разложим $g(\mu)$ на линейные множители

$$g(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2) \dots (\mu - \mu_l). \quad (23)$$

Подставим в обе части этого тождества вместо μ матрицу A :

$$g(A) = a_0(A - \mu_1 E)(A - \mu_2 E) \dots (A - \mu_l E). \quad (24)$$

Переходя к определителям в обеих частях равенства (24) и используя равенства (22) и (23), получим:

$$\begin{aligned} |g(A)| &= a_0^n |A - \mu_1 E| |A - \mu_2 E| \dots |A - \mu_l E| = \\ &= (-1)^{nl} a_0^n \Delta(\mu_1) \Delta(\mu_2) \dots \Delta(\mu_l) = \\ &= (-1)^{nl} a_0^n \prod_{i=1}^l \prod_{k=1}^n (\mu_i - \lambda_k) = g(\lambda_1) g(\lambda_2) \dots g(\lambda_n). \end{aligned}$$

Заменив в равенстве

$$|g(A)| = g(\lambda_1) g(\lambda_2) \dots g(\lambda_n) \quad (25)$$

многочлен $g(\mu)$ на $\lambda - g(\mu)$, где λ — некоторый параметр, найдем:

$$|\lambda E - g(A)| = [\lambda - g(\lambda_1)] [\lambda - g(\lambda_2)] \dots [\lambda - g(\lambda_n)]. \quad (26)$$

Из этого равенства вытекает следующая

Теорема 3. Если $\lambda_1, \lambda_2, \dots, \lambda_n$ — все характеристические числа (с учетом кратностей) матрицы A , а $g(\mu)$ — некоторый скалярный многочлен, то $g(\lambda_1), g(\lambda_2), \dots, g(\lambda_n)$ — все характеристические числа матрицы $g(A)$.

В частности, если матрица A имеет характеристические числа $\lambda_1, \lambda_2, \dots, \lambda_n$, то матрица A^k имеет характеристические числа $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ ($k = 0, 1, 2, 3, \dots$).

3. Укажем эффективную формулу, выражающую присоединенную матрицу $B(\lambda)$ через характеристический многочлен $\Delta(\lambda)$.

Пусть

$$\Delta(\lambda) = \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_n. \quad (27)$$

Разность $\Delta(\lambda) - \Delta(\mu)$ делится без остатка на $\lambda - \mu$. Поэтому

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^{n-1} + (\mu - p_1) \lambda^{n-2} + (\mu^2 - p_1\mu - p_2) \lambda^{n-3} + \dots \quad (28)$$

есть многочлен относительно λ и μ .

Тождество

$$\Delta(\lambda) - \Delta(\mu) = \delta(\lambda, \mu) (\lambda - \mu) \quad (29)$$

не нарушится, если в него вместо λ и μ подставить перестановочные между собой матрицы λE и A . Тогда, поскольку согласно теореме Гамильтона — Кэли $\Delta(A) = 0$,

$$\Delta(\lambda) E = \delta(\lambda E, A) (\lambda E - A). \quad (30)$$

Сопоставляя между собой равенства (20') и (30), из однозначности частного получаем искомую формулу

$$B(\lambda) = \delta(\lambda E, A). \quad (31)$$

Отсюда в силу (28)

$$B(\lambda) = \lambda^{n-1} + B_1 \lambda^{n-2} + B_2 \lambda^{n-3} + \dots + B_{n-1}, \quad (32)$$

где

$$B_1 = A - p_1 E, \quad B_2 = A^2 - p_1 A - p_2 E, \dots$$

и вообще

$$B_k = A^k - p_1 A^{k-1} - p_2 A^{k-2} - \dots - p_k E \quad (k = 1, 2, \dots, n-1). \quad (33)$$

Матрицы B_1, B_2, \dots, B_{n-1} можно вычислять последовательно, исходя из рекуррентных соотношений

$$B_k = AB_{k-1} - p_k E \quad (k = 1, 2, \dots, n-1; B_0 = E). \quad (34)$$

При этом

$$AB_{n-1} - p_n E = 0^1). \quad (35)$$

Соотношения (34) и (35) непосредственно получаются из тождества (20), если в обеих частях этого тождества приравнять между собой коэффициенты при одинаковых степенях λ .

