Матрицы и системы линейных уравнений

П.А. Кожевников

27 мая 2016 г.

Оглавление

Предисловие

Это учебное пособие пособие написано на основе лекций по курсу аналитической геометрии и линейной алгебры, который читается автором на факультете общей и прикладной физики МФТИ с 2007 года.

По традиции, сложившейся на Физтехе, матрицы и системы линейных уравнений систематически изучаются до введения абстрактных понятий векторного пространства и линейного отображения. При дальнейшем освоении абстрактных понятий матрицы, наряду с наглядной геометрией, становятся источником мотивировок и примеров. Такая методика преподавания отражена в этом издании.

Изложение в основном тексте ведется последовательно. Предполагается, что читатель знаком с элементарной математикой и основами теории множеств (некоторые начальные сведения собраны в приложении). Основной текст снабжен сносками-комментариями, многие из которых даны с целью связать излагаемый фрагмент теории с другими сюжетами из алгебры и геометрии.

В тексте символами $\triangleright$ и $\square$ отмечаются соответственно начало и конец доказательства утверждений (теорем, предложений и лемм). В случае доказательства эквивалентности $A \Leftrightarrow B$ условий A и B символы $\boxed{\Rightarrow}$ и $\boxed{\Leftarrow}$ обозначают начало доказательства соответственно следствий $A \Rightarrow B$ и $B \Rightarrow A$. Курсивом выделяются формулировки утверждений, а также определяемые понятия.

Автор благодарен своим учителям и коллегам из МГУ и МФТИ, в процессе работы с которыми складывались понимание математики и стиль ее преподавания. Также автор выражает благодарность своим слушателям — интерес к предмету и увлеченность многих студентов МФТИ способствовали работе над этим изданием.

Доцент МФТИ
П. Кожевников

Введение

Пусть даны натуральные числа m и n. *Матрицей* размера $m \times n$ будем называть упорядоченный набор из mn действительных[1] чисел — *элементов* матрицы, записанный в виде таблицы с m *строками* и n *столбцами*.

Множество матриц (с действительными числами) размера $m \times n$ будем обозначать[2] $\mathbf{M}_{m\times n}$. Матрицы размера 1×1 иногда позволим себе отождествлять с числами.[3]

Матрица $A \in \mathbf{M}_{m\times n}$ с элементами a_{ij}, где $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, выглядит так:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}. \tag{1}$$

Коротко пишем $A = (a_{ij})$. Строку с номером i будем обозначать $a_{i\,\bullet}$: $a_{i\bullet} = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix}$. Столбец с номером j будем обозначать $a_{\bullet j}$: $a_{\bullet j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$. Строки и столбцы представляют собой матрицы размера $1\times n$ и $m\times 1$: $a_{i\bullet} \in \mathbf{M}_{1\times n}$, $a_{\bullet j} \in \mathbf{M}_{m\times 1}$. Строки и столбцы иногда

[1] Излагаемая здесь теория остается без изменений для матриц с рациональными или комплексными числами, или матриц с элементами из произвольного поля K. Иногда рассматриваются матрицы, элементами которых являются не числа, а другие объекты — например векторы, или даже некоторые отображения. Встречаются и так называемые *блочные матрицы* — матрицы, элементы которых сами являются матрицами.

[2] Во многих книгах для множества матриц размера $m \times n$ с элементами из поля (или кольца) K принято обозначение $\mathbf{Mat}_{m\times n}(K)$.

[3] Как линейное пространство множество $\mathbf{M}_{m\times n}$ естественно отождествляется с $\mathbb{R}^{mn}$ (*изоморфно* $\mathbb{R}^{mn}$).

называют *векторами-строками* и *векторами-столбцами*. Для матрицы (1) можно использовать *строчную запись* $A = \begin{pmatrix} a_{1\,\bullet} \\ a_{2\,\bullet} \\ \vdots \\ a_{m\,\bullet} \end{pmatrix}$, или *столбцовую запись* $A = \begin{pmatrix} a_{\bullet 1} & a_{\bullet 2} & \cdots & a_{\bullet n} \end{pmatrix}$.

Элементы матрицы A, находящиеся на пересечении некоторых строк $a_{i_1\,\bullet}, a_{i_2\,\bullet}, \ldots, a_{i_k\,\bullet}$ $(1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant m)$ и некоторых столбцов $a_{\bullet j_1}, a_{\bullet j_2}, \ldots, a_{\bullet j_l}$ $(1 \leqslant j_1 < j_2 < \ldots < j_l \leqslant n)$, образуют *подматрицу* (размера $k \times l$):

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \ldots & a_{i_1 j_l} \\ a_{i_2 j_1} & a_{i_2 j_2} & \ldots & a_{i_2 j_l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k j_1} & a_{i_k j_2} & \ldots & a_{i_k j_l} \end{pmatrix}.$$

В частности, строки и столбцы матрицы являются ее подматрицами.[4] Фиксированный элемент матрицы также можно считать подматрицей размера 1×1.

Если $m = n$, то матрица $A \in \mathbf{M}_{m\times n}$ называется *квадратной*, число n называется *порядком* квадратной матрицы.

В матрице $A = (a_{ij}) \in \mathbf{M}_{m\times n}$ элементы a_{11}, a_{22}, …, a_{kk}, где $k = \min\{m, n\}$, называются *диагональными*, они образуют *главную диагональ*.

Квадратная матрица $A = (a_{ij})$ порядка n называется *верхнетреугольной*, если для любых $1 \leqslant j < i \leqslant n$ выполнено $a_{ij} = 0$. Иначе говоря, матрица верхнетреугольная, если все ее элементы под главной диагональю равны нулю. Аналогично определяются *нижнетреугольные* матрицы как квадратные матрицы, в которых все элементы над главной диагональю равны нулю. Квадратная матрица $A = (a_{ij}) \in \mathbf{M}_{n\times n}$ называется *диагональной*, если при $i \neq j$ выполнено $a_{ij} = 0$. Таким образом, в диагональной матрице все элементы вне главной диагонали нулевые. Диагональную матрицу

[4] Иногда подматрицу $k \times l$ называют *минором* $k \times l$, хотя чаще минором называют определитель квадратной подматрицы.

$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$, у которой $a_{ii} = \lambda_i$ для $i = 1, 2, \dots, n$, обозначаем $\operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Матрица $\operatorname{diag}(\lambda, \lambda, \dots, \lambda)$ называется *скалярной*, а матрица $\operatorname{diag}(1, 1, \dots, 1)$ — *единичной*. Единичную матрицу порядка n будем обозначать E_n или E, если из контекста ясен ее порядок.

Матрицы $A = (a_{ij}) \in \mathbf{M}_{m\times n}$ и $B = (b_{ij}) \in \mathbf{M}_{k\times l}$ считаются *равными* (пишем $A = B$), если они имеют одинаковые размеры, и на соответствующих местах у них равные элементы (то есть $k = m$, $l = n$ и $a_{ij} = b_{ij}$ для всех $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$).

В дальнейших рассмотрениях нам встретятся *наборы*, или *системы матриц* (в частности, системы векторов-строк и векторов-столбцов).[5] Если в конечной системе матриц зафиксирован порядок, в котором они перечисляются, то говорят об *упорядоченной системе матриц*.

§ 1. Линейные операции над матрицами, транспонирование матриц

Сложение и умножение на число

Матрицы одинакового размера можно складывать, матрицу можно умножать на число. Эти операции определены самым естественным образом — "покомпонентно".

Определение. Пусть даны матрицы $A, B \in \mathbf{M}_{m\times n}$, $A = (a_{ij})$, $B = (b_{ij})$. Матрица $(c_{ij}) \in \mathbf{M}_{m\times n}$ называется *суммой* матриц A и B, если $c_{ij} = a_{ij} + b_{ij}$ для всех $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

Определение. Пусть дана матрица $A \in \mathbf{M}_{m\times n}$, $A = (a_{ij})$, и число $\lambda \in \mathbb{R}$. Матрица $(d_{ij}) \in \mathbf{M}_{m\times n}$ называется *произведением* матрицы A на число λ, если $d_{ij} = \lambda a_{ij}$ для всех $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

---

[5] Отличие системы (набора) от множества в том, что в ней один элемент может содержаться в нескольких экземплярах. Аналогично понятию подмножества вводится понятие *подсистемы* (поднабора). Можно говорить об операциях объединения и пересечения систем.

Сумму матриц A и B обозначим $A+B$, произведение матрицы A на число λ обозначим λA.

Операцию, сопоставляющую паре матриц A и B матрицу $A+B$, будем называть *сложением*, а операцию, сопоставляющую матрице A произведение A на фиксированное число λ, будем называть *умножением* [6] на число λ.

Определение. Матрица из $\mathbf{M}_{m\times n}$, все элементы которой равны 0, называется *нулевой матрицей.*

Нулевые матрицы обычно обозначают O.

Определение. Матрица $(-1)A$ называется *противоположной* для матрицы A.

Матрицу, противоположную для матрицы A, обозначаем $-A$.

Можно ввести операцию *вычитания*, то есть говорить о *разности* матриц, полагая $A-B=A+(-B)$.

