ЛИНЕЙНАЯ АЛГЕБРА

Выпуск 2

Учебно-методическое пособие для студентов I *курса*

Настоящий выпуск 2 учебно-методического пособия по линейной алгебре представляет собой единое целое с выпуском 1. Здесь содержится окончание гл. 1 "Матрицы и определители", а также гл. 2 "Системы линейных алгебраических уравнений". Этим исчерпывается "конкретная" часть курса линейной алгебры. Последующие главы должны содержать более абстрактный материал: теорию конечномерных векторных пространств и линейных операторов в них.

Напомним, что запись $a := b$ означает определение величины, стоящей слева. Значок • означает конец доказательства. Дополнительный материал помечен верхним значком $*$. Как правило, этот материал не входит в лекционный курс. При ссылках на формулы, теоремы и пункты из другого параграфа применяется двойная нумерация, а из другой главы — тройная.

Глава 1. Матрицы и определители. *Окончание*

§ 7. Обратная матрица. Формулы Крамера. Метод Гаусса

1. Понятие обратной матрицы. Пусть $A \in M^{m,n}$. Матрица $B \in M^{n,m}$ называется *левой обратной* для A, если $BA = I_n$. Матрица $C \in M^{n,m}$ называется *правой обратной* для A, если $AC = I_m$. Сразу же отметим (см. ниже п. 2.6.1), что при $m \neq n$ матрица $A \in M^{m,n}$ не может иметь одновременно левую и правую обратную. Иначе обстоит дело для **квадратных матриц**, *о которых и пойдет речь до конца этой главы.*

Пусть $A \in M^n$ и пусть существует левая обратная, $BA = I$. Тогда $\det A \neq 0$. Действительно, по теореме умножения определителей, $1 = \det I = \det B \det A$, а потому $\det A \neq 0$. Аналогично, если существует правая обратная, $AC = I$, то $\det A \neq 0$. Матрицу с отличным от нуля определителем называют *неособой*. Предположим теперь, что $A \in M^n$ имеет как левую, так и правую обратные матрицы:

$$BA = AC = I. \tag{1}$$

Тогда $B = C$. Действительно, $B = BI = B(AC) = (BA)C = IC = C$. Далее, при условиях (1) как правая обратная, так и левая обратная, может быть только одна. Это следует из того, что, как мы видели, каждая левая обратная матрица совпадает с каждой правой обратной. Наконец, если выполнено (1), то $B = C$ является *двусторонней обратной* матрицей. Такая матрица (если она есть) единственна. Ее обозначают через A^{-1} и называют просто *обратной* матрицей:

$$A^{-1}A = AA^{-1} = I. \tag{2}$$

Из (2) сразу следует, что существует $(A^{-1})^{-1} = A$. Отметим, что $I^{-1} = I$. Затем, если существуют A_1^{-1} и A_2^{-1}, то A_1A_2 имеет обратную и

$$(A_1A_2)^{-1} = A_2^{-1}A_1^{-1}. \tag{3}$$

Действительно, $(A_2^{-1}A_1^{-1})(A_1A_2) = A_2^{-1}(A_1^{-1}A_1)A_2 = A_2^{-1}IA_2 = A_2^{-1}A_2 = I$ и, аналогично, $(A_1A_2)(A_2^{-1}A_1^{-1}) = I$. Далее, вместе с A^{-1} существует

$$(A^t)^{-1} = (A^{-1})^t, \tag{4}$$

поскольку из $AA^{-1} = I$ следует $(A^{-1})^tA^t = I$, а из $A^{-1}A = I$ вытекает $A^t(A^{-1})^t = I$. Аналогично проверяется, что

$$(A^*)^{-1} = (A^{-1})^*. \tag{5}$$

Наконец, из (2) прямо следует равенство $(\det A^{-1})(\det A) = 1$, а потому

$$\det A^{-1} = (\det A)^{-1}. \tag{6}$$

Наиболее содержательным результатом относительно обратных матриц является

Теорема 1. *Пусть $A \in M^n$. Следующие утверждения равносильны:*

1) *У матрицы A существует левая обратная.*

2) *У матрицы A существует правая обратная.*

3) *У матрицы A существует двусторонняя обратная A^{-1}.*

4) *Матрица A — неособая:* $\det A \neq 0$.

Мы уже видели, что 1) $\Rightarrow$ 4) и 2) $\Rightarrow$ 4). Ясно также, что 3) $\Rightarrow$ 1) и 3) $\Rightarrow$ 2). Остается проверить, что 4) $\Rightarrow$ 3). В отличие от предыдущих, последняя импликация доказывается на основании явной конструкции: для A^{-1} в следующем пункте будут предъявлены формулы. Эти формулы, конечно, имеют самостоятельное значение.

2. Присоединенная матрица. Построение обратной для неособой матрицы. Будем исходить из формул (5.14), которые выпишем заново:

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \sum_{j=1}^{n} a_{ji} A_{jk} = \delta_{ik} \det A, \quad i, k = 1, \ldots, n. \tag{7}$$

Здесь $\{a_{ij}\} = A \in M^n$, A_{jk} — алгебраическое дополнение элемента a_{jk} матрицы A. Введем в рассмотрение матрицу $\check{A}$ с элементами $[\check{A}]_{jk} = A_{kj}$. Таким образом, матрица $\check{A}$ (ее называют *присоединенной* для матрицы A) есть транспонированная матрица алгебраических дополнений для элементов матрицы A. Ясно, что соотношения (7) теперь можно записать в виде

$$A\check{A} = \check{A}A = (\det A)\, I. \tag{8}$$

Если матрица A — неособая, т. е. $\det A \neq 0$, то из (8) следует:

$$AX = XA = I, \quad \text{где} \quad X = (\det A)^{-1}\check{A}.$$

Таким образом,

$$A^{-1} = (\det A)^{-1}\check{A}, \quad (\det A \neq 0). \tag{9}$$

Одновременно с (9) мы установили импликацию 4) $\Rightarrow$ 3) и этим *закончили доказательство теоремы* 1. Для обратной матрицы (9) выполнены соотношения (3)–(6).

Пример. Для неособых матриц второго порядка

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = (ad-bc)^{-1}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad ad-bc \neq 0.$$

3. Линейные системы с неособой квадратной матрицей коэффициентов. Формулы Крамера. Рассмотрим систему n линейных уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = f_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = f_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = f_n. \end{cases} \tag{10}$$

Систему (10) запишем в матричной форме

$$A\mathbf{x} = \mathbf{f}, \quad A \in M^n, \tag{11}$$

где $\mathbf{x}$ — столбец неизвестных величин, а $\mathbf{f}$ — столбец правых частей. Предположим, что матрица A — неособая,

$$\det A \neq 0. \tag{12}$$

Теорема 2. *При условии* (12) *уравнение* (11) *имеет решение* $\mathbf{x} = A^{-1}\mathbf{f}$ *при любой правой части* $\mathbf{f}$. *Это решение единственно.*

Доказательство. Действительно, $A\mathbf{x} = A(A^{-1}\mathbf{f}) = (AA^{-1})\mathbf{f} = \mathbf{f}$, т. е. $\mathbf{x} = A^{-1}\mathbf{f}$ — решение уравнения (11). Далее, если $\mathbf{x}$ — решение (11), то $A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{f}$, $(A^{-1}A)\mathbf{x} = A^{-1}\mathbf{f}$ и $\mathbf{x} = A^{-1}\mathbf{f}$. Таким образом, решение единственно. •

Формула (9) для A^{-1} позволяет описать решение уравнения (11) в более конкретной форме. Обозначим через Δ_k определитель матрицы, получаемой из матрицы A заменой столбца с номером k столбцом $\mathbf{f}$:

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,k-1} & f_1 & a_{1,k+1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2,k-1} & f_2 & a_{2,k+1} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n,k-1} & f_n & a_{n,k+1} & \dots & a_{nn} \end{vmatrix}, \quad k = 1, \dots, n. \tag{13}$$

Теорема 3. (Формулы Крамера). *Решение системы* (10) *при условии* (12) *дается формулами*

$$x_k = \Delta_k / \Delta, \quad \Delta = \det A, \quad k = 1, \dots, n. \tag{14}$$

Доказательство. В силу теоремы 2 и формулы (9),

$$x_k = (A^{-1}\mathbf{f})_k = \Delta^{-1} \sum_{j=1}^{n} [\check{A}]_{kj} f_j = \Delta^{-1} \sum_{j=1}^{n} A_{jk} f_j. \tag{15}$$

Сопоставим два выражения

$$\sum_{j=1}^{n} A_{jk} a_{jk}, \quad \sum_{j=1}^{n} A_{jk} f_j.$$

Первое из них есть разложение $\det A$ по элементам столбца с номером k. Поэтому второе есть разложение определителя Δ_k по элементам столбца с тем же номером. Отсюда следует, что правая часть в (15) совпадает с Δ_k/Δ. •

Формулы Крамера (14) имеют, в основном, теоретическое значение. Если n велико, их использование в практических вычислениях невыгодно. К меньшему объему вычислений приводят способы решения, опирающиеся на метод Гаусса (метод последовательного исключения неизвестных).

4. Метод Гаусса. Рассмотрим систему (10) (в матричной записи — уравнение (11)) при условии (12). Будем опираться на следующее очевидное предложение.

Предложение 4. *Система* (10) *перейдет в эквивалентную, если ее подвергнуть одному из следующих преобразований:* 1) *умножить одно из уравнений системы на ненулевую постоянную;* 2) *прибавить к одному из уравнений системы какую-либо линейную комбинацию остальных уравнений системы;* 3) *поменять какие-либо уравнения системы местами.*

Упомянутые в предложении 4 преобразования системы будем называть *допустимыми*. Легко понять, что *условие* (12) *сохраняется при допустимых преобразованиях*.

При записи и преобразованиях системы нет надобности явно выписывать неизвестные $x_1, \ldots, x_n$. Достаточно выписать *расширенную матрицу системы*

$$(A|\mathbf{f}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & f_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & f_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & f_n \end{array}\right) \qquad (16)$$

и проводить допустимые преобразования над ее строками.

Теорема 5. *Если допустимыми преобразованиями строк матрица* (16) *переведена в матрицу* $(I|\mathbf{b})$, *то столбец* $\mathbf{b}$ *есть решение уравнения* (11).

Доказательство. Переводя матрицу $(A|\mathbf{f})$ допустимыми преобразованиями в матрицу $(I|\mathbf{b})$, мы переведем уравнение (11) в эквивалентное уравнение $I\mathbf{x} = \mathbf{b}$. Это и означает, что $x_k = b_k$, $k = 1, \ldots, n$. ●

Опишем теперь общую схему метода Гаусса. Поскольку определитель матрицы системы (11) отличен от нуля, среди элементов ее первого столбца есть ненулевой. Перестановкой уравнений (строк расширенной матрицы) можно добиться того, чтобы $a_{11} \neq 0$. После этого из второго уравнения вычтем первое, умноженное на a_{21}/a_{11}, из третьего — первое, умноженное на a_{31}/a_{11} и так далее. При таком преобразовании коэффициенты при x_1 во всех уравнениях, кроме первого, обратятся в нуль, то есть система перейдет в эквивалентную систему с расширенной матрицей вида

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & f_1 \\ 0 & \widetilde{a}_{22} & \dots & \widetilde{a}_{2n} & \widetilde{f}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \widetilde{a}_{n2} & \dots & \widetilde{a}_{nn} & \widetilde{f}_n \end{array}\right).$$

Определитель матрицы полученной системы отличен от нуля, а потому среди коэффициентов $\widetilde{a}_{22}, \dots, \widetilde{a}_{n2}$ есть ненулевой. Считая $\widetilde{a}_{22} \neq 0$ (этого можно добиться перестановкой строк) и вычитая второе уравнение, умноженное на подходящие множители, из всех последующих, мы исключим x_2 из всех уравнений, начиная с третьего. Действуя подобным образом дальше, найдем, что исходная система эквивалентна системе с треугольной матрицей $\widetilde{A}$ и расширенной матрицей

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1,n-1} & a_{1n} & f_1 \\ 0 & \widetilde{a}_{22} & \widetilde{a}_{23} & \dots & \widetilde{a}_{2,n-1} & \widetilde{a}_{2n} & \widetilde{f}_2 \\ 0 & 0 & \widetilde{a}_{33} & \dots & \widetilde{a}_{3,n-1} & \widetilde{a}_{3n} & \widetilde{f}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \widetilde{a}_{n-1,n-1} & \widetilde{a}_{n-1,n} & \widetilde{f}_{n-1} \\ 0 & 0 & 0 & \dots & 0 & \widetilde{a}_{nn} & \widetilde{f}_n \end{array}\right).$$

Мы исключили из всех уравнений переменные с номерами, меньшими номера уравнения. При этом $\det \widetilde{A} \neq 0$, так что $a_{11}\widetilde{a}_{22} \dots \widetilde{a}_{nn} \neq 0$.