Если A — неособенная матрица, то

$$p_n = (-1)^{n-1} |A| \neq 0$$

и из (35) следует:

$$A^{-1} = \frac{1}{p_n} B_{n-1}. \quad (36)$$

Пусть λ_0 — характеристическое число матрицы A , т. е. $\Delta(\lambda_0) = 0$. Подставив в (20) вместо λ значение λ_0 , найдем:

$$(\lambda_0 E - A) B(\lambda_0) = 0. \quad (37)$$

Допустим, что матрица $B(\lambda_0) \neq 0$, и обозначим через b любой ненулевой столбец этой матрицы. Тогда из (37) $(\lambda_0 E - A) b = 0$, или

$$Ab = \lambda_0 b. \quad (38)$$

Следовательно, любой ненулевой столбец матрицы $B(\lambda_0)$ определяет собственный вектор, соответствующий характеристическому числу λ_0 ²⁾.

Таким образом, если коэффициенты характеристического многочлена известны, то присоединенная матрица может быть найдена по формуле (31). Если данная матрица A неособенная, то по формуле (36) находится обратная матрица A^{-1} . Если λ_0 — характеристическое число матрицы A , то ненулевые столбцы матрицы $B(\lambda_0)$ являются собственными векторами матрицы A для $\lambda = \lambda_0$.

Пример.

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

$$\Delta(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda - 2 & 1 & -1 \\ 0 & \lambda - 1 & -1 \\ 1 & -1 & \lambda - 1 \end{vmatrix} = \lambda^3 - 4\lambda^2 + 5\lambda - 2,$$

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^2 + \lambda(\mu - 4) + \mu^2 - 4\mu + 5,$$

$$B(\lambda) = \delta(\lambda E, A) = \lambda^2 E + \lambda \underbrace{(A - 4E)}_{B_1} + \underbrace{A^2 - 4A + 5E}_{B_2}.$$

1) Из (34) следуют равенства (33). Если в (35) подставить выражение для B_{n-1} , из (33), то получим: $\Delta(A) = 0$. Этот вывод теоремы Гамильтона — Кэли не опирается явным образом на обобщенную теорему Безу, но неявно содержит в себе эту теорему.

2) См. гл. III, § 7. Если характеристическому числу λ_0 соответствует d_0 линейно независимых собственных векторов ($n - d_0$ — ранг матрицы $\lambda_0 E - A$), то ранг матрицы $B(\lambda_0)$ не превосходит d_0 . В частности, если числу λ_0 соответствует только одно собственное направление, то в матрице $B(\lambda_0)$ элементы любых двух столбцов пропорциональны.

Но

$$B_1 = A - 4E = \begin{vmatrix} -2 & -1 & 1 \\ 0 & -3 & 1 \\ -1 & 1 & -3 \end{vmatrix},$$

$$B_2 = AB_1 + 5E = \begin{vmatrix} 0 & 2 & -2 \\ -1 & +3 & -2 \\ 1 & -1 & 2 \end{vmatrix},$$

$$B(\lambda) = \begin{vmatrix} \lambda^2 - 2\lambda & -\lambda + 2 & \lambda - 2 \\ -1 & \lambda^2 - 3\lambda + 3 & \lambda - 2 \\ -\lambda + 1 & \lambda - 1 & \lambda^2 - 3\lambda + 2 \end{vmatrix},$$

$$|A| = 2, \quad A^{-1} = \frac{1}{2} B_2 = \begin{vmatrix} 0 & 1 & -1 \\ -\frac{1}{2} & \frac{3}{2} & -1 \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{vmatrix}.$$

Далее

$$\Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2)$$

Первый столбец матрицы $B(+1)$ дает собственный вектор $(+1, +1, 0)$ для характеристического числа $\lambda = 1$.