Предложение 1.1. $\forall\ A,B,C\in\mathbf{M}_{m\times n},\ \forall\ \lambda,\mu\in\mathbb{R}$ *выполнены следующие свойства:*[7]

1) $(A+B)+C=A+(B+C)$;
2) $A+B=B+A$;
3) $A+O=A$ *(здесь* $O\in\mathbf{M}_{m\times n}$ *— нулевая матрица);*
4) $A+(-A)=O$;
5) $(\lambda+\mu)A=\lambda A+\mu A$;
6) $\lambda(A+B)=\lambda A+\lambda B$;
7) $1\cdot A=A$;
8) $(\lambda\mu)A=\lambda(\mu A)$.

$\triangleright$ Все перечисленные свойства легко вытекают из соответствующих свойств для чисел — так как операции сложения и умножения на число выполняются "покомпонентно". $\square$

[6] Более точно, сложение — это отображение $\psi:\mathbf{M}_{m\times n}\times\mathbf{M}_{m\times n}\to\mathbf{M}_{m\times n}$, при котором $\psi(A,B)=A+B$, а умножение на λ — отображение $\delta_\lambda:\mathbf{M}_{m\times n}\to\to\mathbf{M}_{m\times n}$, при котором $\delta_\lambda(A)=\lambda A$.

[7] Знакомые с началами линейной алгебры легко переформулируют это предложение следующим образом: множество $\mathbf{M}_{m\times n}$ является векторным пространством относительно операций сложения и умножения на число.
Свойства 1) и 2) называются *ассоциативностью* и *коммутативностью* сложения, 5) и 6) — свойствами *линейности*, или *дистрибутивности*.

Следствие. $\forall\ A, B \in \mathbf{M}_{m\times n}$, $\forall\ \lambda, \mu \in \mathbb{R}$ *выполнено:*

1) $0 \cdot A = \lambda \cdot O = O$;

2) $-(\lambda A) = (-\lambda)A = \lambda(-A)$;

3) $(\lambda - \mu)A = \lambda A - \mu A$;

4) $\lambda(A - B) = \lambda A - \lambda B$.

Линейная комбинация и линейная оболочка

Определение. Пусть A_1, A_2, $\dots$, $A_k \in \mathbf{M}_{m\times n}$, λ_1, λ_2, $\dots$, $\lambda_k \in \mathbb{R}$. Сумма $\sum\limits_{i=1}^{k} \lambda_i A_i$ называется *линейной комбинацией* матриц $A_1, A_2, \dots$ $\dots, A_k$ с *коэффициентами* $\lambda_1, \lambda_2, \dots, \lambda_k$.

Если хотя бы один из коэффициентов в сумме $\sum\limits_{i=1}^{k} \lambda_i A_i$ не равен 0, то говорят, что линейная комбинация *нетривиальная*. Ясно, что *тривиальная* линейная комбинация равна нулевой матрице. Если матрица $B \in \mathbf{M}_{m\times n}$ равна линейной комбинации матриц[8] A_1, A_2, $\dots$, A_k, говорят, что матрица B *раскладывается* по A_1, A_2, $\dots$, A_k, или *линейно выражается* через $A_1, A_2, \dots, A_k$.

Определение. Пусть $\mathcal{A}$ — система матриц из $\mathbf{M}_{m\times n}$. Множество матриц из $\mathbf{M}_{m\times n}$, которые линейно выражаются через несколько (конечное число) матриц из $\mathcal{A}$, называется *линейной оболочкой* системы $\mathcal{A}$.

Линейная оболочка[9] системы $\mathcal{A}$ обозначается $\langle\mathcal{A}\rangle$. В частности, $\langle A_1, A_2, \dots, A_k\rangle$ — линейная оболочка конечной системы матриц A_1, A_2, $\dots$, A_k.

---

[8] Свойства из предложения 1.1 позволяют утверждать, что в сумме $\sum\limits_{i=1}^{k} \lambda_i A_i$ получится матрица, не зависящая от порядка выполнения операций сложения и умножения на число.

[9] В терминах линейной алгебры $\langle\mathcal{A}\rangle$ — это наименьшее подпространство векторного пространства, содержащее все векторы из $\mathcal{A}$. Формально можно записать $\langle\mathcal{A}\rangle = \left\{ \sum\limits_{i=1}^{k} \lambda_i A_i \,\middle|\, k \in \mathbb{N},\ A_1, A_2, \dots, A_k \in \mathcal{A},\ \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R} \right\}$.

Итак, по определению $B \in \langle \mathcal{A} \rangle \Leftrightarrow B$ линейно выражается через несколько матриц из $\mathcal{A}$. Условимся считать, что $\langle \varnothing \rangle = O$. Ясно, что для любого множества $\mathcal{A} \subset \mathbf{M}_{m\times n}$ выполнено $\mathcal{A} \subset \langle \mathcal{A} \rangle$. Также понятно, что для любой системы матриц $\mathcal{A}$ выполнено $O \in \langle \mathcal{A} \rangle$.

Транспонирование

Определение. Пусть дана матрица $A \in \mathbf{M}_{m\times n}$, $A = (a_{ij})$. Матрица $(b_{ij}) \in \mathbf{M}_{n\times m}$ называется *транспонированной* к матрице A, если $b_{ij} = a_{ji}$ для всех $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

Матрицу, транспонированную к A, будем обозначать A^T.

Операцию, сопоставляющую каждой матрице A матрицу A^T, будем называть *транспонированием*.[10]

Нетрудно представить транспонирование наглядно: это симметрия относительно главной диагонали. Столбцы матрицы A^T — это транспонированные строки матрицы A, то есть если $A = \begin{pmatrix} a_{1\bullet} \\ a_{2\bullet} \\ \vdots \\ a_{m\bullet} \end{pmatrix}$, то $A^T = \begin{pmatrix} a_{1\bullet}^T & a_{2\bullet}^T & \ldots & a_{m\bullet}^T \end{pmatrix}$.[11]

Предложение 1.2. $\forall\, A, B \in \mathbf{M}_{m\times n}$, $\forall\, \lambda \in \mathbb{R}$ *выполнено:* [12]

1) $(A^T)^T = A$;

2) $(A + B)^T = A^T + B^T$

3) $(\lambda A)^T = \lambda A^T$.

$\triangleright$ Непосредственная проверка.[13] $\square$

[10]Говоря более точно, транспонирование — это отображение (очевидно, биективное) $\varphi : \mathbf{M}_{m\times n} \to \mathbf{M}_{n\times m}$, при котором $\varphi(A) = A^T$.

[11]По этому правилу удобно выписывать матрицу, транспонированную к данной.

[12]Свойства 2) и 3) означают, что операция транспонирования линейна.

[13]Свойство 1) очевидно. Для 2) и 3) приведем соответствующую проверку:
2) Пусть $A = (a_{ij})$, $B = (b_{ij})$, $A^T = (c_{ij})$, $B^T = (d_{ij})$, $A + B = S = (s_{ij})$, $S^T = (r_{ij})$. Тогда $r_{ij} = s_{ji} = a_{ji} + b_{ji} = c_{ij} + d_{ij}$, что и требуется.
3) Пусть $A = (a_{ij})$, $\lambda A = B = (b_{ij})$, $B^T = (c_{ij})$, $A^T = (d_{ij})$. Тогда $c_{ij} = b_{ji} = \lambda a_{ji} = \lambda d_{ij}$.

Задачи и упражнения

1. Вычислите $A + 3B^T$, где $A = \begin{pmatrix} 4 & -2 & 0 \\ 2 & -5 & -1 \end{pmatrix}$, $B = \begin{pmatrix} -3 & 3 \\ 2 & -6 \\ 1 & 7 \end{pmatrix}$.

2. Укажите некоторую систему из четырех матриц A_1, A_2, A_3, $A_4 \in \mathbf{M}_{2\times 2}$ такую, что $\langle A_1, A_2, A_3, A_4\rangle = \mathbf{M}_{2\times 2}$.

3. Докажите, что множество строк $\{(x_1, x_2, x_3, x_4) \,|\, x_1 + x_2 + x_3 + x_4 = 0\}$ совпадает с линейной оболочкой строк $(-1, 1, 0, 0)$, $(-1, 0, 1, 0)$, $(-1, 0, 0, 1)$.

4. Пусть $\mathcal{A}$ — произвольная система столбцов из $\mathbf{M}_{n\times 1}$. Докажите, что $\langle\langle \mathcal{A}\rangle\rangle = \langle \mathcal{A}\rangle$.

5. Квадратная матрица A называется *симметрической*, если $A^T = A$, и *кососимметрической*, если $A^T = -A$. Докажите, что любая квадратная матрица представляется единственным образом в виде суммы $B + C$, где B — симметрическая, а C — кососимметрическая матрицы.

§ 2. Линейная зависимость и ранг

Линейная зависимость

В этом и следующем пунктах речь пойдет о линейной зависимости и ранге для систем векторов-столбцов. Все определения и утверждения без изменений переносятся на системы векторов-строк.[14]

Определение. Система столбцов $A_1, A_2, \ldots, A_k$ из $\mathbf{M}_{m\times 1}$ называется *линейно зависимой*, если некоторая их нетривиальная линейная комбинация равна O, и *линейно независимой* в противном случае.

[14] Все остается в силе и для матриц фиксированного размера, и вообще для систем элементов произвольного векторного пространства. Для векторов геометрического векторного пространства понятие линейной зависимости имеет следующий смысл: система из двух векторов $\mathbf{a}_1, \mathbf{a}_2$ линейно зависима $\Leftrightarrow$ $\mathbf{a}_1$ и $\mathbf{a}_2$ коллинеарны; система из трех векторов $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ линейно зависима $\Leftrightarrow$ $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ компланарны.