Описанная процедура носит название *прямого хода* метода Гаусса. *Обратный ход* заключается в том, чтобы допустимыми преобразованиями свести систему к системе с диагональной матрицей. Последнее уравнение системы теперь приняло вид $\widetilde{a}_{nn}x_n = \widetilde{f}_n$, а потому $x_n = \widetilde{f}_n/\widetilde{a}_{nn}$. Зная x_n, из предпоследнего уравнения определяем x_{n-1}: $\widetilde{a}_{n-1,n-1}x_{n-1} = \widetilde{f}_{n-1} - \widetilde{a}_{n-1,n}x_n$. Таким же образом последовательно вычисляем $x_{n-2}, \dots, x_1$. Эту процедуру можно описать следующим образом: сперва из уравнений с номерами $k = 1, \ldots, n-1$ вычтем последнее, умноженное на коэффициенты $\widetilde{a}_{kn}/\widetilde{a}_{nn}$, тем самым исключив переменную x_n из всех уравнений системы, кроме последнего. Тогда система изобразится расширенной матрицей вида

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & \ldots & a_{1,n-1} & 0 & \widetilde{f}_1 \\ 0 & \widetilde{a}_{22} & \widetilde{a}_{23} & \ldots & \widetilde{a}_{2,n-1} & 0 & \widetilde{f}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \widetilde{a}_{n-1,n-1} & 0 & \widetilde{f}_{n-1} \\ 0 & 0 & 0 & \ldots & 0 & \widetilde{a}_{nn} & \widetilde{f}_n \end{array}\right).$$

Аналогичным образом вычитаем предпоследнее уравнение (оно содержит только переменную x_{n-1}) с подходящими множителями из уравнений с номерами $1, \ldots, n-2$, исключая из них x_{n-1}. Далее проводим описанную процедуру, пока из каждого уравнения не окажутся исключенными неизвестные с номерами, *большими* номера уравнения. Полученная система имеет диагональную матрицу, а расширенная матрица принимает вид

$$\left(\begin{array}{cccc|c} a_{11} & 0 & \ldots & 0 & g_1 \\ 0 & \widetilde{a}_{22} & \ldots & 0 & g_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \widetilde{a}_{nn} & g_n \end{array}\right).$$

Отметим, что процедура обратного хода метода Гаусса не изменяет диагональных элементов матрицы системы. Теперь можно разделить первое уравнение на $a_{11} \neq 0$, второе — на $\widetilde{a}_{22}$ и так далее. При этом получится система, эквивалентная исходной, с расширенной матрицей вида

$$\left(\begin{array}{cccc|c} 1 & 0 & \ldots & 0 & b_1 \\ 0 & 1 & \ldots & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & b_n \end{array}\right). \tag{17}$$

По теореме 5 это означает, что $x_k = b_k$, $k = 1, \dots, n$. Таким образом, метод Гаусса привел к решению исходной системы. При этом матрицу (17) можно записать в виде $(I|A^{-1}\mathbf{f})$.

Замечание. При выполнении допустимых преобразований над матрицей $(A|\mathbf{f})$ все действия определялись лишь матрицей A. Вектор $\mathbf{f}$ оставался "пассивным" объектом преобразований.

5. Вычисление обратной матрицы методом Гаусса. Начнем со следующей леммы.

Лемма 6. *Пусть* $A \in M^{m,l}$, $B \in M^{l,n}$, $C \in M^{m,n}$. *Пусть* $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathbb{C}^l$ *— последовательные столбцы матрицы* B *и* $\mathbf{c}_1, \dots, \mathbf{c}_n \in \mathbb{C}^m$ *— последовательные столбцы матрицы* C. *Тогда равенство* $C = AB$ *равносильно соотношениям*

$$\mathbf{c}_k = A\mathbf{b}_k, \quad k = 1, \dots, n. \tag{18}$$

Доказательство. Произведение AB строится по правилу "строка на столбец" (см. п. 1.4, формулы (1.9)). Это правило равносильно набору формул (18). •

Следствие 7. *Пусть* $A \in M^n$, $\det A \neq 0$, *и пусть* $\mathbf{e}_1, \dots, \mathbf{e}_n$ *— стандартные орты в* $\mathbb{C}^n$. *Пусть* $\mathbf{c}_k$ *— решение уравнения* $A\mathbf{c}_k = \mathbf{e}_k$, $k = 1, \dots, n$, *и пусть* $C = (\mathbf{c}_1, \dots, \mathbf{c}_n)\,(\in M^n)$. *Тогда* $C = A^{-1}$.

Доказательство. Поскольку $\mathbf{c}_k = A^{-1}\mathbf{e}_k$, $k = 1, \dots, n$, и $I = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, применима лемма 6 с заменой A на A^{-1} и B на I. Тогда $C = A^{-1}I = A^{-1}$. •

Будем теперь рассматривать матрицы класса $M^{n,2n}$, записывая их в виде $(A|B)$, где $A, B \in M^n$.

Теорема 8. *Пусть* $A \in M^n$, $\det A \neq 0$, *и матрица* $(A|I)$ *приведена допустимыми преобразованиями строк к виду* $(I|C)$. *Тогда* $C = A^{-1}$.

Доказательство. Запишем исходную матрицу в виде $(A|\mathbf{e}_1, \dots, \mathbf{e}_n)$. Приводя ее допустимыми преобразованиями к виду $(I|\mathbf{c}_1, \dots, \mathbf{c}_n)$, мы, в силу теоремы 5, *одновременно* найдем все векторы $\mathbf{c}_k = A^{-1}\mathbf{e}_k$, $k = 1, \dots, n$. Тогда следствие 7 означает, что $(\mathbf{c}_1, \dots, \mathbf{c}_n) = C = A^{-1}$. •

Формулировка теоремы 8 носит условный характер. Однако описанная в п. 4 схема метода Гаусса показывает, что приведение $(A|I)$ к виду $(I|C)$ действительно возможно. При этом *одновременно* (ср. замечание в конце п. 4) *вычисляются все столбцы матрицы* C. Такая схема довольно экономна в вычислительном отношении.

Пример. $A = \begin{pmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{pmatrix}$. Найдем матрицу A^{-1}. Расширенная матрица системы для определения обратной матрицы имеет вид

$$\left(\begin{array}{ccc|ccc} 2 & 7 & 3 & 1 & 0 & 0 \\ 3 & 9 & 4 & 0 & 1 & 0 \\ 1 & 5 & 3 & 0 & 0 & 1 \end{array}\right).$$

Переставим третью строку на первое место, а из двух остальных вычтем ее с коэффициентами 2 и 3 соответственно. Получим:

$$\left(\begin{array}{rrr|rrr} 1 & 5 & 3 & 0 & 0 & 1 \\ 0 & -3 & -3 & 1 & 0 & -2 \\ 0 & -6 & -5 & 0 & 1 & -3 \end{array}\right).$$

Теперь из третьей строки вычтем удвоенную вторую

$$\left(\begin{array}{rrr|rrr} 1 & 5 & 3 & 0 & 0 & 1 \\ 0 & -3 & -3 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right).$$

Прямой ход завершен. Обратный ход: к первой и второй строкам прибавляем третью, умноженную на 3 и -3 соответственно:

$$\left(\begin{array}{rrr|rrr} 1 & 5 & 0 & 6 & -3 & -2 \\ 0 & -3 & 0 & -5 & 3 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right).$$

А теперь умножим первую строку на 3 и прибавим к ней вторую, умноженную на 5:

$$\left(\begin{array}{rrr|rrr} 3 & 0 & 0 & -7 & 6 & -1 \\ 0 & -3 & 0 & -5 & 3 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right).$$

Умножим первую строку на 1/3, а вторую — на $-1/3$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -7/3 & 2 & -1/3 \\ 0 & 1 & 0 & 5/3 & -1 & -1/3 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right).$$

Правый блок последней матрицы есть A^{-1}.

Упражнения. 1) Покажите, что решение матричного уравнения $AX = B$, где $A, B \in M^n$, $\det A \neq 0$, (равное $X = A^{-1}B$) может быть найдено преобразованием расширенной матрицы $(A|B)$ к виду $(I|X)$.

2) Как решить матричное уравнение $YA = B$?

§ 8. Характеристический многочлен и спектр квадратной матрицы. Функции от матриц

1. Определение характеристического многочлена. Наряду с квадратной матрицей $A \in M^n$ рассмотрим семейство матриц $A - \lambda I$, где λ — числовой параметр, вообще говоря, комплексный.

Определение. *Характеристическим многочленом* матрицы A называется многочлен от переменной λ: $d_A(\lambda) := \det(A - \lambda I)$.

Легко видеть, что $d_A(\lambda)$ — многочлен степени n. Действительно,

$$d_A(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{22} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix}, \qquad (1)$$

поэтому только один моном, входящий в $\det(A - \lambda I)$, будет содержать λ в степенях n и $n - 1$, а именно моном, равный произведению всех диагональных элементов матрицы $A - \lambda I$. Все остальные мономы содержат по крайней мере по два недиагональных элемента $A - \lambda I$ и, соответственно, имеют по λ степень не выше $n - 2$. Таким образом, два старших коэффициента (при λ^n и λ^{n-1}) у полинома $d_A(\lambda)$ — те же, что у монома $(a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda)$ и равны $(-1)^n$ и $(-1)^{n-1} \sum_{k=1}^{n} a_{kk} = (-1)^{n-1} \mathrm{Tr}\, A$ соответственно. Свободный член полинома $d_A(\lambda)$ совпадает со значением $d_A(0)$ и равен $\det A$. Итак,

$$d_A(\lambda) = \det (A - \lambda I) = \delta_n \lambda^n + \delta_{n-1} \lambda^{n-1} + \ldots + \delta_1 \lambda + \delta_0,$$

$$\delta_n = (-1)^n; \quad \delta_{n-1} = (-1)^{n-1} \mathrm{Tr}\, A; \quad \delta_0 = \det A. \tag{2}$$

Выражения для других коэффициентов — более громоздкие, и мы их здесь не приводим.