Первый столбец матрицы $B(+2)$ дает собственный вектор $(0, 1, 1)$, соответствующий характеристическому числу $\lambda = 2$.

§ 4. Метод Д. К. Фаддеева одновременного вычисления коэффициентов характеристического многочлена и присоединенной матрицы

Д. К. Фаддеев¹⁾ предложил метод одновременного определения скалярных коэффициентов p_1, p_2, \dots, p_n характеристического многочлена

$$\Delta(\lambda) = \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_n \quad (39)$$

и матричных коэффициентов B_1, B_2, \dots, B_{n-1} присоединенной матрицы $B(\lambda)$.

Для изложения метода Д. К. Фаддеева²⁾ введем понятие о следе матрицы.

Под следом матрицы $A = \|a_{ik}\|_1^n$ (обозначение: $\text{Sp } A$) понимают сумму диагональных элементов этой матрицы:

$$\text{Sp } A = \sum_{i=1}^n a_{ii}. \quad (40)$$

Нетрудно видеть, что

$$\text{Sp } A = p_1 = \sum_{i=1}^n \lambda_i, \quad (41)$$

1) См. [32], стр. 160.

2) С другим эффективным методом вычисления коэффициентов характеристического многочлена, с методом А. Н. Крылова, мы познакомим читателя в гл. VII, § 8.

если $\lambda_1, \lambda_2, \dots, \lambda_n$ — характеристические числа матрицы A , т. е.

$$\Delta(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n). \quad (42)$$

Так как, согласно теореме 3, степень матрицы A^k имеет своими характеристическими числами степени $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ ($k = 0, 1, 2, \dots$), то

$$\text{Sp } A^k = s_k = \sum_{i=1}^n \lambda_i^k \quad (k = 0, 1, 2, \dots). \quad (43)$$

Суммы s_k ($k = 1, 2, \dots, n$) степеней корней многочлена (39) связаны с коэффициентами этого уравнения формулами Ньютона¹⁾

$$kp_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1 \quad (k = 1, 2, \dots, n). \quad (44)$$

Если вычислить следы s_1, s_2, \dots, s_n матриц A, A^2, \dots, A^n , то затем можно из уравнений (44) последовательно определить коэффициенты p_1, p_2, \dots, p_n . В этом состоит метод Леверье определения коэффициентов характеристического многочлена по следам степеней матрицы.

Д. К. Фаддеев предложил вместо следов степеней A, A^2, \dots, A^n вычислить последовательно следы некоторых других матриц A_1, A_2, \dots, A_n и с их помощью определить p_1, p_2, \dots, p_n и B_1, B_2, \dots, B_n следующими формулами:

$$\left. \begin{array}{l} A_1 = A, \quad p_1 = \text{Sp } A_1, \quad B_1 = A_1 - p_1 E, \\ A_2 = AB_1, \quad p_2 = \frac{1}{2} \text{Sp } A_2, \quad B_2 = A_2 - p_2 E, \\ \dots \quad \dots \quad \dots \\ A_{n-1} = AB_{n-2}, \quad p_{n-1} = \frac{1}{n-1} \text{Sp } A_{n-1}, \quad B_{n-1} = A_{n-1} - p_{n-1} E, \\ A_n = AB_{n-1}, \quad p_n = \frac{1}{n} \text{Sp } A_n, \quad B_n = A_n - p_n E = 0. \end{array} \right\} \quad (45)$$

Последнее равенство $B_n = A_n - p_n E = 0$ может быть использовано для контроля вычислений.