Полагают, что пустая система векторов-столбцов линейно независима — формально это согласуется с определением.

Предложение 2.1. *Система столбцов* A_1, A_2, $\ldots, A_k \in \mathbf{M}_{m\times 1}$ $(k \geqslant 2)$ *линейно зависима* $\Leftrightarrow$ *среди столбцов* $A_1, A_2, \ldots, A_k$ *найдется столбец, который линейно выражается через остальные.*

$\triangleright$ $\boxed{\Rightarrow}$ Пусть $\sum\limits_{i=1}^{k} \lambda_i A_i = O$, и не все коэффициенты равны 0, скажем, $\lambda_k \neq 0$. Тогда поделим равенство на $-\lambda_k$ и перенесем A_k в другую часть; получим $A_k = \sum\limits_{i=1}^{k-1} \mu_i A_i$, где $\mu_i = -\dfrac{\lambda_i}{\lambda_k}$, $i = 1, 2, \ldots$ $\ldots, k-1$.

$\boxed{\Leftarrow}$ Пусть, скажем, A_k раскладывается по столбцам A_1, A_2, $\ldots$, A_{k-1}: $A_k = \sum\limits_{i=1}^{k-1} \mu_i A_i$. Тогда $\sum\limits_{i=1}^{k-1} \mu_i A_i - A_k$ — нетривиальная линейная комбинация, равная O. $\square$

Предложение 2.2. *1) Если в системе* A_1, A_2, $\ldots$, $A_k \in \mathbf{M}_{m\times 1}$ *некоторая подсистема линейно зависима, то и вся система линейно зависима.*

2) Подсистема линейно независимой системы линейно независима.

$\triangleright$ 1) Пусть, скажем, для системы столбцов A_1, A_2, $\ldots$, A_k ее подсистема A_1, A_2, $\ldots$, A_s $(s \leqslant k)$ линейно зависима, и некоторая нетривиальная линейная комбинация $\sum\limits_{i=1}^{s} \mu_i A_i$ равна O. Тогда $\sum\limits_{i=1}^{s} \mu_i A_i + 0 \cdot A_{s+1} + \ldots + 0 \cdot A_k$ — нетривиальная линейная комбинация, равная O.

2) Это переформулировка утверждения 1). $\square$

Следствие. *1) Система столбцов из* $\mathbf{M}_{m\times 1}$, *содержащая* O, *является линейно зависимой.*

2) Система столбцов из $\mathbf{M}_{m\times 1}$, *содержащая два одинаковых столбца, является линейно зависимой.*

Предложение 2.3. *Пусть даны столбец $B \in \mathbf{M}_{m\times 1}$ и линейно независимая система A_1, A_2, $\ldots$, $A_k \in \mathbf{M}_{m\times 1}$. Если $B = \sum\limits_{i=1}^{k} \lambda_i A_i$, то коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_k$ определяются однозначно.*

$\triangleright$ Предположим, что B разложен по столбцам $A_1, A_2, \ldots, A_k$ еще каким-то способом: $B = \sum\limits_{i=1}^{k} \mu_i A_i$. Вычитая из одного разложения другое, получаем $\sum\limits_{i=1}^{k} (\lambda_i - \mu_i) A_i = O$. Так как $A_1, A_2, \ldots, A_k$ — линейно независимая система, то левая часть последнего равенства — тривиальная линейная комбинация, откуда $\lambda_i = \mu_i$, $i = 1, 2, \ldots, k$. $\square$

Определение. Пусть $\mathcal{A}$ — система векторов-столбцов из $\mathbf{M}_{m\times 1}$. Конечную подсистему A_1, A_2, $\ldots$, A_r системы $\mathcal{A}$ будем называть *базисной* для $\mathcal{A}$, если она линейно независима, и любой столбец из $\mathcal{A}$ принадлежит $\langle A_1, A_2, \ldots, A_r \rangle$.

Ниже дадим описание базисных[15] подсистем для произвольных систем векторов-столбцов, а пока укажем пример базисной подсистемы в $\mathbf{M}_{m\times 1}$. В множестве $\mathbf{M}_{m\times 1}$ рассмотрим систему столбцов $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, $\ldots$, $e_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ (у столбца e_i на i-м месте единица, а остальные элементы равны нулю).

Предложение 2.4. *Система столбцов $e_1, \ldots, e_m$ является базисной подсистемой*[16] *в $\mathbf{M}_{m\times 1}$.*

---

[15]Предлагаемый термин "базисная подсистема"находится в согласии с общепринятым определением базиса в линейной алгебре: действительно, упорядоченная базисная подсистема в системе $\mathcal{A}$ — это базис подпространства $\langle \mathcal{A} \rangle$. Для линейного отображения $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, заданном матрицей A, базисная система столбцов матрицы A — это базис в $\operatorname{Im} \varphi$.

[16]Упорядоченную систему столбцов $e_1, \ldots, e_m$ называют *стандартным базисом* в $\mathbb{R}^m = \mathbf{M}_{m\times 1}$.

$\triangleright$ Из равенства $\sum\limits_{i=1}^{m} \lambda_i e_i = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$ видно, что если $\sum\limits_{i=1}^{m} \lambda_i e_i = O$, то $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$, поэтому $e_1, \ldots, e_m$ — линейно независимая система. Из того же равенства видно, что любой столбец из $\mathbf{M}_{m\times 1}$ можно разложить по столбцам $e_1, \ldots, e_m$. $\square$

Ранг

Определение. Целое неотрицательное число r называется *рангом* непустой (возможно, бесконечной) системы $\mathcal{A}$ векторов-столбцов из $\mathbf{M}_{m\times 1}$, если в $\mathcal{A}$ можно выбрать r столбцов, являющихся линейно независимой системой, но нельзя выбрать $r+1$ столбцов, являющихся линейно независимой системой.

Обозначение для ранга:[17] $\operatorname{rg}\mathcal{A}$. В частности, $\operatorname{rg}(A_1, A_2, \ldots, A_k)$ — ранг конечной системы столбцов $A_1, A_2, \ldots, A_k$.

Итак, ранг системы векторов-столбцов $\mathcal{A}$ — это максимальное количество векторов-столбцов, которое может быть в линейно независимой подсистеме $\mathcal{A}$. В частности, система из одного или нескольких нулевых столбцов имеет ранг 0. Ясно, что ранг определен для любой конечной системы векторов-столбцов и не превосходит количества столбцов в системе.[18] В случае, если в системе имеются несколько равных столбцов, можно оставить один из них, удалив "копии— ранг при этом не изменится.

Предложение 2.5. *Конечная система $\mathcal{A}$, состоящая из k столбцов, линейно независима $\Leftrightarrow \operatorname{rg}\mathcal{A} = k$.*

$\triangleright$ Сразу следует из определения. $\square$

[17]Если $\mathcal{A}$ является подпространством векторного пространства, то ранг принято называть *размерностью* подпространства $\mathcal{A}$ и обозначать $\dim\mathcal{A}$.

[18]Ниже мы увидим, что и для любой бесконечной системы векторов-столбцов из $\mathbf{M}_{m\times 1}$ определен ранг, причем этот ранг не превосходит m. В отличие от пространства столбцов высоты m, существуют векторные пространства, содержащие системы векторов бесконечного ранга (это означает, что для любого r можно выбрать линейно независимую подсистему из r векторов).

Предложение 2.6. *Пусть* $\mathcal{A}_1$, $\mathcal{A}_2$ — *две системы столбцов из* $\mathbf{M}_{m\times1}$, *причем* $\operatorname{rg}\mathcal{A}_1 = r_1$, $\operatorname{rg}\mathcal{A}_2 = r_2$. *Тогда* $\operatorname{rg}(\mathcal{A}_1 \bigcup \mathcal{A}_2) \leqslant r_1 + r_2$.

$\triangleright$ Пусть это не так, и в объединении наборов $\mathcal{A}_1$ и $\mathcal{A}_2$ нашлась линейно независимая подсистема из $r_1 + r_2 + 1$ столбцов. Но (по определению ранга и предложению 2.2) среди этих столбцов не более r_1 столбцов из $\mathcal{A}_1$ и не более r_2 столбцов из $\mathcal{A}_2$. Противоречие. $\square$

Предложение 2.7. *Пусть* $\mathcal{A}, \mathcal{B}$ — *две системы столбцов из* $\mathbf{M}_{m\times1}$, *причем* $\mathcal{A} \subset \mathcal{B}$ *и* $\operatorname{rg}\mathcal{B} = r$. *Тогда* $\operatorname{rg}\mathcal{A} \leqslant r$.

$\triangleright$ Сразу следует из определения. $\square$

Предыдущее предложение почти очевидно: если к системе столбцов $\mathcal{A}$ добавить некоторые столбцы (расширить $\mathcal{A}$ до системы $\mathcal{B}$), то ранг не уменьшится. Оказывается, ранг не изменится, если к $\mathcal{A}$ добавлять линейные комбинации столбцов из $\mathcal{A}$ (и наоборот, ранг не изменится, если из системы столбцов удалить столбец, который раскладывается по оставшимся столбцам). Обобщение этого факта составляет содержание следующей основной теоремы о рангах. Доказательству теоремы предпошлем две леммы (которые являются частными случаями теоремы).