2. Собственные значения квадратной матрицы. В соответствии с основной теоремой алгебры, характеристический полином матрицы $A \in M^n$ может быть представлен в виде

$$d_A(\lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n), \tag{3}$$

где $\lambda_1, \ldots, \lambda_n$ — корни $d_A(\lambda)$, вообще говоря, комплексные. Корни многочлена $d_A(\lambda)$ называют *собственными значениями* матрицы A. Среди них могут быть совпадающие. В этом случае применяют и другой способ нумерации: обозначают через $\mu_1, \ldots, \mu_p$, $p \leq n$, все *различные* корни $d_A(\lambda)$, а через $\sigma_1, \ldots, \sigma_p$ их кратности (так называемые *алгебраические* кратности). Тогда

$$d_A(\lambda) = (-1)^n (\lambda - \mu_1)^{\sigma_1} (\lambda - \mu_2)^{\sigma_2} \ldots (\lambda - \mu_p)^{\sigma_p},$$

$$\sigma_1 + \sigma_2 + \ldots + \sigma_p = n.$$

Если $\sigma_k = 1$, то собственное значение μ_k называют *простым*. Коэффициенты характеристического многочлена можно выразить через собственные значения. В частности, из (3) следует, что коэффициент δ_{n-1} при λ^{n-1} равен $(-1)^{n-1} \sum_{k=1}^{n} \lambda_k$, а свободный член δ_0 равен $\prod_{k=1}^{n} \lambda_k$. Учитывая (2), видим, что

$$\mathrm{Tr}\, A = \sum_{k=1}^{n} \lambda_k \ , \quad \det A = \prod_{k=1}^{n} \lambda_k.$$

Как следствие, получаем, что *матрица A — неособая в том и только в том случае, когда все ее собственные значения отличны от нуля.*

Совокупность собственных значений $\mu_1, \ldots, \mu_p$ матрицы A (вместе с их кратностями $\sigma_1, \ldots, \sigma_p$) называют *спектром* матрицы A. Таким образом, матрица A — неособая тогда и только тогда, когда число 0 не принадлежит ее спектру. Отыскание спектра матрицы требует решения алгебраического уравнения $d_A(\lambda) = 0$ степени n. Иногда эта задача упрощается. В частности, если матрица A — *треугольная*, то ее *собственные значения совпадают с элементами главной диагонали*. Это видно непосредственно из определения (1) для d_A.

3. Функции от квадратной матрицы. По *определению* положим $A^2 = AA$, $A^3 = A^2 A$, $\ldots$, $A^m = A^{m-1}A$, $A^0 = I$. Ясно, что

$$A^m = A^{m-1}A = A^{m-2}AA = A^{m-2}A^2 = A^{m-3}AA^2 = A^{m-3}A^3 = \cdots,$$

а потому $A^m = A^r A^s$, где $r+s = m$. Таким образом, $A^r A^s = A^s A^r$, иначе говоря, степени матрицы коммутируют между собой. Здесь m, r, s — неотрицательные целые числа.

Определение. Пусть дан полином $P_m(x) = p_0 + p_1 x + \cdots + p_m x^m$. Тогда значением полинома P_m от матричного аргумента A называется матрица

$$P_m(A) = p_0 I + p_1 A + \cdots + p_m A^m.$$

В силу перестановочности друг с другом степеней матрицы, справедливы соотношения

$$(P+Q)(A) = P(A) + Q(A)\ ,\ (PQ)(A) = P(A)Q(A).$$

Далее, если A — неособая матрица, то можно ввести в рассмотрение обратную матрицу A^{-1}. Под отрицательной степенью матрицы A будем понимать положительную степень обратной матрицы: $A^{-m} = (A^{-1})^m$. Это определение дает возможность ввести в рассмотрение матрицы $(A - \lambda I)^{-m}$, если λ не принадлежит спектру A.

Пусть теперь $P(x)$, $Q(x)$ — многочлены, причем корни полинома Q не принадлежат спектру A. Рассмотрим рациональную функцию $R(x) = P(x)/Q(x)$ и воспользуемся ее *разложением на простейшие дроби.* Это разложение (оно устанавливается в курсе анализа) дает для $R(x)$ представление в виде линейной комбинации многочлена и натуральных степеней функций $(x-x_k)^{-1}$, где x_k — корни многочлена Q. Основываясь на этом, можно определить матрицу $R(A)$. Легко видеть (проверьте!), что для двух рациональных функций R_1, R_2 выполнено

$$R_1(A)+R_2(A)=(R_1+R_2)(A), \quad R_1(A)R_2(A)=(R_1R_2)(A). \qquad (4)$$

В частности, $R_1(A)$ и $R_2(A)$ перестановочны.

Для любой матрицы $A \in M^n$ можно также определить $f(A)$, если функция $f(x)$ разлагается в сходящийся при любом x степенной ряд (ряд Тейлора)

$$f(x)=\sum_{k=0}^{\infty} f_k x^k \ , \quad f_k=\frac{f^{(k)}(0)}{k!}.$$

Тогда по определению

$$f(A)=\sum_{k=0}^{\infty} f_k A^k,$$

где сходимость ряда понимается поэлементно: $[f(A)]_{ij} = \sum_{k=0}^{\infty} f_k[A^k]_{ij}$, $i,j=1,\ldots,n$. Для функций $f_1(A)$, $f_2(A)$ справедлив аналог соотношений (4):

$$f_1(A)+f_2(A)=(f_1+f_2)(A), \quad f_1(A)f_2(A)=(f_1f_2)(A).$$

Примеры. 1) Пусть $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$. Тогда $A^m=\begin{pmatrix}1&m\\0&1\end{pmatrix}$. Действительно, это верно при $m=0$ и $m=1$, а для последующих значений m получаем по индукции:

$$A^m=A^{m-1}A=\begin{pmatrix}1&m-1\\0&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&m\\0&1\end{pmatrix}.$$

Тогда

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sum\limits_{k=0}^{\infty} \frac{1}{k!} & \sum\limits_{k=0}^{\infty} \frac{k}{k!} \\ 0 & \sum\limits_{k=0}^{\infty} \frac{1}{k!} \end{pmatrix} = \begin{pmatrix} e & e \\ 0 & e \end{pmatrix}.$$

2) Пусть теперь $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $f(x) = (x+1)/(x-1)$. Тогда

$$f(A) = (A+I)(A-I)^{-1} = \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 3 & 3 \end{pmatrix}^{-1} =$$

$$= -\frac{1}{6} \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ -3 & 0 \end{pmatrix} = -\frac{1}{6} \begin{pmatrix} 0 & -4 \\ -6 & -6 \end{pmatrix} = \begin{pmatrix} 0 & 2/3 \\ 1 & 1 \end{pmatrix}.$$

С другой стороны, так как $f(x) = 1 + 2(x-1)^{-1}$, то $f(A) = I + 2(A-I)^{-1}$:

$$f(A) = I + 2 \begin{pmatrix} 0 & 2 \\ 3 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 3 & -2 \\ -3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2/3 \\ 1 & 1 \end{pmatrix}.$$

Упражнения. 1) Пусть A, B — две неособые матрицы. Проверьте, что

$$B^{-1} - A^{-1} = B^{-1}(A-B)A^{-1} = A^{-1}(A-B)B^{-1}.$$

2) Если A, B перестановочны, то перестановочны любые их натуральные степени. Если при этом A, B — неособые, то перестановочны любые их целые степени. Проверьте.

4. Тождество Кэли. Так называется соотношение, составляющее содержание следующей теоремы.

Теорема 1. Пусть $A \in M^n$ и d_A — характеристический многочлен матрицы A. Тогда

$$d_A(A) = \mathbb{O}. \tag{5}$$

Доказательство. Пусть $(A-\lambda I)\check{}$ — матрица, присоединенная к $A-\lambda I$. В силу (7.8),

$$(A-\lambda I)(A-\lambda I)\check{} = \det(A-\lambda I)\, I = d_A(\lambda) I. \qquad (6)$$

Элементы матрицы $(A-\lambda I)\check{}$, с точностью до знака, совпадают с минорами матрицы $A-\lambda I$ порядка $n-1$, а следовательно, являются полиномами от λ степени не выше $n-1$. Введем матрицы $B_0, B_1, \ldots, B_{n-1}$, элементы которых совпадают с коэффициентами этих полиномов при соответствующих степенях λ:

$$[(A-\lambda I)\check{}]_{ij} = b_{0,ij} + b_{1,ij}\lambda + \cdots + b_{n-1,ij}\lambda^{n-1},$$

то есть

$$(A-\lambda I)\check{} = B_0 + B_1\lambda + \cdots + B_{n-1}\lambda^{n-1}.$$

В соответствии с (6),

$$(A-\lambda I)(B_0+B_1\lambda+\cdots+B_{n-1}\lambda^{n-1}) = d_A(\lambda)I = (\delta_0+\delta_1\lambda+\cdots+\delta_n\lambda^n)I. \qquad (7)$$

Приравнивая друг другу коэффициенты при одинаковых степенях λ в правой и левой частях (7), находим:

$$\begin{cases} AB_0 = \delta_0 I \\ AB_1 - B_0 = \delta_1 I \\ AB_2 - B_1 = \delta_2 I \\ \dots\dots\dots\dots\dots\dots \\ AB_{n-1} - B_{n-2} = \delta_{n-1} I \\ -B_{n-1} = \delta_n I. \end{cases} \qquad (8)$$

Умножим теперь слева k-ое равенство в (8) на A^{k-1}. Тогда

$$\begin{cases} AB_0 = \delta_0 I \\ A^2B_1 - AB_0 = \delta_1 A \\ A^3B_2 - A^2B_1 = \delta_2 A^2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ A^nB_{n-1} - A^{n-1}B_{n-2} = \delta_{n-1} A^{n-1} \\ -A^nB_{n-1} = \delta_n A^n. \end{cases} \qquad (9)$$

Сложим равенства (9). Все слагаемые в левой части сокращаются, а потому

$$\delta_0 I + \delta_1 A + \cdots + \delta_n A^n = d_A(A) = \mathbb{O},$$

что совпадает с (5). •

Одним из полезных следствий тождества Кэли является то, что матрица A^n (а следовательно, и более высокие степени матрицы A) суть линейные комбинации матриц $I, A, A^2, \cdots, A^{n-1}$. Поэтому вычисление любого полинома $p(A)$ можно свести к вычислению полинома от A степени не выше $n-1$. Действительно, $p(x) = s(x)d_A(x) + r_{n-1}(x)$, где $s(x)$ — частное от деления p на d_A, а r_{n-1} — остаток. Поэтому

$$p(A) = s(A)d_A(A) + r_{n-1}(A) = r_{n-1}(A).$$

Пример. $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. Найдем матрицу A^{10}. Поскольку

$$d_A(\lambda) = \begin{vmatrix} 1-\lambda & 1 \\ -2 & 4-\lambda \end{vmatrix} = \lambda^2 - 5\lambda + 6 = (\lambda-2)(\lambda-3);$$

$$\frac{x^{10}}{(x-2)(x-3)} = x^{10}\left(\frac{1}{x-3} - \frac{1}{x-2}\right) =$$
$$\frac{x^{10}-3^{10}}{x-3} - \frac{x^{10}-2^{10}}{x-2} + \frac{3^{10}}{x-3} - \frac{2^{10}}{x-2},$$

то $x^{10} = d_A(x)s(x) + (3^{10}(x-2) - 2^{10}(x-3))$, где $s(x) = (x^{10}-3^{10})(x-3)^{-1} - (x^{10}-2^{10})(x-2)^{-1}$ — полином степени 9. Следовательно, $A^{10} = 3^{10}(A-2I) - 2^{10}(A-3I)$.

5. Подобие матриц. Здесь мы обсудим два важных понятия — *подобие* и *диагонализуемость* матриц.

Определение. Матрицы $A, B \in M^n$ называются *подобными*, если существует неособая матрица $X \in M^n$ такая, что $B = X^{-1}AX$ (а тогда и $A = XBX^{-1}$). При этом X называется *матрицей, осуществляющей подобие*, или *аффинитетом.*

Отметим основные свойства подобия.