Для того чтобы убедиться, что числа p_1, p_2, \dots, p_n и матрицы B_1, B_2, \dots, B_{n-1} , последовательно определяемые по формулам (45), являются коэффициентами $\Delta(\lambda)$ и $B(\lambda)$, заметим, что из (45) вытекают следующие формулы для A_k и B_k ($k = 1, 2, \dots, n$):

$$A_k = A^k - p_1 A^{k-1} - \dots - p_{k-1} A - p_k E, \quad B_k = A^k - p_1 A^{k-1} - \dots - p_{k-1} A - p_k E. \quad (46)$$

Приравняем между собой следы левой и правой частей первой из этих формул; получим:

$$kp_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1.$$

Но эти формулы совпадают с формулами Ньютона (44), по которым последовательно определяются коэффициенты характеристического многочлена $\Delta(\lambda)$. Следовательно, числа p_1, p_2, \dots, p_n , определяемые по формулам (45), и являются коэффициентами $\Delta(\lambda)$. Но тогда вторые формулы (46) совпадают с формулами (33), по которым определяются матричные коэффициенты B_1, B_2, \dots, B_{n-1} присоединенной матрицы $B(\lambda)$. Следовательно, формулы (45) определяют и коэффициенты B_1, B_2, \dots, B_{n-1} матричного многочлена $B(\lambda)$.

1) См., например, [20], стр. 224.

Пример 1).

$$\begin{aligned}
 A &= \begin{vmatrix} 2 & -1 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 4 & 3 \end{vmatrix}, \quad p_1 = \text{Sp } A = 4, \quad B_1 = A - 4E = \begin{vmatrix} -2 & -1 & 1 & 2 \\ 0 & -3 & 1 & 0 \\ -1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -4 \end{vmatrix}, \\
 A_2 = AB_1 &= \begin{vmatrix} -3 & 4 & 0 & -3 \\ -1 & -2 & -2 & 1 \\ 2 & 0 & -2 & -5 \\ -3 & -3 & -1 & 3 \\ -5 & -1 & -5 & -4 \end{vmatrix}, \quad p_2 = \frac{1}{2} \text{Sp } A_2 = -2, \quad B_2 = A_2 + 2E = \begin{vmatrix} -1 & 4 & 0 & -3 \\ -1 & 0 & -2 & 1 \\ 2 & 0 & 0 & -5 \\ -3 & -3 & -1 & 3 \end{vmatrix}; \\
 A_3 = AB_2 &= \begin{vmatrix} -5 & 2 & 0 & -2 \\ 1 & 0 & -2 & -4 \\ -1 & -7 & -3 & 4 \\ 0 & 4 & -2 & -7 \\ -5 & -1 & -7 & -9 \end{vmatrix}, \quad p_3 = \frac{1}{3} \text{Sp } A_3 = -5, \quad B_3 = A_3 + 5E = \begin{vmatrix} 0 & 2 & 0 & -2 \\ 1 & 5 & -2 & 4 \\ -1 & -7 & 2 & 4 \\ 0 & 4 & -2 & -2 \end{vmatrix}; \\
 A_4 = AB_3 &= \begin{vmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{vmatrix}, \quad p_4 = -2, \\
 \Delta(\lambda) &= \lambda^4 - 4\lambda^3 + 2\lambda^2 + 5\lambda + 2, \\
 |A| &= 2, \quad A^{-1} = \frac{1}{p_4} B_3 = \begin{vmatrix} 0 & -1 & 0 & 1 \\ -\frac{1}{2} & -\frac{5}{2} & 1 & -2 \\ \frac{1}{2} & \frac{7}{2} & -1 & -2 \\ 0 & -2 & 1 & 1 \end{vmatrix}.
 \end{aligned}$$

Замечание. Если мы хотим определить p_1, p_2, p_3, p_4 и только первые столбцы в B_1, B_2, B_3 , то достаточно вычислить в A_2 элементы 1-го столбца и только диагональные элементы остальных столбцов, в A_3 — только элементы 1-го столбца, в A_4 — только два первых элемента 1-го столбца.