Лемма 1. *Пусть* $\mathcal{A}$ — *система векторов-столбцов из* $\mathbf{M}_{m\times1}$, *причем* $\operatorname{rg}\mathcal{A} = r$. *Тогда любая линейно независимая подсистема из* r *столбцов является базисной для* $\mathcal{A}$.

$\triangleright$ Пусть $A_1, A_2, \ldots, A_r$ — некоторая линейно независимая подсистема в $\mathcal{A}$, A — произвольный столбец из $\mathcal{A}$. Достаточно доказать, что A линейно выражается через столбцы A_1, $A_2, \ldots$, A_r.

Из определения ранга следует, что система из $r+1$ столбцов A, A_1, A_2, $\ldots$, A_r линейно зависима. Тогда найдутся числа μ, $\lambda_1, \ldots$, $\lambda_r \in \mathbb{R}$, не все равные нулю и такие, что $\mu A + \sum\limits_{i=1}^{r} \lambda_i A_i = O$. При этом $\mu \neq 0$, иначе система A_1, A_2, $\ldots$, A_r была бы линейно зависимой. Отсюда $A = -\sum\limits_{i=1}^{r} \dfrac{\lambda_i}{\mu} A_i$. $\square$

Лемма 2. *Пусть* $A_1, A_2, \ldots, A_k$ — *конечная система векторов-столбцов из* $\mathbf{M}_{m\times1}$, $B \in \langle A_1, A_2, \ldots, A_k\rangle$. *Тогда* $\operatorname{rg}(A_1, A_2, \ldots, A_k) =$ $= \operatorname{rg}(A_1, A_2, \ldots, A_k, B)$.

$\triangleright$ Положим $r = \mathrm{rg}(A_1, A_2, \ldots, A_k)$ $(r \leqslant k)$. Предположим, что утверждение неверно, и в системе A_1, A_2, $\ldots$, A_k, B нашлась линейно независимая подсистема из $r+1$ векторов-столбцов. Тогда один из этих $r+1$ столбцов — это B (так как в системе A_1, A_2,$\ldots$ $\ldots$, A_k нет линейно независимой подсистемы из $r+1$ векторов-столбцов). Итак, пусть для определенности $B, A_1, A_2, \ldots, A_r$ — линейно независимая система. Из предложения 2.2 следует, что система $A_1, A_2, \ldots, A_r$ линейно независима, значит, по лемме 1 каждый из векторов-столбцов $A_1, A_2, \ldots, A_k$ лежит в $\langle A_1, A_2, \ldots, A_r\rangle$. По условию B равен линейной комбинации векторов-столбцов $A_1, A_2, \ldots, A_k$: $B = \sum\limits_{i=1}^{k} \lambda_i A_i$. Подставив в это выражение разложения $A_1, A_2, \ldots$ $\ldots, A_k$ по столбцам $A_1, A_2, \ldots, A_r$, получим, что B раскладывается по векторам-столбцам $A_1, A_2, \ldots, A_r$. Но это противоречит линейной независимости системы $B, A_1, A_2, \ldots, A_r$ (см. предложение 2.1). $\square$

Теорема 2.1 (основная теорема о рангах)**.** *Пусть $\mathcal{A}$ и $\mathcal{B}$ — две такие системы векторов-столбцов из $\mathbf{M}_{m\times 1}$, что $\mathcal{A} \subset \mathcal{B}$. Пусть* $\mathrm{rg}\,\mathcal{A} = r$. *Тогда* $\mathrm{rg}\,\mathcal{B} = r \Leftrightarrow$ *любой столбец из $\mathcal{B}$ принадлежит $\langle\mathcal{A}\rangle$.*

$\triangleright$ $\boxed{\Rightarrow}$ В системе $\mathcal{A}$ зафиксируем некоторую линейно независимую подсистему $A_1, A_2, \ldots, A_r$ из r столбцов. Так как $A_1, A_2, \ldots$ $\ldots, A_r \in \mathcal{B}$ и $\mathrm{rg}\,\mathcal{B} = r$, то по лемме 1 (примененной к системе $\mathcal{B}$) любой столбец из $\mathcal{B}$ лежит в $\langle A_1, A_2, \ldots, A_r\rangle$ и, следовательно, в $\langle\mathcal{A}\rangle$.

$\boxed{\Leftarrow}$ Предположим, что утверждение неверно, и в системе $\mathcal{B}$ нашлась линейно независимая подсистема из $r+1$ векторов-столбцов $B_1, B_2, \ldots, B_{r+1}$. Каждый из них линейно выражается через несколько (конечное число) векторов-столбцов из $\mathcal{A}$, поэтому можно выбрать конечную систему столбцов $A_1, A_2, \ldots, A_l$ из $\mathcal{A}$, через которые линейно выражается каждый из векторов-столбцов B_1, B_2, $\ldots$, B_{r+1}. Тогда $\mathrm{rg}(A_1, A_2, \ldots, A_l, B_1, B_2, \ldots, B_{r+1}) \geqslant \mathrm{rg}(B_1, B_2, \ldots, B_{r+1}) =$ $= r+1$. С другой стороны, применяя многократно лемму 2, имеем: $\mathrm{rg}(A_1, A_2, \ldots, A_l, B_1, B_2, \ldots, B_{r+1}) = \mathrm{rg}(A_1, A_2, \ldots, A_l, B_1, B_2, \ldots$ $\ldots, B_r) = \mathrm{rg}(A_1, A_2, \ldots, A_l, B_1, B_2, \ldots, B_{r-1}) = \ldots = \mathrm{rg}(A_1, A_2, \ldots$ $\ldots, A_l) \leqslant \mathrm{rg}\,\mathcal{A} = r$. Противоречие. $\square$

Следствие 1 (описание базисных подсистем)**.** *Пусть $\mathcal{A}$ — система векторов-столбцов из $\mathbf{M}_{m\times 1}$, и* $\mathrm{rg}\,\mathcal{A} = r$. *Тогда базисными для*

$\mathcal{A}$ являются в точности линейно независимые подсистемы из r векторов-столбцов.

Следствие 2. *Для любой системы $\mathcal{A}$ векторов-столбцов из $\mathbf{M}_{m\times 1}$ ее ранг определен, причем $\operatorname{rg}\mathcal{A} \leqslant m$.*

$\triangleright$ Очевидно, $\operatorname{rg}\mathcal{A} \leqslant \operatorname{rg}\mathbf{M}_{m\times 1}$. Согласно предложению 2.4, в множестве $\mathbf{M}_{m\times 1}$ всех столбцов высоты m имеется базисная подсистема из m столбцов. Отсюда $\operatorname{rg}\mathbf{M}_{m\times 1} = m$. $\square$

Следствие 3. *Для любой системы векторов-столбцов $\mathcal{A}$ из $\mathbf{M}_{m\times 1}$ выполнено $\operatorname{rg}\mathcal{A} = \operatorname{rg}\langle\mathcal{A}\rangle$.*

Ранг матрицы

Определение. *Столбцовым (строчным) рангом* матрицы называется ранг системы ее столбцов (строк).

Столбцовый и строчный и ранг матрицы A обозначим соответственно $\operatorname{rg}_v A$ и $\operatorname{rg}_h A$. Итак, для матрицы (1) по определению $\operatorname{rg}_v A = \operatorname{rg}\{a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}\}$, $\operatorname{rg}_h A = \operatorname{rg}\{a_{1\,\bullet}, a_{2\,\bullet}, \ldots, a_{m\,\bullet}\}$. В §4 мы докажем, что строчный и столбцовый ранги любой матрицы равны, после чего сможем говорить о ранге матрицы $\operatorname{rg} A$ (не уточняя, какой из рангов $\operatorname{rg}_v A$ и $\operatorname{rg}_h A$ имеется в виду).

Предложение 2.8. *Если A' — некоторая подматрица матрицы A, то $\operatorname{rg}_v A' \leqslant \operatorname{rg}_v A$ и $\operatorname{rg}_h A' \leqslant \operatorname{rg}_h A$.*

$\triangleright$ Достаточно показать, что после вычеркивания строки (или столбца) ранги rg_v и rg_h не увеличатся.

Очевидно, что после вычеркивания строки rg_h не увеличится.

После вычеркивания строки линейно зависимая подсистема столбцов остается линейно зависимой (если некоторая нетривиальная линейная комбинация нескольких столбцов была равна O, то это свойство сохранится и после вычеркивания строки), следовательно, rg_v не увеличивается. $\square$

Предложение 2.9. *Пусть $A, B \in \mathbf{M}_{m\times n}$. Тогда $\operatorname{rg}_v(A+B) \leqslant \operatorname{rg}_v A + \operatorname{rg}_v B$ (и $\operatorname{rg}_h(A+B) \leqslant \operatorname{rg}_h A + \operatorname{rg}_h B$).*

$\triangleright$ Каждый столбец матрицы $A + B$ линейно выражается через столбцы $a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}, b_{\bullet 1}, b_{\bullet 2}, \ldots, b_{\bullet n}$. Поэтому, пользуясь теоремой 2.1, $\mathrm{rg}_v(A + B) \leqslant \mathrm{rg}\langle a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}, b_{\bullet 1}, b_{\bullet 2}, \ldots, b_{\bullet n}\rangle = \mathrm{rg}(a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}, b_{\bullet 1}, b_{\bullet 2}, \ldots, b_{\bullet n})$. Но из предложения 2.6 следует, что последний ранг не превосходит $\mathrm{rg}(a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}) + \mathrm{rg}(b_{\bullet 1}, b_{\bullet 2}, \ldots, b_{\bullet n}) = \mathrm{rg}_v A + \mathrm{rg}_v B$.