1. Характеристический многочлен (а, следовательно, и определитель, след, спектр) у подобных матриц один и тот же. Действительно, $d_B(\lambda) = \det(B - \lambda I) = \det(X^{-1}AX - \lambda X^{-1}IX) = \det(X^{-1}(A - \lambda I)X) = \det X^{-1} \det(A - \lambda I) \det X = d_A(\lambda)$.

Отметим, что подобные матрицы, в силу равенства их определителей, могут быть неособыми только одновременно.

2. Если матрицы A и B подобны, то подобны и их степени, причем подобие осуществляется той же матрицей. Действительно, если m — натуральное число, то

$$B^m = (X^{-1}AX)^m = X^{-1}AXX^{-1}AX \ldots X^{-1}AX = X^{-1}A^mX.$$

Далее, если матрицы A и B неособые, то $B^{-1} = (X^{-1}AX)^{-1} = X^{-1}A^{-1}(X^{-1})^{-1} = X^{-1}AX$, а следовательно, и $B^m = X^{-1}A^mX$ для всех $m \in \mathbb{Z}$. Как следствие, любые полиномы или рациональные функции от подобных матриц также подобны с тем же аффинитетом.

3. Если матрицы A и B подобны, то подобны также и матрицы A^t и B^t, A^* и B^*. Действительно, $B^t = (X^{-1}AX)^t = X^tA^t(X^t)^{-1}, B^* = X^*A^*(X^*)^{-1}$. При этом, однако, аффинитет меняется.

Определение. Матрица называется *диагонализуемой*, если она подобна некоторой диагональной матрице.

Поскольку спектр диагональной матрицы

$$A = \operatorname{diag}\{a_{11}, a_{22}, \ldots, a_{nn}\}$$

совпадает с числами $a_{11}, a_{22}, \ldots, a_{nn}$, то и для диагонализуемой матрицы $B = X^{-1}AX$ числа $a_{11}, \ldots, a_{nn}$ суть ее собственные значения.

Не все квадратные матрицы диагонализуемы. В качестве примера приведем матрицу $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Действительно, ее характеристический многочлен $d_A(\lambda)$ есть

$$d_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2.$$

Единственная диагональная матрица с таким же характеристическим многочленом — нулевая. Однако нулевая матрица не может быть подобна ненулевой: если $A = X^{-1}\mathbb{O}X$, то $A = \mathbb{O}$. Из рассмотренного примера следует также, что совпадение характеристических многочленов не является достаточным (а лишь необходимым) условием подобия матриц.

Вместе с тем, "большинство" матриц диагонализуемы. Приведем здесь *достаточные* условия диагонализуемости, которые будут доказаны в последующих главах курса. Диагонализуемы все матрицы *с простым спектром*, все *эрмитовы* и все *унитарные* (см. ниже п. 9.4) матрицы. Эти классы даже в совокупности не исчерпывают все множество диагонализумых матриц.

Легко понять, что для $A = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ будет $A^m = \mathrm{diag}\{\lambda_1^m, \ldots, \lambda_n^m\}$ при натуральном m и $m = 0$, а если все λ_j отличны от нуля, то это верно и для любых целых m. Поэтому, если $R(x)$ — рациональная функция, такая, что $R(A)$ имеет смысл, то $R(A) = \mathrm{diag}\{R(\lambda_1), \ldots, R(\lambda_n)\}$. Теперь естественно для *произвольной* функции $f(x)$ принять равенство

$$f(A) = \mathrm{diag}\{f(\lambda_1), \ldots, f(\lambda_n)\}$$

за **определение** функции $f(A)$ от диагональной матрицы A. Это определение согласовано с определением рациональной функции от матрицы и с определением через ряды. Заметим также, что для диагональной матрицы тождество Кэли очевидно, поскольку

$$d_A(A) = \mathrm{diag}\{d_A(\lambda_1), \ldots, d_A(\lambda_n)\} = \mathrm{diag}\{0, \ldots, 0\} = \mathbb{O}.$$

Это рассуждение переносится и на диагонализуемые матрицы. Действительно, если $B = X^{-1}AX$, где $A = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$, то для любого полинома $p(x)$ верно $p(B) = X^{-1}p(A)X$, откуда $d_B(B) = d_A(B) = X^{-1}d_A(A)X = \mathbb{O}$. Формулу

$$f(B) = X^{-1}\mathrm{diag}\{f(\lambda_1), \ldots, f(\lambda_n)\}X$$

естественно принять за **определение** произвольной функции $f(x)$ от диагонализуемой матрицы B.

§ 9. Специальные классы квадратных матриц

1. Симметричные и кососимметричные матрицы. Матрица $A \in M^n$ называется *симметричной*, если $A^t = A$, матрица B называется *кососимметричной* (или *антисимметричной*), если $B^t = -B$.

Всякую матрицу $C \in M^n$ можно единственным образом представить в виде $C = A + B$, где A — симметричная, B — кососимметричная матрицы. Действительно, $2C = (C + C^t) + (C - C^t)$, и можно положить $2A = (C + C^t)$, $2B = (C - C^t)$. Очевидно, $A = A^t$, $B = -B^t$. С другой стороны, если $C = A+B$, где $A = A^t$ и $B = -B^t$, то $C^t = A^t + B^t = A - B$. Тем самым $2A = C + C^t$, $2B = C - C^t$, что показывает единственность представления.

Отметим некоторые свойства симметричных и кососимметричных матриц.

1. Если матрица A симметрична (кососимметрична), то и матрица cA также симметрична (соответственно, кососимметрична).

2. Если две симметричные матрицы коммутируют, то их произведение симметрично: $(A_1A_2)^t = A_2^tA_1^t = A_2A_1 = A_1A_2$.

3. Собственные значения вещественной симметричной матрицы вещественны, а сами такие матрицы диагонализуемы. Это будет доказано в одной из последующих глав.

4. Диагональные элементы кососимметричных матриц равны нулю: $b_{kj} = -b_{jk}$, а потому $b_{kk} = 0$.

5. Для характеристического многочлена кососимметричной матрицы верно соотношение $d_B(\lambda) = (-1)^n d_B(-\lambda)$. Действительно,

$$d_B(\lambda) = \det\,(B - \lambda I) = \det\,(B - \lambda I)^t = \\ \det\,(-B - \lambda I) = (-1)^n \det\,(B + \lambda I).$$

Тем самым, спектр кососимметричной матрицы симметричен относительно нуля: если $d_B(\lambda) = 0$, то и $d_B(-\lambda) = 0$. Кратности ненулевых собственных значений λ и $-\lambda$ также совпадают. Если n нечетно, то $\det B = -\det B$, и нуль является собственным значением матрицы B.

Упражнение. Проверьте, что произведение коммутирующих кососимметричных матриц есть симметричная матрица.

2. Эрмитовы матрицы. Напомним, что *сопряженной*[1] к матрице F называется матрица F^*, для которой $[F^*]_{ij} = \overline{f_{ji}}$. По **определению**, матрица $F \in M^n$ называется *эрмитовой*, если $F = F^*$. Для вещественных матриц понятия симметричности и эрмитовости совпадают, для комплексных это уже не так. В числе основных свойств эрмитовых матриц, отметим (пока без доказательства), что их собственные значения вещественны, а сами они диагонализуемы. Заметим также, что всякая матрица $F \in M^n$ (единственным образом) представима в виде $F = G + iH$, где G и H — эрмитовы матрицы. Достаточно положить $2G = F + F^*$, $2iH = F - F^*$.

Следующие свойства предоставляется проверить читателю в качестве простых упражнений.

Упражнения. 1) Покажите, что диагональные элементы эрмитовых матриц вещественны.

2) Проверьте, что матрица, обратная к неособой эрмитовой матрице — эрмитова.

3) Покажите, что матрица iB, где B — вещественная кососимметричная матрица, эрмитова.

4) Докажите, что произведение коммутирующих эрмитовых матриц — эрмитова матрица.

3. Ортогональные матрицы. Вещественная матрица $V \in M^n$ называется *ортогональной*, если выполнено любое из трех (а тогда и два остальных) равносильных друг другу условий

$$V^tV = I, \quad VV^t = I, \quad V^{-1} = V^t.$$

Равносильность этих условий вытекает из теоремы 7.1.

Отметим, что равенство $V^tV = I$ означает *ортогональность столбцов* матрицы V:

$$\sum_{k=1}^{n} v_{ki}v_{kj} = \delta_{ij}, \quad i, j = 1, \ldots, n,$$

[1] Иногда используют равносильный термин "эрмитово-сопряженная" матрица.

а равенство $VV^t = I$ аналогичным образом означает *ортогональность строк* матрицы V. Таким образом, ортогональность матрицы по столбцам (по строкам) влечет ее ортогональность по строкам (соответственно по столбцам).

Примером ортогональной матрицы при $n = 2$ является матрица

$$V = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix},$$

описывающая поворот плоскости на угол α.

Отметим следующие свойства ортогональных матриц.

1. Единичная матрица ортогональна, так как $I^t = I = I^{-1}$.

2. Матрица, обратная к ортогональной, ортогональна. Действительно,

$$(V^{-1})^t V^{-1} = (V^t)^t V^t = VV^t = I.$$

3. Произведение ортогональных матриц ортогонально :

$$(V_1V_2)^t(V_1V_2) = (V_2^tV_1^t)(V_1V_2) = V_2^t(V_1^tV_1)V_2 = V_2^tIV_2 = I.$$

(Свойства 1–3 означают, что ортогональные матрицы образуют группу относительно матричного умножения; см. ниже п. 5).

4. Определитель ортогональной матрицы равен либо 1, либо -1. Действительно,

$$1 = \det(V^tV) = \det V \det V^t = (\det V)^2.$$

5. Все собственные значения ортогональных матриц по модулю равны 1 (они, однако, не обязательно вещественны). Это свойство пока оставляем без доказательства.

Ортогональные матрицы с определителем, равным 1, называют *собственно ортогональными*.

В пространстве $\mathbb{R}^3$ ортогональные матрицы описывают замены декартовых систем координат при сохранении начала координат. Действительно, пусть $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ — новая тройка ортов в $\mathbb{R}^3$. Рассмотрим матрицу, составленную из столбцов координат $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ относительно исходных ортов. Эта матрица, очевидно, ортогональна по столбцам. Условие, что матрица собственно ортогональна, означает, что тройка векторов $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ ориентирована так же, как исходная тройка ортов.

4. Унитарные матрицы. Комплексным аналогом ортогональных матриц являются матрицы *унитарные*. Матрица $U \in M^n$ называется *унитарной*, если выполнено любое из трех равносильных условий:

$$U^*U = I, \quad UU^* = I, \quad U^{-1} = U^*.$$

Всякая ортогональная матрица является унитарной, поскольку для вещественных матриц $V^t = V^*$.

Отметим основные свойства унитарных матриц.

1. Столбцы (строки) унитарной матрицы ортогональны как комплексные векторы:

$$\sum_{k=1}^{n} u_{ki}\overline{u_{kj}} = \sum_{k=1}^{n} u_{ik}\overline{u_{jk}} = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

2. Единичная матрица унитарна.

3. Матрица, обратная к унитарной, унитарна.

4. Произведение унитарных матриц — унитарная матрица.

Свойства 2–4 проверяются так же, как аналогичные свойства ортогональных матриц. Мы предоставляем сделать это читателю самостоятельно.

5. Определитель унитарной матрицы по модулю равен 1, поскольку

$$1 = \det(U^*U) = \det U \det U^* = \det U \,\overline{\det U}.$$

Унитарные матрицы, определитель которых равен 1, называются *собственно унитарными*.