§ 5. Минимальный многочлен матрицы

Определение 1. Скалярный многочлен $f(\lambda)$ называется *аннулирующим многочленом квадратной матрицы* A , если

$$f(A) = 0.$$

Аннулирующий многочлен $\phi(\lambda)$ наименьшей степени со старшим коэффициентом, равным единице, называется *минимальным многочленом* матрицы A .

Согласно теореме Гамильтона — Кэли характеристический многочлен $\Delta(\lambda)$ матрицы A является аннулирующим для этой матрицы. Однако, как будет показано ниже, в общем случае он не является минимальным.

1) Для контроля за вычислениями мы под каждой из матриц A_1, A_2, A_3 подпisyvаем ее суммарную строку. Произведения суммарной строки первого сомножителя на столбцы второго должны дать элементы суммарной строки произведения.

Разделим произвольный аннулирующий многочлен $f(\lambda)$ на минимальный:

$$f(\lambda) = \psi(\lambda) q(\lambda) + r(\lambda),$$

где степень $r(\lambda)$ меньше степени $\psi(\lambda)$. Отсюда имеем:

$$f(A) = \psi(A) q(A) + r(A).$$

Поскольку $f(A) = 0$ и $\psi(A) = 0$, то, значит, и $r(A) = 0$. Но степень $r(\lambda)$ меньше степени минимального многочлена $\psi(\lambda)$. Поэтому $r(\lambda) \equiv 0^1$. Таким образом, *произвольный аннулирующий многочлен матрицы всегда делится без остатка на ее минимальный многочлен*.

Пусть два многочлена $\psi(\lambda)$ и $\psi^1(\lambda)$ являются минимальными для одной и той же матрицы. Тогда каждый из них делится на другой многочлен без остатка, т. е. эти многочлены отличаются постоянным множителем. Этот постоянный множитель равен единице, поскольку равны единице старшие коэффициенты в $\psi(\lambda)$ и $\psi^1(\lambda)$. Мы доказали *единственность минимального многочлена для данной матрицы A*.

Выведем формулу, связывающую минимальный многочлен с характеристическим.

Обозначим через $D_{n-1}(\lambda)$ наибольший общий делитель всех миноров $(n-1)$ -го порядка характеристической матрицы $\lambda E - A$, т. е. наибольший общий делитель всех элементов присоединенной матрицы $B(\lambda) = \|b_{ik}(\lambda)\|_1^n$ (см. предыдущий параграф); при этом старший коэффициент в $D_{n-1}(\lambda)$ берем равным единице. Тогда

$$B(\lambda) = D_{n-1}(\lambda) C(\lambda), \quad (47)$$

где $C(\lambda)$ — некоторая многочленная матрица, «приведенная» присоединенная матрица для $\lambda E - A$. Из (20) и (47) находим:

$$\Delta(\lambda) E = (\lambda E - A) C(\lambda) D_{n-1}(\lambda). \quad (48)$$

Отсюда следует, что $\Delta(\lambda)$ делится без остатка на $D_{n-1}(\lambda)^2$:

$$\frac{\Delta(\lambda)}{D_{n-1}(\lambda)} = \psi(\lambda), \quad (49)$$

где $\psi(\lambda)$ — некоторый многочлен. Обе части тождества (48) можно сократить на $D_{n-1}(\lambda)^2$:

$$\psi(\lambda) E = (\lambda E - A) C(\lambda). \quad (50)$$

Поскольку $\psi(\lambda) E$ делится без остатка слева на $\lambda E - A$, то в силу обобщенной теоремы Безу

$$\psi(A) = 0.$$

Таким образом, многочлен $\psi(\lambda)$, определенный формулой (49), является аннулирующим многочленом для матрицы A . Докажем, что он является минимальным многочленом.

1) В противном случае существовал бы аннулирующий многочлен, степень которого была бы меньше степени минимального многочлена.

2) В этом можно и непосредственно убедиться, разлагая характеристический определитель $\Delta(\lambda)$ по элементам какой-либо строки.