Рассуждения для rg_h аналогичны. $\square$

Задачи и упражнения

1. Пусть известно, что некоторый столбец B раскладывается по столбцам A_1, A_2, …, A_k единственным образом. Докажите, что система A_1, A_2, …, A_k линейно независима.

2. Пусть $A, B \in \mathbf{M}_{m\times n}$. Может ли оказаться, что $\mathrm{rg}\, A = 10$, $\mathrm{rg}\, B = 5$, $\mathrm{rg}(A + B) = 3$?

3. Докажите, что матрицу ранга r можно представить в виде суммы r матриц ранга 1.

4. Найдите наибольшее количество матриц, которое может быть в линейно независимой подсистеме а) симметрических; б) кососимметрических матриц порядка n.

5. Докажите, что любую линейно независимую подсистему в системе $\mathcal{A}$ столбцов высоты n можно дополнить до базисной подсистемы.

§ 3. Умножение матриц

Определение, основные свойства и примеры

Определим умножение строки длины n на столбец высоты n.

Определение. Произведением $\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ называется число (то есть матрица размера 1×1) $c = \sum\limits_{i=1}^{n} a_i b_i$.

Определение. Произведением матрицы $A = (a_{ij}) \in \mathbf{M}_{m\times n}$ на матрицу $B = (b_{jk}) \in \mathbf{M}_{n\times p}$ называется матрица $(c_{ik}) \in \mathbf{M}_{m\times p}$, в которой $c_{ik} = a_{i\bullet} b_{\bullet k}$ для всех $i \in \{1, 2, \dots, m\}$, $k \in \{1, 2, \dots, p\}$.

Таким образом, произведение AB определено, если число столбцов матрицы A равно числу строк матрицы B. Операцию, сопоставляющую упорядоченной паре матриц A и B их произведение, будем называть операцией *умножения.*[19]

Теорема 3.1. $\forall\ A, A' \in \mathbf{M}_{m\times n}$, $\forall\ B, B' \in \mathbf{M}_{n\times p}$, $\forall\ C \in \mathbf{M}_{p\times q}$, $\forall\ \lambda \in \mathbb{R}$ *выполнены равенства:*[20]

1) $(AB)C = A(BC)$*;*
2) $A(B + B') = AB + AB'$*;*
3) $(A + A')B = AB + A'B$*;*
4) $\lambda(AB) = (\lambda A)B = A(\lambda B)$*;*
5) $(AB)^T = B^T A^T$*.*

[19] Более точно, умножение — это отображение $\varphi : \mathbf{M}_{m\times n} \times \mathbf{M}_{n\times p} \to \mathbf{M}_{m\times p}$, при котором $\varphi(A, B) = AB$.
Приведем несколько примеров, мотивирующих такое определение умножения матриц.
1. Если X и Y — координатные столбцы векторов $\mathbf{a}$ и $\mathbf{b}$ в ортонормированном базисе, то скалярное произведение $(\mathbf{a}, \mathbf{b})$ равно $X^T Y$.
2. Произведение матриц перехода от базиса e к e' и от e' к e'' — матрица перехода от e к e''.
3. Матрица композиции линейных отображений векторных пространств равна произведению матриц этих линейных отображений.
4. Матричная запись билинейной функции $X^T B Y$, где X и Y — координатные столбцы векторов, а B — матрица билинейной функции. Также в матричном виде иногда удобно записываются квадратичные функции, уравнения алгебраических кривых и поверхностей второго порядка.
5. Матричная запись $AX = b$ системы линейных уравнений — см. §5.

[20] Свойство 1) — ассоциативность умножения, свойства 2) и 3) — дистрибутивности умножения по сложению.

$\triangleright$ Непосредственная проверка.[21] $\square$

Рассмотрим некоторые матрицы, умножение на которые выглядит особенно просто.

Умножение любой матрицы справа или слева на нулевую матрицу (подходящего размера) очевидно приводит к нулевой матрице соответствующего размера.

При умножении матрицы $A = (a_{ij}) \in \mathbf{M}_{m\times n}$ слева на $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$ i-я строчка умножается на λ_i, а при умножении на $\mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_n)$ справа j-й столбец умножается на μ_j.[22] Так, умножение матрицы слева или справа на скалярную матрицу $\mathrm{diag}(\lambda, \lambda, \ldots, \lambda)$ равносильно умножению на λ, в частности, $E_m A = AE_n = A$.[23]

Как следует из пункта 1) теоремы 3.1, произведение нескольких матриц (если это произведение возможно выполнить) $A_1 A_2 \ldots A_k$ не зависит от расстановки скобок.[24]

Если $A \in \mathbf{M}_{n\times n}$, то можно определить s-ю степень матрицы для натуральных s: $A^s = \underbrace{AA\ldots A}_{s \text{ букв } A}$. Также положим $A^0 = E$. Нетрудно видеть, что для целых неотрицательных s, t верны равенства $A^{s+t} = A^s A^t$, $A^{st} = (A^s)^t$.

Равенство $AB = BA$ для матриц, вообще говоря, не выполнено. Например, произведение AB может быть определено, а BA — нет. В

[21] Проверим, например, 1), 3) и 5). Пусть $A = (a_{ij})$, $A' = (a'_{ij})$, $B = (b_{jk})$, $C = (c_{kl})$.
1) Пусть $AB = (d_{ik})$, $(AB)C = (f_{il})$, $BC = (g_{jl})$, $A(BC) = (h_{il})$. Тогда $f_{il} = \sum\limits_{k=1}^{p} d_{ik}c_{kl} = \sum\limits_{k=1}^{p}\sum\limits_{j=1}^{n} a_{ij}b_{jk}c_{kl} = \sum\limits_{j=1}^{n}\sum\limits_{k=1}^{p} a_{ij}b_{jk}c_{kl} = \sum\limits_{j=1}^{n} a_{ij}g_{jl} = h_{il}$.
3) Пусть $AB = (d_{ik})$, $A + A' = (f_{ij})$, $(A + A')B = (g_{ik})$, $A'B = (d'_{ik})$, $AB + A'B = (h_{ik})$. Тогда $g_{ik} = \sum\limits_{j=1}^{n} f_{ij}b_{jk} = \sum\limits_{j=1}^{n} (a_{ij} + a'_{ij})b_{jk} = \sum\limits_{j=1}^{n} a_{ij}b_{jk} + \sum\limits_{j=1}^{n} a'_{ij}b_{jk} = d_{ik} + d'_{ik} = h_{ik}$.
5) Пусть $AB = (d_{ik})$, $(AB)^T = (f_{ki})$, $A^T = (g_{ji})$, $B^T = (h_{kj})$, $B^T A^T = (t_{ki})$. Тогда $f_{ki} = d_{ik} = \sum\limits_{j=1}^{n} a_{ij}b_{jk} = \sum\limits_{j=1}^{n} g_{ji}h_{kj} = t_{ki}$.

[22] Общая ситуация видна в предложении 3.2.

[23] Отсюда понятен смысл терминов "скалярная"и "единичная".

[24] Формальное доказательство этого факта можно провести так же, как и в случае произведения отображений (см. приложение).

случае $A \in \mathbf{M}_{m\times n}$, $B \in \mathbf{M}_{n\times m}$, $m \neq n$, оба произведения AB и BA определены, но являются матрицами разных размеров. Даже если A и B — квадратные матрицы порядка n (тогда оба произведения AB и BA определены и являются квадратными матрицами порядка n), возможно AB не равно BA, например:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{но} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Из приведенного примера видим также, что произведение двух ненулевых матриц может оказаться нулевой матрицей.

Говорят, что матрицы A и B *перестановочны*, если $AB = BA$.

Можно определить значение многочлена $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ от квадратной матрицы A: положим $f(A) = \sum\limits_{i=0}^{n} a_i A^i$. Нетрудно поверить, что для многочленов f, g, h таких, что $h(x) = f(x)g(x)$, верно $h(A) = f(A)g(A) = g(A)f(A)$. В частности, матрица A перестановочна с любым многочленом от нее.

Иногда матрицы естественно разбиваются на блоки. Например, рассмотрим *блочную матрицу* $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, где $A_{ij} \in \mathbf{M}_{k_i\times l_j}$ (k_1, k_2, l_1, l_2 — некоторые натуральные числа). Оказывается, можно осуществлять "блочное"перемножение матриц с согласованной блочной структурой. Например, произведение блочных матриц AB, где B — блочная матрица $B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$ ($B_1 \in \mathbf{M}_{l_1\times p}$, $B_2 \in \mathbf{M}_{l_2\times p}$), можно вычислять[25] как $AB = \begin{pmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{pmatrix}$.