6. Собственные значения унитарных матриц по модулю равны 1, а сами эти матрицы диагонализуемы. Это свойство будет доказано в одной из последующих глав.

5. Группы матриц. Приведем определение группы.

Определение. Множество G называется *группой*, если

1. На G задана бинарная операция $G \times G \to G$ (называемая обычно *умножением*), которая ассоциативна: для любых $a, b, c \in G$ выполнено $a(bc) = (ab)c$.

2. Существует нейтральный элемент (единица) $e \in G$, такой что $ae = ea = a$ для всех $a \in G$.

3. Для любого элемента $a \in G$ существует *обратный* элемент a^{-1}: $a^{-1}a = aa^{-1} = e$.

Умножение квадратных матриц ассоциативно и существует нейтральный элемент — единичная матрица. Какое-либо множество матриц будет образовывать группу в том и только в том случае, когда оно замкнуто относительно умножения и вместе с каждой матрицей A содержит обратную матрицу A^{-1}. Ясно, что все матрицы, входящие в группу, должны быть неособыми.

В качестве примеров укажем следующие группы матриц.

1. $GL(n)$ — группа всех неособых матриц порядка n.

2. $SL(n)$ — группа матриц порядка n с определителем, равным 1.

3. $O(n)$ — группа ортогональных матриц порядка n.

4. $SO(n)$ — группа собственно ортогональных матриц порядка n.

5. $U(n)$ — группа унитарных матриц порядка n.

6. $SU(n)$ — группа собственно унитарных матриц порядка n.

7. Группа неособых верхнетреугольных (нижнетреугольных) матриц.

Проверка того, что в каждом из перечисленных случаев указанное множество матриц действительно образует группу, не должно вызвать у читателя затруднений. Отметим, что использованные в примерах 1–6 обозначения являются стандартными.

Глава 2. Системы линейных алгебраических уравнений

§ 1. Ранг прямоугольной матрицы. Теорема о ранге

Материал этого параграфа относится к теории матриц. Он помещен, однако, в настоящую главу, поскольку понятие ранга матрицы имеет основное значение при исследовании линейных алгебраических систем. Мы дадим три различных определения ранга матрицы, а затем докажем, что все три определения эквивалентны. Это утверждение и составляет содержание теоремы о ранге. В качестве основного варианта мы будем рассматривать комплексный вариант теории. Соответствующие утверждения для вещественных матриц и векторов прямо следуют из "комплексных" рассмотрений.

1. Миноры. Первое определение ранга матрицы. Для $(m \times n)$-матриц введем понятие *минора порядка* $r \leq \min(m, n)$. Этим будет обобщено понятие минора (порядка $n-1$), введенное для квадратных матриц в § 5 главы 1. Пусть $A \in M^{m,n}$ и пусть у A вычеркнуто $(m-r)$ каких-либо строк и $(n-r)$ каких-либо столбцов. Останется квадратная матрица порядка r. Ее определитель называется *минором порядка* r матрицы A.

Определение 1. *Число $r = r_A$ называется рангом матрицы A, если хотя бы один ее минор порядка r отличен от нуля, а все миноры порядка $r+1$ равны нулю.*

Для неособой квадратной матрицы $A \in M^n$ ранг равен ее порядку $r_A = n$. Если $r = r_A = \min(m, n)$, то миноров порядка $r+1$ нет; в этом случае достаточно убедиться, что есть минор порядка r, отличный от нуля.

Если $r = r_A < \min(m, n)$, то не только миноры порядка $r+1$, но и все миноры более высоких порядков (если они есть) равны нулю. Например, для миноров порядка $r+2$ это видно, если каждый из них разложить по элементам какой-либо строки; соответствующие алгебраические дополнения выражаются через миноры порядка $r+1$. Затем так же убеждаемся, что равны нулю все миноры порядка $r+3$ и т. д. Таким образом, *ранг r_A есть наивысший порядок отличных от нуля миноров матрицы A.*

Учитывая свойства определителей и определение 1 ранга матрицы, отметим, что операция транспонирования, а также какая-либо перестановка местами строк (либо перестановка столбцов) не меняют величины r_A.

2. Линейная зависимость и независимость столбцов (строк). Второе и третье определения ранга матрицы. Дадим понятие линейной независимости системы векторов в $\mathbb{C}^m$. Пусть $\mathbf{x}_k \in \mathbb{C}^m$, $k = 1, \ldots, s$; число s назовем *длиной* этой системы (набора) векторов. Система векторов $\mathbf{x}_1, \ldots, \mathbf{x}_s$, называется *линейно зависимой*, если найдется набор чисел $c_1, \ldots, c_s$, не все из которых нули, такой что

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_s\mathbf{x}_s = 0, \quad \sum_{k=1}^{s} |c_k|^2 \neq 0.$$

Система векторов $\mathbf{x}_1, \ldots, \mathbf{x}_s$ называется *линейно независимой*, если она не является линейно зависимой. Иначе говоря, система $\mathbf{x}_1, \ldots, \mathbf{x}_s$ линейно независима, если из равенства

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_s\mathbf{x}_s = 0$$

следует, что $c_1 = c_2 = \ldots = c_s = 0$.

Будем рассматривать столбцы матрицы A как векторы из $\mathbb{C}^m$. Рангом матрицы A назовем наибольшую из длин линейно независимых наборов, составленных из столбцов матрицы A. Говоря более кратко, мы принимаем следующее

Определение 2. *Рангом матрицы A называется максимальное число линейно независимых столбцов матрицы A.*

Ранг по столбцам (в смысле второго определения) обозначим через $\widetilde{\rho} = \widetilde{\rho}_A$.

Подобным же образом будем рассматривать линейно независимые наборы строк матрицы A. Примем

Определение 3. *Рангом матрицы A называется максимальное число линейно независимых строк матрицы A.*

Ранг по строкам (в смысле третьего определения) обозначим через $\rho = \rho_A$. Сразу же ясно, что $\rho_A = \widetilde{\rho}_{A^t}$.

Очевидно, что какая-либо перестановка строк местами не меняет максимального числа линейно независимых строк ρ_A. Если же переставить местами, например, i-ый и j-ый столбцы матрицы A, то это приведет к одновременной перестановке i-го и j-го элементов в каждой строке матрицы A. Ясно, что при этом максимальное число ρ_A линейно независимых строк не изменится. Вообще, какая-либо перестановка местами строк (столбцов) не меняет ни ранга по строкам ρ_A, ни ранга по столбцам $\widetilde{\rho}_A$. Мы уже отмечали, что при этом не изменяется также r_A.

3. Теорема о ранге (формулировка). Следствия. Содержательность понятию ранга матрицы придает следующая

Теорема 1 (о ранге). *Все три определения ранга матрицы равносильны.*

Мы докажем эту теорему в п. 4, а сейчас обсудим ее следствия. Наиболее важным является

Следствие 2. *У всякой матрицы максимальное число ее линейно независимых строк равно максимальному числу ее линейно независимых столбцов.*

Иначе говоря, определения 2 и 3 равносильны. Это сильное утверждение, которое сравнительно трудно получить прямо. Оно проще всего получается за счет того, что оба эти определения эквивалентны определению 1.

Следствие 3. *Определитель матрицы $A \in M^n$ равен нулю тогда и только тогда, когда между ее строками, а также между ее столбцами есть линейная зависимость.*

Действительно, каждое из этих трех утверждений означает (в смысле одного из трех равносильных определений), что ранг матрицы A меньше n.

Следствие 4. *В координатном пространстве $\mathbb{C}^m$ всякий линейно независимый набор содержит не более m векторов.*

Доказательство. Пусть $\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{x}_{m+1} \in \mathbb{C}^m$. Составим из этих столбцов матрицу $X = (\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{x}_{m+1}) \in M^{m,m+1}$. Ее ранг не превосходит числа ее строк m. Поэтому и число линейно независимых столбцов не превосходит m, что и требовалось. •

Следствие 5. *В $\mathbb{C}^m$ заведомо существуют линейно незави-*

симые наборы из m векторов. Их образуют столбцы любой неособой матрицы $A \in M^m$.

Напротив, выписывая подряд m линейно независимых векторов (столбцов) из $\mathbb{C}^m$, мы получим *неособую* матрицу класса M^m. В частности, единичной матрице I соответствует линейно независимый набор "стандартных ортов"

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, \ \mathbf{e}_m = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Другой пример: если матрица A — треугольная и на ее главной диагонали нет нулей, то ее столбцы линейно независимы.

В "вещественном" варианте при $m = 3$ следствие 5 соответствует тому, что три вектора в $\mathbb{R}^3$ некомпланарны тогда и только тогда, когда их смешанное произведение отлично от нуля.

4. Доказательство теоремы о ранге потребует несколько шагов. Докажем эквивалентность определений 1 и 3, то есть равенство $r_A = \rho_A$.

Лемма 6. *Справедливо неравенство*

$$r_A \leq \rho_A. \tag{1}$$

Доказательство. Рассмотрим какой-либо минор порядка $\rho+1$ (здесь $\rho = \rho_A$) и обозначим строки матрицы A, входящие в этот минор, через $\omega_1, \ldots, \omega_\rho, \omega_{\rho+1}$. Эти строки обязательно линейно зависимы: найдется ненулевой набор чисел $c_1, \ldots, c_{\rho+1}$, такой что

$$c_1\omega_1 + c_2\omega_2 + \ldots + c_\rho\omega_\rho + c_{\rho+1}\omega_{\rho+1} = 0. \tag{2}$$

Пусть при этом $c_s \neq 0$ (такой номер s обязательно найдется). Прибавим к строке ω_s остальные строки, умноженные на числа $c_k(c_s)^{-1}$. Вследствие свойств определителей, это не изменит минор. При этом в силу (2) в s-ой строке окажутся одни нули, и минор равен нулю. Отсюда следует (1). •

Обратное неравенство доказывается сложнее. Назовем минор матрицы A порядка $r = r_A$ *базовым*, если этот минор не равен нулю.