3) В данном случае наряду с (50) имеет место и тождество [см. (20')]

$$\psi(\lambda) E = C(\lambda) (\lambda E - A),$$

т. е. $C(\lambda)$ является одновременно и левым и правым частным от деления $\psi(\lambda) E$ на $\lambda E - A$.

Обозначим минимальный многочлен через $\psi^*(\lambda)$. Тогда $\psi(\lambda)$ делится без остатка на $\psi^*(\lambda)$:

$$\psi(\lambda) = \psi^*(\lambda) \chi(\lambda). \quad (51)$$

Поскольку $\psi^*(A) = 0$, то в силу обобщенной теоремы Безу матричный многочлен $\psi^*(\lambda) E$ делится слева без остатка на $\lambda E - A$:

$$\psi^*(\lambda) E = (\lambda E - A) C^*(\lambda). \quad (52)$$

Из (51) и (52) следует:

$$\psi(\lambda) E = (\lambda E - A) C^*(\lambda) \chi(\lambda). \quad (53)$$

Тождества (50) и (53) показывают, что как $C(\lambda)$, так и $C^*(\lambda) \chi(\lambda)$ являются левыми частными при делении $\psi(\lambda) E$ на $\lambda E - A$. В силу однозначности деления

$$C(\lambda) = C^*(\lambda) \chi(\lambda).$$

Отсюда следует, что $\chi(\lambda)$ является общим делителем всех элементов многочленной матрицы $C(\lambda)$. Но, с другой стороны, наибольший общий делитель всех элементов приведенной присоединенной матрицы $C(\lambda)$ равен единице, поскольку эта матрица была получена из $B(\lambda)$ путем деления на $D_{n-1}(\lambda)$. Поэтому $\chi(\lambda) = \text{const}$. Так как старшие коэффициенты в $\psi(\lambda)$ и $\psi^*(\lambda)$ равны единице, то в (51) $\chi(\lambda) = 1$, т. е. $\psi(\lambda) = \psi^*(\lambda)$, что и требовалось доказать.

Мы установили следующую формулу для минимального многочлена:

$$\psi(\lambda) = \frac{\Delta(\lambda)}{D_{n-1}(\lambda)}. \quad (54)$$

Для приведенной присоединенной матрицы $C(\lambda)$ имеем формулу, аналогичную формуле (31) (на стр. 94):

$$C(\lambda) = \Psi(\lambda E, A), \quad (55)$$

где многочлен $\Psi(\lambda, \mu)$ определяется равенством

$$\Psi(\lambda, \mu) = \frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu} \quad (56)$$

Кроме того,

$$(\lambda E - A) C(\lambda) = \psi(\lambda) E. \quad (57)$$

Переходя к определителям в обеих частях равенства (57), получаем:

$$\Delta(\lambda) | C(\lambda) | = [\psi(\lambda)]^n. \quad (58)$$

Таким образом, $\Delta(\lambda)$ делится без остатка на $\psi(\lambda)$, а некоторая степень $\psi(\lambda)$ делится без остатка на $\Delta(\lambda)$, т. е. совокупность всех различных между собой корней у многочлена $\Delta(\lambda)$ и $\psi(\lambda)$ одна и та же. Другими словами, корнями $\psi(\lambda)$ служат все различные между собой характеристические числа матрицы A .

Если

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \dots (\lambda - \lambda_s)^{n_s} \quad (59)$$

$$(\lambda_i \neq \lambda_j \text{ при } i \neq j; n_i > 0, i, j = 1, 2, \dots, s),$$

¹⁾ Формула (55) выводится совершенно так же, как и формула (31). В обе части тождества $\psi(\lambda) - \psi(\mu) = (\lambda - \mu) \Psi(\lambda, \mu)$ вместо λ и μ подставляются матрицы λE и A и полученное матричное равенство сопоставляется с (50).