[25] В общем случае, пусть $A = (A_{ij})$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots n$, где A_{ij} — матрица размера $r_i \times s_j$, $B = (B_{jk})$, $j = 1, 2, \ldots, n$, $k = 1, 2, \ldots p$, где B_{jk} — матрица размера $s_j \times t_k$. Тогда $AB = (C_{ik})$, где $C_{ik} = \sum\limits_{j=1}^{n} A_{ij}B_{jk}$. Доказательство проводится непосредственной проверкой.

Обратимые матрицы. Обратная матрица

Для некоторых квадратных матриц удается определить операцию обращения.[26]

Определение. Матрица $B \in \mathbf{M}_{n\times n}$ называется *обратной* для матрицы $A \in \mathbf{M}_{n\times n}$, если $BA = AB = E$.

Определение. Матрица $A \in \mathbf{M}_{n\times n}$ называется *обратимой*, если для нее существует обратная матрица.

Можно попробовать ослабить условие обратимости: скажем, что матрица $B \in \mathbf{M}_{n\times n}$ является *левой обратной* для матрицы $A \in \mathbf{M}_{n\times n}$, если $BA = E$; квадратные матрицы, имеющие левую обратную, назовем *обратимыми слева*. Аналогично определим *обратимые справа* матрицы. Однако ниже будет доказано (см. теорему 4.3), что для обратимости квадратной матрицы достаточно потребовать, чтобы она была обратима слева или справа.[27]

Предложение 3.1. *Для обратимой матрицы $A \in \mathbf{M}_{n\times n}$ существует единственная левая обратная матрица и единственная правая обратная матрица, причем они равны.*

$\triangleright$ Пусть B — некоторая левая обратная матрица для A, C — некоторая правая обратная матрица для A. Тогда $B = BE = B(AC) = (BA)C = EC = C$. Итак, каждая левая обратная матрица совпадает с C, и, значит, она единственна. Аналогично, правая обратная матрица единственна. $\square$

Таким образом, для обратимой матрицы A существует единственная обратная матрица (она же единственная левая обратная и единственная правая обратная) — будем обозначать ее A^{-1}.

Теорема 3.2. *Пусть $A, B \in \mathbf{M}_{n\times n}$ — обратимые матрицы. Тогда*

1) A^{-1} обратима, причем $(A^{-1})^{-1} = A$;

[26] Ниже мы увидим, что невозможность обратить матрицу эквивалентна ее вырожденности (§ 4) или равенству нулю определителя (§ 6).

[27] Аналогичное утверждение верно для обратимости линейных отображений конечномерных векторных пространств (ввиду соответствия умножения матриц и линейных отображений), но неверное для произвольных отображений — см. приложение.

2) AB обратима, причем $(AB)^{-1} = B^{-1}A^{-1}$;

3) A^T обратима, причем $(A^T)^{-1} = (A^{-1})^T$.

$\triangleright$ Непосредственно проверяется, что матрицы A, $B^{-1}A^{-1}$, $(A^{-1})^T$ являются обратными для матриц A^{-1}, AB, A^T соответственно.[28] $\square$

Следствие. *Если $A_1, A_2, \ldots, A_k \in \mathbf{M}_{n\times n}$ — обратимые матрицы, то $A_1A_2\ldots A_k$ обратима, причем $(A_1A_2\ldots A_k)^{-1} = A_k^{-1}A_{k-1}^{-1}\ldots \ldots A_1^{-1}$.*

Если $A \in \mathbf{M}_{n\times n}$ — обратимая матрица, то можно определить A^s и для неположительных целых s: положим $A^0 = E_n$, $A^s = (A^{-1})^{-s}$, если s — целое отрицательное. Нетрудно проверить, что $\forall\ s, t \in \mathbb{Z}$ верны равенства $A^{s+t} = A^sA^t$ и $A^{st} = (A^s)^t$.

Умножение матриц и ранг

Предложение 3.2. *Пусть $A \in \mathbf{M}_{m\times n}$, $B \in \mathbf{M}_{n\times p}$. Тогда*

1) каждый столбец матрицы AB линейно выражается через столбцы матрицы A;

2) каждая строка матрицы AB линейно выражается через строки матрицы B.

$\triangleright$ 1) Пусть $B = (b_{jk})$, и $(a_{\bullet 1}\ldots a_{\bullet n})$, $(c_{\bullet 1}\ldots c_{\bullet p})$ — столбцовые записи матриц A и $C = AB$. Тогда по определению умножения матриц $c_{\bullet k} = \sum\limits_{j=1}^{n} a_{\bullet j}b_{jk}$, то есть k-й столбец матрицы C равен линейной комбинации столбцов A с коэффициентами из k-го столбца матрицы B.

2) Доказывается аналогично 1), либо сводится к 1) транспонированием матриц ($(AB)^T = B^TA^T$). $\square$

Предложение 3.3. *Пусть $A \in \mathbf{M}_{m\times n}$, $B \in \mathbf{M}_{n\times p}$. Тогда $\mathrm{rg}_v(AB) \leqslant \leqslant \mathrm{rg}_v A$ и $\mathrm{rg}_h(AB) \leqslant \mathrm{rg}_h B$.*

---

[28] Проверим, например, 2) и 3).
2) $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E$. Аналогично проверяется, что $(AB)(B^{-1}A^{-1}) = E$.
3) $(A^{-1})^TA^T = (AA^{-1})^T = E^T = E$. Аналогично $A^T(A^{-1})^T = E$.

$\rhd$ Докажем неравенство про rg_h (рассуждения про rg_v аналогичны).

Пусть $c_{1\bullet}, c_{2\bullet}, \ldots, c_{m\bullet}$ — строки матрицы AB. Согласно предложению 3.2, каждая из строк $c_{1\bullet}, c_{2\bullet}, \ldots, c_{m\bullet}$ равна линейной комбинации строк $b_{1\bullet}, b_{2\bullet}, \ldots, b_{n\bullet}$, поэтому $\mathrm{rg}_h(AB) = \mathrm{rg}(c_{1\bullet}, c_{2\bullet}, \ldots$ $\ldots, c_{m\bullet}) \leqslant \mathrm{rg}\langle b_{1\bullet}, b_{2\bullet}, \ldots, b_{n\bullet}\rangle$. Но из теоремы 2.1 вытекает, что $\mathrm{rg}\langle b_{1\bullet}, b_{2\bullet}, \ldots, b_{n\bullet}\rangle = \mathrm{rg}(b_{1\bullet}, b_{2\bullet}, \ldots, b_{n\bullet}) = \mathrm{rg}_h B$. $\square$

Следующее предложение показывает, что ранг не меняется при домножении на обратимую матрицу.

Предложение 3.4. *Пусть $B \in \mathbf{M}_{m\times n}$ — произвольная матрица, а $A \in \mathbf{M}_{m\times m}$ и $C \in \mathbf{M}_{n\times n}$ — обратимые матрицы. Тогда $\mathrm{rg}_h(AB) = \mathrm{rg}_h B$ и $\mathrm{rg}_v(BC) = \mathrm{rg}_v B$.*

$\rhd$ По предложению 3.3 $\mathrm{rg}_h B \geqslant \mathrm{rg}_h(AB) \geqslant \mathrm{rg}_h(A^{-1}AB) = \mathrm{rg}_h(EB) = \mathrm{rg}_h B$. Аналогично, $\mathrm{rg}_v B \geqslant \mathrm{rg}_v(BC) \geqslant \mathrm{rg}_v(BCC^{-1}) = \mathrm{rg}_h(BE) = \mathrm{rg}_v B$. $\square$

Задачи и упражнения

1. Положим $A = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 3 & -1 \end{pmatrix}$, $C = \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}$. Вычислите ACA^T.

2. Вычислите $(SAS^{-1})^{100}$, если $A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$, $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

3. Сумма всех диагональных элементов матрицы A называется ее *следом* и обозначается $\mathrm{tr}\, A$. Докажите, что для любых матриц $A \in \mathbf{M}_{m\times n}$ и $B \in \mathbf{M}_{n\times m}$ выполнено равенство $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

4. Докажите, что произведение двух симметрических матриц является симметрической матрицей тогда и только тогда, когда эти матрицы перестановочны.

5. Для $A \in \mathbf{M}_{n\times n}$ положим $C_A = \{X \in \mathbf{M}_{n\times n} \mid AX = XA\}$.

   а) Докажите, что $C_A = \mathbf{M}_{n\times n} \Leftrightarrow A$ — скалярная матрица.

б) Докажите, что если $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, где $\lambda_1, \ldots, \lambda_n$ — попарно различные числа, то C_A — множество всех диагональных матриц $n \times n$.

6. Пусть $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, где λ_i — попарно различные числа. Докажите, что любая диагональная матрица порядка n является многочленом от A.[29]

7. а) Докажите, что произведение двух верхнетреугольных матриц — верхнетреугольная матрица.

   б) Докажите, что если $A \in \mathbf{M}_{n\times n}$ — верхнетреугольная матрица, у которой все элементы главной диагонали равны 0, то $A^n = O$.

   в) Вычислите $(E + A)^{10}$, если $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

8. Пусть $A \in \mathbf{M}_{n\times n}$ такова, что $A^m = O$. Докажите, что матрица $E - A$ обратима.

9. Докажите, что при $m < n$ матрица $A \in \mathbf{M}_{m\times n}$ не может иметь "левой обратной"матрицы (то есть не существует матрицы $B \in \in \mathbf{M}_{n\times m}$ такой, что $BA = E_n$).