Лемма 7 (о базовом миноре). *Строки матрицы A, не входящие в базовый минор, линейно зависят от строк, входящих в базовый минор.*

Доказательство. Выделим базовый минор, то есть какой-либо минор порядка $r = r_A$, отличный от нуля. Переставим сначала строки, а затем столбцы матрицы A так, чтобы этот минор располагался в левом верхнем углу. При этом оба числа r_A и ρ_A не изменятся. Выпишем получившуюся матрицу, сохраняя для нее прежнее обозначение A

$$A = \left(\begin{array}{cccc|ccc}
a_{11} & a_{12} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2r} & a_{2,r+1} & \cdots & a_{2n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{r1} & a_{r2} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\
\cline{1-4}
a_{r+1,1} & a_{r+1,2} & \cdots & a_{r+1,r} & a_{r+1,r+1} & \cdots & a_{r+1,n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{m1} & a_{m2} & \cdots & a_{mr} & a_{m,r+1} & \cdots & a_{mn}
\end{array}\right)$$

Матрицу, отчеркнутую в левом верхнем углу, обозначим через A_0. По построению $A_0 \in M^r$ и $\det A_0 \neq 0$. Далее, набор $\omega_1, \ldots, \omega_r$ первых r строк линейно независим, так как иначе было бы $\det A_0 = 0$. Покажем, что добавляя к этому набору еще одну строку матрицы A, получим линейно-зависимый набор. Пусть это строка ω_{r+1} (иначе опять переставили бы строки). Дополним матрицу A_0 элементами $(r+1)$-ой строки снизу и s-го столбца справа, причем $s = 1, \ldots, n$. Полученную матрицу обозначим через $A_s \in M^{r+1}$:

$$A_s = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1r} & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2r} & a_{2s} \\ \multicolumn{4}{c|}{\dotfill} & \cdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} & a_{rs} \\ \hline a_{r+1,1} & a_{r+1,2} & \cdots & a_{r+1,r} & a_{r+1,s} \end{array}\right)$$

Заметим, что $\det A_s = 0$. Действительно, если $s \le r$, то A_s содержит два одинаковых столбца. Если же $s > r$, то $\det A_s$ — один из миноров порядка $r+1$ матрицы A, а потому он равен нулю по определению числа $r = r_A$. Разложим $\det A_s$ по последнему столбцу, причем заметим, что алгебраические дополнения элементов этого столбца не зависят от номера s. Обозначим их через $c_1, \dots, c_{r+1}$. Заметим также, что

$$c_{r+1} = \det A_0 \ne 0. \tag{3}$$

Условие $\det A_s = 0$ теперь запишется в виде

$$\det A_s = \sum_{k=1}^{r+1} c_k a_{ks} = 0, \quad s = 1, \dots, n. \tag{4}$$

Совокупность равенств (4) есть развернутая форма соотношения

$$c_1\omega_1 + \ldots + c_r\omega_r + c_{r+1}\omega_{r+1} = 0, \tag{5}$$

причем, в силу (3), набор чисел $c_1, c_2, \dots, c_{r+1}$ — ненулевой. Итак, между строками $\omega_1, \dots, \omega_r, \omega_{r+1}$ есть линейная зависимость. Более того, строка ω_{r+1} линейно выражается через строки $\omega_1, \dots, \omega_r$. •

Лемма 8. *Справедливо неравенство*

$$\rho_A \le r_A. \tag{6}$$

Доказательство. Требуется доказать, что *любой* набор из $r+1$ строк матрицы A — линейно зависимый. Составим из выделенных $r+1$ строк матрицу $B \in M^{r+1,n}$. Ясно, что $r_B \leq r_A$. Выделяя здесь ненулевой минор порядка r_B, заключаем, что некоторые r_B+1 строк из B линейно зависимы (это следует из леммы 2 в применении к B). Тем более линейно зависим весь набор строк матрицы B, что и требовалось. •

Теперь легко завершить **доказательство теоремы о ранге.** Из (1) и (6) следует равенство $\rho_A = r_A$. Для доказательства эквивалентности определений 1 и 2 перейдем от A к транспонированной матрице A^t. При этом $r_A = r_{A^t}$ и $\widetilde{\rho}_A = \rho_{A^t}$. По уже доказанному свойству (в применении к A^t) $r_{A^t} = \rho_{A^t}$. Следовательно, $r_A = \widetilde{\rho}_A$. Теорема о ранге доказана. •

5. О вычислении ранга матрицы. Ранг матрицы не меняется при следующих *элементарных преобразованиях*: 1) транспонирование; 2) какая-либо перестановка строк (или столбцов) местами; 3) умножение строк (или столбцов) на ненулевые множители; 4) добавление к элементам какой-либо строки соответствующих элементов другой строки (то же для столбцов). Отметим, что в п. 1.7.4 вводилось понятие *допустимых* преобразований матрицы. Набор *элементарных* преобразований несколько шире.

В силу теоремы о ранге каждое из этих свойств достаточно проверить для какого-либо одного определения ранга. Мы уже отмечали выше, что r_A не меняется при транспонировании и что ранг (в смысле любого из определений) не меняется при преобразовании 2). Очевидно, что ρ_A не меняется при преобразованиях 3) и 4) относительно строк, а $\widetilde{\rho}_A$ не меняется при преобразованиях 3) и 4) относительно столбцов.

При вычислении ранга матрицы полезно применять к ней элементарные преобразования.

Пример. Вычислим ранг матрицы

$$A = \begin{pmatrix} 1 & 4 & 3 & 2 & 1 \\ 0 & 1 & -1 & -2 & 2 \\ 2 & 9 & 5 & 2 & 4 \\ 2 & 7 & 7 & 6 & 0 \end{pmatrix}.$$

Вычтем из третьей строки четвертую, а из четвертой удвоенную первую. Тогда матрица A перейдет в

$$\begin{pmatrix} 1 & 4 & 3 & 2 & 1 \\ 0 & 1 & -1 & -2 & 2 \\ 0 & 2 & -2 & -4 & 4 \\ 0 & -1 & 1 & 2 & -2 \end{pmatrix}.$$

Теперь прибавим к последней строке вторую, а к третьей строке прибавим удвоенную четвертую:

$$\begin{pmatrix} 1 & 4 & 3 & 2 & 1 \\ 0 & 1 & -1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Очевидно, последняя матрица имеет две линейно независимые строки. Поскольку ранг при сделанных преобразованиях не изменился, $r_A = 2$.

При вычислении ранга матрицы полезен также следующий прием. Пусть найден ненулевой минор порядка l. Рассмотрим только те миноры порядка $l+1$, которые содержат этот минор "внутри себя". Если все эти миноры равны нулю, то и вообще все миноры порядка $l+1$ равны нулю, и $r_A = l$. Почему это так? Вспомним доказательство леммы о базовом миноре. Фактически мы использовали там лишь то, что все миноры порядка $r+1$, "содержащие A_0 внутри себя", равны нулю. Из этого мы вывели, что любые $r+1$ строк линейно зависимы. В рассматриваемой сейчас ситуации это означает, что любые $l+1$ строк матрицы A линейно зависимы, а потому любой минор порядка $l+1$ равен нулю.

§ 2. Системы линейных алгебраических уравнений. Общие свойства

Рассмотрим систему m уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = f_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = f_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = f_m. \end{cases} \qquad (1)$$

Будем рассматривать $x_1, \ldots, x_n$ как координаты неизвестного вектора $\mathbf{x} \in \mathbb{C}^n$, $f_1, \ldots, f_m$ — как координаты известного вектора $\mathbf{f} \in \mathbb{C}^m$. Пусть $A = \{a_{ik}\}_{i=1,\ldots,m}^{k=1,\ldots,n}$ — матрица коэффициентов; ясно, что $A \in M^{m,n}$. Тогда систему (1) можно (и удобно) записать в виде одного уравнения

$$A\mathbf{x} = \mathbf{f} \tag{2}$$

относительно вектора $\mathbf{x}$. Матрица A задает линейное отображение $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^m$. Решить систему — значит для заданного вектора $\mathbf{f} \in \mathbb{C}^m$ найти вектор $\mathbf{x} \in \mathbb{C}^n$, удовлетворяющий (2). Иными словами, необходимо найти *прообраз* вектора $\mathbf{f}$ при отображении $\mathcal{A}$. Система (2) разрешима тогда и только тогда, когда такой прообраз (хотя бы один) существует.

Если $\mathbf{f} = 0$, то система (2) называется *однородной*; если же $\mathbf{f} \neq 0$, то система (2) называется *неоднородной*. Будем интересоваться общими вопросами разрешимости системы: 1) существует ли решение; 2) если решение существует, то единственно ли оно. Для однородной системы ответ на первый вопрос всегда положительный: однородная система

$$A\mathbf{z} = 0 \tag{3}$$

всегда имеет *тривиальное* решение $\mathbf{z} = 0$. Возникает вопрос 3) существует ли нетривиальное решение ($\mathbf{z} \neq 0$) однородной системы (3). Оказывается, вопрос 2) для неоднородной системы (2) равносилен вопросу 3) для однородной системы (3).

Предложение 1. *Пусть $\mathbf{x}_0 \in \mathbb{C}^n$ — некоторое решение системы (2), $\mathbf{z} \in \mathbb{C}^n$ — произвольное решение системы (3). Тогда $\mathbf{x}_0 + \mathbf{z}$ также является решением системы (2).*

Доказательство. Из $A\mathbf{x}_0 = \mathbf{f}$ и $A\mathbf{z} = 0$ следует $A(\mathbf{x}_0 + \mathbf{z}) = \mathbf{f}$. •

Предложение 2. *Пусть $\mathbf{x}_0$ и $\mathbf{x}_1$ — некоторые решения системы (2). Тогда $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{z}$, где $\mathbf{z}$ является решением системы (3).*

Доказательство. Из $A\mathbf{x}_1 = \mathbf{f}$ и $A\mathbf{x}_0 = \mathbf{f}$ следует $A(\mathbf{x}_1 - \mathbf{x}_0) = 0$ и (3) выполнено для $\mathbf{z} = \mathbf{x}_1 - \mathbf{x}_0$. •

Общим решением системы (2) (или (3)) называется совокупность всех ее решений, *частным* решением называется какое-либо фиксированное решение. Из предложений 1 и 2 вытекает

Теорема 3. *Общее решение* $\mathbf{x}$ *неоднородной системы* (2) *представимо в виде суммы частного решения* $\mathbf{x}_0$ *неоднородной системы* (2) *и общего решения* $\mathbf{z}$ *однородной системы* (3): $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$.

Из теоремы 3 вытекает следствие, которое и означает, что вопросы 2) и 3) равносильны.

Следствие 4. *Пусть неоднородная система* (2) *имеет решение* $\mathbf{x}_0$. *Это решение единственно тогда и только тогда, когда однородная система* (3) *имеет лишь тривиальное решение.*

Решения однородной системы обладают следующим свойством линейности.

Предложение 5. *Пусть* $\mathbf{z}_1, \ldots, \mathbf{z}_p$ — *некоторые решения системы* (3). *Тогда их линейная комбинация* $\mathbf{z} = c_1\mathbf{z}_1 + \ldots + c_p\mathbf{z}_p$ *(с произвольными постоянными* $c_1, \ldots, c_p$*) также является решением системы* (3).

Доказательство. Из равенств $A\mathbf{z}_j = 0$, $j = 1, \ldots, p$, прямо следует $A\mathbf{z} = 0$. •

Следующее предложение (принцип суперпозиции) отражает свойство линейности отображения $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^m$, заданного матрицей A. Оно позволяет находить решение системы с правой частью, равной линейной комбинации векторов $\mathbf{f}_j$, $j = 1, \ldots, p$, если известно решение системы для каждой правой части $\mathbf{f}_j$.

Предложение 6. *Пусть* $\mathbf{x}_j$ *является решением системы* $A\mathbf{x}_j = \mathbf{f}_j$, $j = 1, \ldots, p$. *Тогда линейная комбинация* $\mathbf{x} = c_1\mathbf{x}_1 + \ldots + c_p\mathbf{x}_p$ *является решением системы* $A\mathbf{x} = \mathbf{f}$ *с правой частью* $\mathbf{f} = c_1\mathbf{f}_1 + \ldots + c_p\mathbf{f}_p$.

Доказательство. Пользуясь свойством линейности отображения $\mathcal{A}$, имеем

$$A\mathbf{x} = A(c_1\mathbf{x}_1 + \ldots + c_p\mathbf{x}_p) = c_1A\mathbf{x}_1 + \ldots + c_pA\mathbf{x}_p = c_1\mathbf{f}_1 + \ldots + c_p\mathbf{f}_p = \mathbf{f}. \bullet$$

§ 3. Однородные системы линейных алгебраических уравнений

1. Критерий существования нетривиального решения. Рассмотрим однородную систему

$$A\mathbf{z} = 0. \tag{1}$$

Критерий существования нетривиального решения дает следующая теорема.

Теорема 1. *Однородная система* (1) *имеет нетривиальное решение тогда и только тогда, когда ранг матрицы A меньше числа неизвестных n:*

$$r_A < n. \tag{2}$$

Доказательство. Систему (1) запишем в виде

$$z_1\mathbf{a}_1 + z_2\mathbf{a}_2 + \ldots + z_n\mathbf{a}_n = 0, \tag{3}$$

где $\mathbf{a}_k$ — столбцы матрицы A,

$$\mathbf{a}_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}, \quad k = 1, \ldots, n. \tag{4}$$

Если $\mathbf{z}$ — нетривиальное решение, то хотя бы одно из чисел z_j ненулевое. Поэтому существование нетривиального решения системы (1) равносильно линейной зависимости столбцов матрицы A, что, в свою очередь, равносильно неравенству (2).
●

Следствие 2. *Однородная система n уравнений с n неизвестными имеет нетривиальное решение тогда и только тогда, когда определитель матрицы A равен нулю:* $\det A = 0$.

Следствие 3. *Однородная система m уравнений с n неизвестными при $m < n$ всегда имеет нетривиальное решение.*

Действительно, тогда $r_A \le m < n$.

2. Структура общего решения. Итак, пусть выполнено условие (2). Тогда существуют нетривиальные решения системы (1). Опишем общее решение. Пусть $r = r_A$. Переставим уравнения так, чтобы первые r строк матрицы A были линейно независимыми. Остальные $m - r$ строк линейно зависят от первых r строк. Поэтому оставшиеся $m - r$ уравнений можно отбросить: они являются следствием первых r уравнений. Затем согласованно переставим слагаемые в каждом уравнении так, чтобы в левом верхнем углу матрицы оказался базовый минор; такая перестановка отвечает перенумерации неизвестных. Матрицу получившейся системы снова обозначим через $A \in M^{r,n}$. Система теперь принимает вид

$$\begin{cases} a_{11}z_1 + \ldots + a_{1r}z_r + a_{1,r+1}z_{r+1} + \ldots + a_{1n}z_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}z_1 + \ldots + a_{rr}z_r + a_{r,r+1}z_{r+1} + \ldots + a_{rn}z_n = 0. \end{cases} \tag{5}$$

Используем обозначения

$$A_0 = \begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} \end{pmatrix}, \quad C = \begin{pmatrix} a_{1,r+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{r,r+1} & \ldots & a_{rn} \end{pmatrix}.$$

Матрица A_0 соответствует базовому минору матрицы A, поэтому $\det A_0 \neq 0$. Вектор неизвестных $\mathbf{z} \in \mathbb{C}^n$ разделим на "две части" $\widetilde{\mathbf{z}} \in \mathbb{C}^r$, $\widehat{\mathbf{z}} \in \mathbb{C}^{n-r}$:

$$\widetilde{\mathbf{z}} = \begin{pmatrix} z_1 \\ \vdots \\ z_r \end{pmatrix}, \quad \widehat{\mathbf{z}} = \begin{pmatrix} z_{r+1} \\ \vdots \\ z_n \end{pmatrix}.$$

Тогда систему (5) можно записать в виде

$$A_0\widetilde{\mathbf{z}} + C\widehat{\mathbf{z}} = 0. \tag{6}$$

Поскольку матрица A_0 обратима, то "зависимые" переменные $\widetilde{\mathbf{z}}$ выражаются через "свободные" переменные $\widehat{\mathbf{z}}$:

$$\widetilde{\mathbf{z}} = -A_0^{-1}C\widehat{\mathbf{z}}. \tag{7}$$

Вектор свободных переменных $\widehat{\mathbf{z}} \in \mathbb{C}^{n-r}$ может быть выбран произвольно. Таким образом, общее решение однородной системы имеет следующую структуру

$$\mathbf{z} = \begin{pmatrix} -A_0^{-1} C \widehat{\mathbf{z}} \\ \widehat{\mathbf{z}} \end{pmatrix}, \quad \widehat{\mathbf{z}} \in \mathbb{C}^s, \quad s = n - r. \tag{8}$$

Мы видим, что при условии (2) существует бесконечно много нетривиальных решений однородной системы. Их множество параметризуется s-мерным вектором свободных неизвестных.

3. Фундаментальная система решений. Максимальный набор линейно независимых решений однородной системы называется *фундаментальной системой решений*. Построим *стандартную* фундаментальную систему решений. Пусть

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_s = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{9}$$

— стандартный набор ортов в $\mathbb{C}^s$. Подставляя $\widehat{\mathbf{z}} = \mathbf{e}_k$ в (8), получим s решений однородной системы

$$\mathbf{z}_k = \begin{pmatrix} -A_0^{-1} C \mathbf{e}_k \\ \mathbf{e}_k \end{pmatrix}, \quad k = 1, \ldots, s. \tag{10}$$

Очевидно, векторы $\mathbf{z}_1, \ldots, \mathbf{z}_s$ линейно независимы. Столбец $\widehat{\mathbf{z}}$ можно записать в виде

$$\widehat{\mathbf{z}} = c_1 \mathbf{e}_1 + \ldots + c_s \mathbf{e}_s, \tag{11}$$

где $c_1, \ldots, c_s$ — координаты $\widehat{\mathbf{z}}$ в базисе (9). Из (8) и (10) тогда получаем, что общее решение можно представить в виде

$$\mathbf{z} = c_1 \mathbf{z}_1 + \ldots + c_s \mathbf{z}_s, \quad c_1, \ldots, c_s \in \mathbb{C}. \tag{12}$$

Это равенство и означает, что решения (10) образуют фундаментальную систему решений (называемую "стандартной") системы (1). Общее решение (12) есть произвольная линейная комбинация решений из фундаментальной системы.

Если вместо стандартного базиса (9) использовать какой-либо другой базис в $\mathbb{C}^s$ (т.е. s линейно независимых векторов из $\mathbb{C}^s$), то аналог формул (10) приведет нас к другой фундаментальной системе решений. Через эту систему также можно записать общее решение $\mathbf{z}$. Именно общее решение является инвариантом, не зависящим от конкретного выбора фундаментальной системы. Отметим, что даже *стандартная* фундаментальная система не инвариантна, поскольку в выборе свободных переменных есть произвол.

Пример. Рассмотрим однородную систему

$$\begin{cases} z_1 - z_2 + z_3 - z_4 = 0 \\ z_1 + z_2 + 2z_3 + 3z_4 = 0 \\ 2z_1 + 4z_2 + 5z_3 + 10z_4 = 0. \end{cases}$$

Эта система заведомо имеет нетривиальные решения, поскольку $4 = n > m = 3$. Найдем ранг матрицы

$$A = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 2 & 3 \\ 2 & 4 & 5 & 10 \end{pmatrix}.$$

Вычитая из второй строки первую, а из третьей удвоенную вторую, получаем матрицу

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 2 & 1 & 4 \\ 0 & 2 & 1 & 4 \end{pmatrix}$$

с двумя одинаковыми строками. Очевидно, $r_A = 2$. Исходная система равносильна системе из двух уравнений

$$\begin{cases} z_1 - z_2 + z_3 - z_4 = 0 \\ 2z_2 + z_3 + 4z_4 = 0. \end{cases}$$

Число свободных переменных $s = n - r = 2$. Переменные z_3 и z_4 можно принять за свободные; z_1 и z_2 через них выражаются:

$$z_1 = -\frac{3}{2}z_3 - z_4, \quad z_2 = -\frac{1}{2}z_3 - 2z_4.$$

Найдем стандартную фундаментальную систему решений. Полагая $z_3 = 1$, $z_4 = 0$, получим $z_1 = -3/2$, $z_2 = -1/2$. Полагая же $z_3 = 0$, $z_4 = 1$, получим $z_1 = -1$, $z_2 = -2$. Таким образом, стандартная фундаментальная система решений есть

$$\mathbf{z}_1 = \begin{pmatrix} -3/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{z}_2 = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

Общее решение системы имеет вид

$$\mathbf{z} = c_1\mathbf{z}_1 + c_2\mathbf{z}_2 = c_1 \begin{pmatrix} -3/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad c_1,\ c_2 \in \mathbb{C}.$$

§ 4. Неоднородные системы линейных алгебраических уравнений

1. Критерий существования решения неоднородной системы. Изучим теперь разрешимость неоднородной системы $A\mathrm{x} = \mathbf{f}$, которую запишем в развернутом виде:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = f_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = f_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = f_m. \end{cases} \tag{1}$$

С системой (1) связываем матрицу системы $A \in M^{m,n}$ и *расширенную* матрицу $(A|\mathbf{f}) \in M^{m,n+1}$ системы, т. е. матрицу

$$(A|\mathbf{f}) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & f_1 \\ a_{21} & \dots & a_{2n} & f_2 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & f_m \end{array}\right).$$

Следующая теорема дает критерий существования решения системы (1).

Теорема 1 (Кронекера — Капелли). *Для существования решения неоднородной системы* (1) *необходимо и достаточно, чтобы ранг расширенной матрицы* $(A|\mathbf{f})$ *был равен рангу матрицы* A*:*

$$r_{(A|\mathbf{f})} = r_A. \tag{2}$$

Доказательство. Для столбцов матрицы A используем обозначения (3.4). Тогда систему (1) можно записать в виде

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n = \mathbf{f}, \tag{3}$$

причем коэффициенты x_k, $k = 1, \ldots, n$, в (3) совпадают с координатами решения x. Таким образом, существование решения у системы $A\mathrm{x} = \mathbf{f}$ равносильно тому, что столбец $\mathbf{f}$ линейно выражается через столбцы (3.4) матрицы A. Последнее означает, что дописывая к матрице A столбец $\mathbf{f}$, мы не изменяем ранга (в смысле определения 1.2), т. е. выполнено (2). •

Теорема Кронекера — Капелли дает "индивидуальный" критерий существования решения для заданной правой части $\mathbf{f}$. Следующая теорема дает критерий разрешимости системы при любой правой части.

Теорема 2. *Для того, чтобы при любом* $\mathbf{f} \in \mathbb{C}^m$ *существовало решение неоднородной системы* (1) *необходимо и достаточно, чтобы ранг матрицы* A *был равен числу уравнений* m*:*

$$r_A = m. \tag{4}$$

Доказательство сводится к прямой ссылке на предложение 3, которое устанавливается ниже.

2. Условия разрешимости. Выясним, каким условиям надо подчинить правую часть $\mathbf{f}$, чтобы выполнялось условие (2) и система имела решение. Для строк матрицы A используем обозначения (ср. § 1)

$$\omega_j = \begin{pmatrix} a_{j1} & a_{j2} & \ldots & a_{jn} \end{pmatrix}, \quad j = 1, \ldots, m.$$

Пусть $r = r_A$. Переставляя уравнения в системе, можно считать, что первые r строк матрицы A линейно независимы, а остальные строки линейно зависят от них. Систему (1), очевидно, можно записать в виде совокупности равенств

$$\omega_1\mathbf{x} = f_1, \ldots, \omega_r\mathbf{x} = f_r, \omega_{r+1}\mathbf{x} = f_{r+1}, \ldots, \omega_m\mathbf{x} = f_m. \tag{5}$$

Строки ω_{r+i} при $i = 1, \ldots, m - r$ линейно зависят от первых r строк, то есть найдутся такие числа γ_{ij}, что

$$\omega_{r+i} = \sum_{j=1}^{r} \gamma_{ij}\omega_j, \quad i = 1, \ldots, m - r. \tag{6}$$

Предложение 3. *Пусть $r_A = r$, первые r строк матрицы A линейно независимы, и выполнены соотношения* (6). *Тогда для существования решения системы* (1) *необходимо и достаточно, чтобы правая часть* $\mathbf{f}$ *удовлетворяла $\sigma = m - r$ условиям разрешимости*

$$f_{r+i} = \sum_{j=1}^{r} \gamma_{ij} f_j, \quad i = 1, \ldots, \sigma. \tag{7}$$

Доказательство. *Необходимость.* Пусть существует решение $\mathbf{x}$. Тогда выполнено (5). С учетом (6) имеем

$$f_{r+i} = \omega_{r+i}\mathbf{x} = \left(\sum_{j=1}^{r} \gamma_{ij}\omega_j\right)\mathbf{x} = \sum_{j=1}^{r} \gamma_{ij} f_j, \quad i = 1, \ldots, \sigma.$$

Достаточность. Пусть выполнены соотношения (7). Вместе с (6) это означает, что строки расширенной матрицы $(A|\mathbf{f})$ с номерами $r+1, \ldots, m$ линейно зависят от первых r строк. Следовательно, $r_{(A|\mathbf{f})} = r_A$, и, в силу теоремы 1, существует решение системы. •

Ясно, что теорема 2 следует из предложения 3 при $\sigma = 0$. Заметим также, что если число уравнений больше числа неизвестных $m > n$, то заведомо $m > r$ и возникают условия разрешимости.