§ 4. Элементарные преобразования и их применение

Элементарные преобразования и элементарные матрицы

Определим три типа *элементарных преобразований*[30] *строк* матрицы $A \in \mathbf{M}_{m\times n}$. (Аналогично вводятся элементарные преобразования столбцов.)

[29] Интересен вопрос об описании всех матриц A, для которых C_A (см. задачу 5) совпадает с множеством многочленов от A.

[30] Формально говоря, каждое элементарное преобразование — это отображение $\mathbf{M}_{m\times n} \to \mathbf{M}_{m\times n}$.

<u>I тип</u>: строки $a_{i\bullet}$ и $a_{j\bullet}$ $(i \neq j)$ меняются местами (остальные строки остаются без изменения). Обозначим это преобразование $\widetilde{P}_{ij}$.

<u>II тип</u>: строка $a_{i\bullet}$ умножается на число $\lambda \neq 0$. Обозначим это преобразование $\widetilde{D}_i(\lambda)$.

<u>III тип</u>: к строке $a_{i\bullet}$ прибавляется строка $\lambda a_{j\bullet}$ $(i \neq j,\ \lambda \in \mathbb{R})$ (все строки, кроме i-й, остаются без изменения). Обозначим это преобразование $\widetilde{T}_{ij}(\lambda)$.

Определим квадратные *элементарные матрицы* порядка m трех типов как результат применения к единичной матрице $E = E_m$ соответствующего элементарного преобразования.

<u>I тип</u>: $P_{ij} = \widetilde{P}_{ij}(E)$. Заметим, что P_{ij} отличается от единичной матрицы лишь в четырех элементах, расположенных на пересечении i-й и j-й строк с i-м и j-м столбцами; скажем, для $m = 5$ имеем

$$P_{14} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

<u>II тип</u>: $D_i(\lambda) = \widetilde{D}_i(\lambda)(E) = \mathrm{diag}(1, 1, \ldots, 1, \lambda, 1, \ldots, 1)$ (λ на i-м месте).

<u>III тип</u>: $T_{ij}(\lambda) = \widetilde{T}_{ij}(\lambda)(E)$. Матрица $T_{ij}(\lambda)$ отличается от единичной матрицы лишь элементом, равным λ, расположенным на пересечении i-й строки и j-го столбца, скажем, для $m = 5$ имеем

$$T_{25}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \lambda \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Предложение 4.1. *Если S — элементарная матрица, то S^T — элементарная матрица того же типа, что и S.*

$\triangleright$ $P_{ij}^T = P_{ij}$; $D_i(\lambda)^T = D_i(\lambda)$; $T_{ij}(\lambda)^T = T_{ji}(\lambda)$. $\square$

Предложение 4.2. *Элементарное преобразование строк равносильно умножению слева на соответствующую элементарную матрицу.*

$\triangleright$ Для произвольной матрицы $A \in \mathbf{M}_{m\times n}$ непосредственно проверяются (как в доказательстве предложения 3.2) равенства $\widetilde{P}_{ij}(A) = P_{ij}A$, $\widetilde{D}_i(\lambda)(A) = D_i(\lambda)A$, $\widetilde{T}_{ij}(\lambda)(A) = T_{ij}(\lambda)A$. $\square$

Предложение 4.3. *1) Элементарное преобразование строк обратимо, причем обратное преобразование является элементарным преобразованием строк.*

2) Элементарная матрица обратима, причем обратная матрица является элементарной.

$\triangleright$ 1) Легко проверить, что $\widetilde{P}_{ij}^{-1} = \widetilde{P}_{ij}$, $\widetilde{D}_i(\lambda)^{-1} = \widetilde{D}_i(\lambda^{-1})$, $\widetilde{T}_{ij}(\lambda)^{-1} = \widetilde{T}_{ij}(-\lambda)$.

2) Из 1) и предложения 4.2 следует,[31] что ${P_{ij}}^{-1} = P_{ij}$, ${D_i(\lambda)}^{-1} = D_i(\lambda^{-1})$, ${T_{ij}(\lambda)}^{-1} = T_{ij}(-\lambda)$. $\square$

Метод Гаусса

В каждой ненулевой строке матрицы можно выделить первый ненулевой элемент, который будем называть *ведущим.* Таким образом, элемент a_{ij} матрицы A является ведущим элементом ненулевой строки $a_{i\,\bullet}$, если $a_{ij} \neq 0$ и $a_{i1} = a_{i2} = \ldots = a_{i\,j-1} = 0$.

Определение. Скажем, что матрица $A \in \mathbf{M}_{m\times n}$ *ступенчатая* (или *имеет ступенчатый вид*), если для некоторого $r \in \{0, 1, 2, \ldots, m\}$ ее первые r строк ненулевые, а последние $m-r$ строк нулевые, и, кроме того, ведущие элементы $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ первых r строк таковы, что $j_1 < j_2 < \ldots < j_r$.

Например, матрица вида $\begin{pmatrix} 0 & * & x & y & z & t \\ 0 & 0 & 0 & * & u & v \\ 0 & 0 & 0 & 0 & * & w \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, где звездочками обозначены ненулевые числа — ведущие элементы строк, является

[31] Нетрудно проверить и непосредственно.

ступенчатой. Звездочки и элементы правее них образуют "ступеньки"; под "ступеньками"все элементы равны нулю.

Из определения видно, что в ступенчатой матрице квадратная подматрица порядка r, расположенная на пересечении первых r строк $a_{1\bullet}$, $a_{2\bullet}$, $\ldots$, $a_{r\bullet}$ и столбцов $a_{\bullet j_1}$, $a_{\bullet j_2}$, $\ldots$, $a_{\bullet j_r}$, является верхнетреугольной с ненулевыми элементами на диагонали.

Определение. Скажем, что ступенчатая матрица $A \in \mathbf{M}_{m\times n}$ имеет *упрощенный вид*, если в ней ведущие элементы $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ всех ненулевых строк равны 1, и для $k = 1, 2, \ldots r$ в столбце $a_{\bullet j_k}$ все элементы, за исключением a_{kj_k}, равны 0.

Из определения следует, что в матрице упрощенного вида квадратная подматрица порядка r, расположенная на пересечении первых r строк $a_{1\bullet}$, $a_{2\bullet}$, $\ldots$, $a_{r\bullet}$ и столбцов $a_{\bullet j_1}$, $a_{\bullet j_2}$, $\ldots$, $a_{\bullet j_r}$, является единичной матрицей E_r.

Опишем метод Гаусса приведения матрицы элементарными преобразованиями строк к ступенчатому и упрощенному виду. Ниже мы увидим многочисленные применения этого метода.

<u>АЛГОРИТМ</u> приведения матрицы элементарными преобразованиями строк к ступенчатому виду (прямой ход метода Гаусса).

Если A — нулевая матрица, то процесс окончен.

Иначе выполним следующий *шаг* прямого хода метода Гаусса. Найдем ненулевой столбец $a_{\bullet j}$ с наименьшим номером j. Перестановкой строк и умножением первой строки на подходящее число добьемся, чтобы элемент, стоящий в первой строке и j-м столбце, стал равным 1. Итак, предполагаем, что $a_{1j} = 1$. Выполняя последовательно преобразования $\widetilde{T}_{21}(-a_{2j})$, $\widetilde{T}_{31}(-a_{3j})$, $\ldots$, $\widetilde{T}_{m1}(-a_{mj})$, получаем, что первые $j-1$ столбцов остались нулевыми, а в j-м столбце теперь ровно один ненулевой элемент $a_{1j} = 1$. Шаг завершен.

После выполнения шага у матрицы мысленно отбрасываем первую строку. Если оставшаяся матрица ненулевая, выполняем шаг для нее.

Продолжаем так далее, пока не закончатся строки матрицы или пока не получим после очередного шага нулевую матрицу.

В результате выполнения всей процедуры действительно получим ступенчатый вид, так как на каждом новом шаге работаем с матрицей, у которой номер первого ненулевого столбца больше, чем у матрицы на предыдущем шаге.

АЛГОРИТМ приведения матрицы элементарными преобразованиями строк к упрощенному виду (обратный ход метода Гаусса).

Пусть мы уже получили (см. предыдущий алгоритм) ступенчатую матрицу $A = (a_{ij}) \in \mathbf{M}_{m\times n}$, в которой ненулевыми являются первые r строк, $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ — ведущие элементы этих строк $(j_1 < j_2 < \ldots < j_r)$, причем $a_{1j_1} = a_{2j_2} = \ldots = a_{rj_r} = 1$.

Выполняя последовательно $\widetilde{T}_{1r}(-a_{1j_r})$, $\widetilde{T}_{2r}(-a_{2j_r})$, $\ldots$, $\widetilde{T}_{r-1\,r}(-a_{r-1\,j_r})$ (это шаг обратного хода метода Гаусса), превращаем в 0 все элементы j_r-го столбца, за исключением $a_{rj_r} = 1$. При этом матрица остается ступенчатой.

Далее последовательно производим аналогичные действия для j_{r-1}-го, $\ldots$, j_2-го столбцов.

Итак, мы конструктивно доказали следующую теорему.