3. Структура общего решения. Итак, пусть выполнены соотношения (6) и условия разрешимости (7). Тогда исходная система (1) равносильна системе из первых r уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = f_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = f_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rn}x_n = f_r. \end{cases} \tag{8}$$

Теперь согласованно переставим слагаемые (за счет перенумерации неизвестных) в каждом уравнении так, чтобы минор, отвечающий матрице

$$A_0 = \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \dots & \dots & \dots \\ a_{r1} & \dots & a_{rr} \end{pmatrix},$$

был базовым, то есть $\det A_0 \neq 0$. Найдем *частное* решение $\mathbf{x}_0$ системы, полагая

$$x_{r+1} = \ldots = x_n = 0. \tag{9}$$

Тогда неизвестные $x_1, \ldots, x_r$ найдутся из системы r уравнений с неособой матрицей A_0:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix} = A_0^{-1} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_r \end{pmatrix}.$$

В силу теоремы 2.3, общее решение $\mathbf{x}$ неоднородной системы есть сумма частного решения $\mathbf{x}_0$ и общего решения $\mathbf{z}$ однородной системы (см. (3.12))

$$\mathbf{x} = \mathbf{x}_0 + \sum_{k=1}^{s} c_k \mathbf{z}_k, \quad s = n - r, \quad c_1, \ldots, c_s \in \mathbb{C}. \tag{10}$$

Итак, *для разрешимости системы* $A\mathbf{x} = \mathbf{f}$ *из* m *уравнений с* n *неизвестными правая часть* $\mathbf{f}$ *должна быть подчинена*

$\sigma = m - r$ условиям разрешимости; при этом общее решение зависит от $s = n - r$ произвольных постоянных.

Мы не касаемся здесь вычислительных схем для решения систем общего вида. Отметим лишь, что метод Гаусса, описанный в § 1.7 для систем с квадратной неособой матрицей, может быть перенесен и на общий случай.

Пример. Рассмотрим неоднородную систему

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 4 \\ x_1 + x_2 + 2x_3 + 3x_4 = 8 \\ 2x_1 + 4x_2 + 5x_3 + 10x_4 = 20. \end{cases}$$

Соответствующая однородная система была рассмотрена в конце § 3. Здесь $r = 2$, $s = 2$, $\sigma = 1$. Третья строка матрицы системы есть линейная комбинация первых двух: $\omega_3 = -\omega_1 + 3\,\omega_2$. Условие разрешимости $f_3 = -f_1 + 3f_2$, очевидно, выполнено. Система равносильна системе двух уравнений

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 4 \\ 2x_2 + x_3 + 4x_4 = 4. \end{cases}$$

Полагая здесь $x_3 = x_4 = 0$, находим $x_1 = 6$, $x_2 = 2$. Тем самым найдено частное решение

$$\mathbf{x}_0 = \begin{pmatrix} 6 \\ 2 \\ 0 \\ 0 \end{pmatrix}.$$

Общее решение есть

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{z}_1 + c_2\mathbf{z}_2 = \begin{pmatrix} 6 \\ 2 \\ 0 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} -3/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad c_1, c_2 \in \mathbb{C}.$$

§ 5. Системы линейных алгебраических уравнений с квадратной матрицей

Пусть $m = n$, то есть число уравнений равно числу неизвестных. Тогда числа $\sigma = m - r$ и $s = n - r$ совпадают: $s = \sigma$. Свойства разрешимости системы $A\mathbf{x} = \mathbf{f}$ зависят от того, является ли матрица A особой или нет. Именно, есть следующие две взаимоисключающие возможности (альтернатива).

1. Пусть $\det A \neq 0$. Тогда $r = m = n$, $s = \sigma = 0$. Существует единственное решение $\mathbf{x} = A^{-1}\mathbf{f}$ системы $A\mathbf{x} = \mathbf{f}$ при любой правой части $\mathbf{f}$. Однородная система $A\mathbf{z} = 0$ имеет только тривиальное решение $\mathbf{z} = 0$.

2. Пусть $\det A = 0$. Тогда $r < n = m$, $s = \sigma > 0$. Однородная система имеет нетривиальные решения. Число линейно независимых решений равно s. Неоднородная система имеет решения, если выполнено s условий разрешимости, накладываемых на $\mathbf{f}$. При этом ее решение не единственно: оно зависит от s произвольных постоянных.

Обычно противопоставление этих двух возможностей формулируют в следующей "острой" форме.

Альтернатива Фредгольма[2] . *Для системы с квадратной матрицей либо решение неоднородной системы существует при любой правой части, либо существует нетривиальное решение однородной системы.*

§ 6*. Односторонняя обратимость матриц

1. Односторонняя обратимость и ее связь с разрешимостью линейных систем. Напомним (см. п. 1.7.1), что для $A \in M^{m,n}$ матрица $B \in M^{n,m}$ называется левой обратной, если $BA = I_n$. Матрица $C \in M^{n,m}$ называется правой обратной, если $AC = I_m$. Обсудим условия существования односторонних обратных матриц и свяжем эти условия со свойствами уравнений (2.2) и (2.3).

---

[2] И.Фредгольм, шведский математик, сформулировал подобную альтернативу на существенно более сложном материале интегральных уравнений.

Предложение 1. *Если матрица A имеет левую обратную, то однородная система (2.3) имеет только тривиальное решение.*

Доказательство. Если $BA = I$, то из (2.3) следует $(BA)\mathbf{z} = 0$, а потому $\mathbf{z} = 0$. •

Предложение 2. *Если матрица A имеет правую обратную, то неоднородная система (2.2) разрешима при любой правой части.*

Доказательство. Пусть $AC = I$. Тогда вектор $\mathbf{x} = C\mathbf{f}$ дает решение системы (2.2): $A\mathbf{x} = (AC)\mathbf{f} = I\mathbf{f} = \mathbf{f}$. •

Доказанные предложения показывают, что правая обратная матрица "отвечает" за существование (при любом $\mathbf{f}$) решения системы $A\mathbf{x} = \mathbf{f}$, а левая обратная — за единственность решения, коль скоро оно есть. Сопоставляя предложение 1 с теоремой 3.1, получаем

Предложение 3. *Если матрица $A \in M^{m,n}$ имеет левую обратную, то*

$$r_A = n. \tag{1}$$

Точно так же, предложение 2 вместе с теоремой 4.2 влечет

Предложение 4. *Если матрица $A \in M^{m,n}$ имеет правую обратную, то*

$$r_A = m. \tag{2}$$

Из (1) и (2) вытекает

Следствие 5. *Двусторонняя обратимость возможна только для квадратных матриц.*

2. Критерий односторонней обратимости. Для квадратных матриц вопросы обратимости полностью решаются теоремой 1.7.1. Поэтому достаточно рассмотреть случай $A \in M^{m,n}$ при $m \neq n$ (хотя формальной необходимости в последнем условии нет). Наша цель состоит в обращении предложений 3 и 4.

Предложение 6. *При условии (1) матрица $A \in M^{m,n}$ обратима слева, а при условии (2) обратима справа.*

Доказательство. Обсудим условие (1), считая $n < m$. Надо построить матрицу $B \in M^{n,m}$ так, чтобы было

$$\sum_{j=1}^{m} b_{ij} a_{jk} = \delta_{ik}, \quad i,k = 1,\ldots,n. \tag{3}$$

Выберем какие-либо n линейно независимых строк матрицы A. Пусть их номера (в порядке возрастания) $j_1,\ldots,j_n$. Составленная из этих строк матрица класса M^n, очевидно, обратима. Построим обратную к ней матрицу. Последовательные столбцы обратной матрицы примем за столбцы матрицы B с номерами $j_1,\ldots,j_n$. Остальные $m-n$ столбцов матрицы B будем считать нулевыми. Легко понять, что тогда все равенства (3) выполнены.

Рассмотрим теперь случай условия (2). Тогда для матрицы A^t равенство $r_{A^t} = m$ играет роль условия (1). Следовательно, найдется матрица $\widetilde{B} \in M^{m,n}$, такая что $\widetilde{B}A^t = I_m$. Переходя к транспонированным матрицам, получим равенство $AC = I_m$ при $C = \widetilde{B}^t$. •

3. Описание множества односторонних обратных. Если $r_A = m < n$, то правая обратная матрица заведомо не единственна. То же верно для левой обратной при условии $r_A = n < m$. Здесь мы приведем полное описание множеств односторонних обратных матриц. При выводе соответствующих формул мы укажем лишь схему доказательства. Читателю рекомендуется восполнить подробности в качестве упражнения.

Начнем со случая $r_A = m < n$. Если C_0 — какая-либо правая обратная матрица для $A \in M^{m,n}$, то $C = C_0 + G$, где $G \in M^{n,m}$ — произвольное решение матричного уравнения

$$AG = \mathbb{O}. \tag{4}$$

Опишем множество этих решений. Сейчас $s = n - r_A = n - m > 0$, а потому однородное уравнение (2.3) имеет нетривиальные решения. Пусть столбцы $\mathbf{z}_1,\ldots,\mathbf{z}_s$ составляют какую-либо фундаментальную систему решений уравнения (2.3) и пусть

$$Z := (\mathbf{z}_1,\ldots,\mathbf{z}_s) \in M^{n,s}.$$

Тогда общее решение G уравнения (4) имеет представление

$$G = Z\Gamma, \quad \Gamma \in M^{s,m}, \tag{5}$$

где матрица Γ произвольна.

Пусть теперь $r_A = n < m$ и $\sigma = m - n$. Если $B_0 A = I$, то множество всех левых обратных матриц B имеет вид $B = B_0 + D$, где $D \in M^{n,m}$ — решение матричного уравнения $DA = \mathbb{O}$. Последнее равносильно уравнению $A^t D^t = \mathbb{O}$, которое имеет вид (4). Поэтому для D^t справедливо представление, аналогичное (5). В результате приходим к равенству

$$D = \widetilde{\Gamma}\widetilde{Z}^t,$$

где матрица $\widetilde{Z} \in M^{m,\sigma}$ составлена из столбцов фундаментальной системы решений для "транспонированного" однородного уравнения $A^t\widetilde{\mathbf{z}} = 0$, а $\widetilde{\Gamma} \in M^{n,\sigma}$ — произвольная матрица.

Упражнение. Проверьте, что всякая левая обратная матрица к матрице

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

имеет вид

$$B = \begin{pmatrix} -2 & 1 & 0 \\ 3/2 & -1/2 & 0 \end{pmatrix} + \begin{pmatrix} c_1 & -2c_1 & c_1 \\ c_2 & -2c_2 & c_2 \end{pmatrix}, \quad c_1, c_2 \in \mathbb{C}.$$

Содержание