Теорема 4.1. *Матрица $A \in \mathbf{M}_{m\times n}$ может быть приведена элементарными преобразованиями строк 1) к ступенчатому виду; 2) к упрощенному виду.*

Элементарные преобразования и ранг

Пусть матрица $A = (a_{\bullet 1} \ \ldots \ a_{\bullet n})$ после некоторого элементарного преобразования строк перешла в матрицу $A' = (a'_{\bullet 1} \ \ldots \ a'_{\bullet n})$. Пусть некоторая линейная комбинация столбцов $a_{\bullet j_1}, \ldots, a_{\bullet j_l}$ равна O: $\sum\limits_{s=1}^{l} \lambda_s a_{\bullet j_s} = O$, то есть $\sum\limits_{s=1}^{l} \lambda_s a_{ij_s} = 0$ для всех i. Тогда, как нетрудно видеть, $\sum\limits_{s=1}^{l} \lambda_s a'_{ij_s} = 0$ для всех i, то есть линейная комбинация столбцов $a'_{\bullet j_1}, \ldots, a'_{\bullet j_l}$ с теми же коэффициентами $\lambda_1, \ldots, \lambda_l$ равна O. Из этого соображения вытекает

Предложение 4.4. *Пусть $1 \leqslant j_1 < j_2 < \ldots < j_l \leqslant n$. При элементарных преобразованиях строк матрицы $A \in \mathbf{M}_{m\times n}$ сохраняется линейная зависимость или независимость системы столбцов с номерами $j_1, j_2, \ldots, j_l$.*

$\triangleright$ Пусть матрица $A = (a_{\bullet 1} \ldots a_{\bullet n})$ после некоторого элементарного преобразования строк перешла в матрицу $A' = (a'_{\bullet 1} \ldots a'_{\bullet n})$.

Если подсистема столбцов $a_{\bullet j_1}, \ldots, a_{\bullet j_l}$ матрицы $A \in \mathbf{M}_{m\times n}$ линейно зависима, то, как показано выше, подсистема столбцов $a'_{\bullet j_1}, \ldots, a'_{\bullet j_l}$ также линейно зависима.

Предположим теперь, что система столбцов $a_{\bullet j_1}, \ldots, a_{\bullet j_l}$ линейно независима. Если система $a'_{\bullet j_1}, \ldots, a'_{\bullet j_l}$ линейно зависима, то, рассматривая обратное элементарное преобразование (см. предложение 4.3), получаем по доказанному, что и система столбцов $a_{\bullet j_1}, \ldots, a_{\bullet j_l}$ линейно зависима — противоречие нашему предположению. $\Box$

Предложение 4.5. *Столбцовый ранг матрицы не изменяется при элементарных преобразованиях строк матрицы.*

$\triangleright$ Сразу следует из предыдущего предложения. $\Box$

Следствие. *Пусть A — ступенчатая матрица, в которой ровно r ненулевых строк. Тогда $\mathrm{rg}_v\, A = r$.*

$\triangleright$ Пусть после удаления нулевых строк из матрицы A получается матрица A'. Тогда $\mathrm{rg}_v\, A = \mathrm{rg}_v\, A'$ и $r \geqslant \mathrm{rg}_v\, A'$, поскольку в матрице A' столбцы высоты r (см. предложение 2.4). С помощью обратного хода метода Гаусса из A' можно получить матрицу упрощенного вида B, имеющую подматрицу E_r (см. алгоритм). Используя предложения 4.5, 2.8, имеем $\mathrm{rg}_v\, A' = \mathrm{rg}_v\, B \geqslant \mathrm{rg}_v\, E_r = r$. Итак, мы получили, что $r \geqslant \mathrm{rg}_v\, A = \mathrm{rg}_v\, A' \geqslant r$, значит $\mathrm{rg}_v\, A = r$. $\Box$

Предложение 4.6. *Строчный ранг матрицы не изменяется при элементарных преобразованиях строк.*

$\triangleright$ Следует из предложений 3.4 и 4.2. $\Box$

Теорема 4.2 (о ранге матрицы)**.** *Для любой матрицы $A \in \mathbf{M}_{m\times n}$ выполнено $\mathrm{rg}_v\, A = \mathrm{rg}_h\, A$.*

$\triangleright$ Докажем вначале, что для любой матрицы $\mathrm{rg}_v\, A \geqslant \mathrm{rg}_h\, A$.

Элементарными преобразованиями строк приведем A к ступенчатому виду A', и пусть в A' ровно r ненулевых строк. Тогда $\mathrm{rg}_h\, A' \leqslant r$, $\mathrm{rg}_v\, A' = r$ (см. предыдущее следствие). Из сохранения рангов при элементарных преобразованиях (предложения 4.5 и 4.6) следует, что $\mathrm{rg}_v A = \mathrm{rg}_v A'$ и $\mathrm{rg}_h A = \mathrm{rg}_h A'$, откуда $\mathrm{rg}_v A = r \geqslant \mathrm{rg}_h A$.

Воспользовавшись доказанным неравенством для A^T, имеем $\mathrm{rg}_h A = \mathrm{rg}_v(A^T) \geqslant \mathrm{rg}_h(A^T) = \mathrm{rg}_v A$. $\Box$

Теперь вместо $\mathrm{rg}_v A$ и $\mathrm{rg}_h A$ мы используем одно обозначение $\mathrm{rg}\, A$.

Следствие. *Для любой матрицы A выполнено* $\mathrm{rg}\, A = \mathrm{rg}\, A^T$.

Из доказанного вытекает

АЛГОРИТМ нахождения ранга матрицы и некоторой базисной подсистемы столбцов.[32]

Данную матрицу A приведем элементарными преобразованиями строк к ступенчатому виду A' (методом Гаусса).

Пусть в A' ровно r ненулевых строк (r "ступенек"), и $j_1 < j_2 < \ldots < j_r$ — номера столбцов, содержащих ведущие элементы строк.

Тогда $\mathrm{rg}\, A = r$ (см. следствие из предложения 4.5).

Кроме того, столбцы матрицы A с номерами $j_1, j_2, \ldots, j_r$ являются базисной подсистемой столбцов матрицы A.

Действительно, в ступенчатом (и в упрощенном) виде столбцы с номерами $j_1, j_2, \ldots, j_r$ образуют базисную подсистему, тогда (см. предложение 4.4) и в исходной матрице A система столбцов с номерами $j_1, j_2, \ldots, j_r$ является базисной.

Невырожденные матрицы

Определение. Квадратная матрица $A \in \mathbf{M}_{n\times n}$ называется *невырожденной*, если $\mathrm{rg}\, A = n$, и *вырожденной* в противном случае.

Иначе говоря, в невырожденной матрице система всех столбцов (или строк) линейно независима.

Теорема 4.3 (критерий невырожденности-1). *Для квадратной матрицы $A \in \mathbf{M}_{n\times n}$ следующие условия эквивалентны:*

1) A — невырожденная;

[32] Конечно, в системе столбцов может быть не единственная базисная подсистема.

2) A приводится элементарными преобразованиями строк к единичной матрице E_n;

3) A равна произведению нескольких элементарных матриц;

4) A — обратимая;

5) A — обратимая слева; 5') A — обратимая справа.

$\triangleright$ 5) (или 5')) $\Rightarrow$ 1) Если $BA = E$ (или $AC = E$), то из предложения 3.3 о ранге произведения имеем: $\operatorname{rg} A \geqslant \operatorname{rg} E = n$.

1) $\Rightarrow$ 2) Приведем A к упрощенному виду A'. Так как $\operatorname{rg} A' = \operatorname{rg} A = n$, то A' должна содержать единичную подматрицу порядка n, значит, $A' = E$.

2) $\Rightarrow$ 3) Из 2) и предложения 4.2 следует равенство $S_k S_{k-1} \ldots S_1 A = E$, где S_i — элементарные матрицы. Тогда домножая это равенство слева последовательно на элементарные матрицы $S_k^{-1}, \ldots, S_1^{-1}$ (см. предложение 4.3), получим $A = S_1^{-1} S_2^{-1} \ldots S_k^{-1}$.

3) $\Rightarrow$ 4) Вытекает из теоремы 3.2 (произведение обратимых матриц — обратимая матрица).

4) $\Rightarrow$ 5) (или 5')) Очевидно. $\square$

Предложение 4.7. *Пусть $A \in \mathbf{M}_{n\times n}$ — обратимая матрица, $B \in \mathbf{M}_{n\times p}$. Пусть (блочная) матрица $(A\,B) \in \mathbf{M}_{n\times(n+p)}$ приведена элементарными преобразованиями строк к виду $(E_n\,C)$. Тогда $C = A^{-1}B$.*[33]

$\triangleright$ Выполнение элементарных преобразований эквивалентно домножению слева на некоторую (обратимую) матрицу $R \in \mathbf{M}_{n\times n}$ (см. предложение 4.2): $(E\,C) = R(A\,B)$. Тогда $RA = E$, $RB = C \Rightarrow R = A^{-1}$ и $C = A^{-1}B$. $\square$

Последнее предложение дает, в частности,

<u>АЛГОРИТМ</u> отыскания A^{-1}.

К данной матрице $A \in \mathbf{M}_{n\times n}$ припишем справа единичную матрицу $E = E_n$, получим матрицу $(A\,E)$. Элементарными преобразованиями строк (метод Гаусса) приводим матрицу $(A\,E)$ к виду

[33] Это предложение дает, в частности, рецепт решения матричного уравнения $AX = B$ с невырожденной матрицей A.

$(A'\,B)$, где A' — ступенчатая матрица. Если в A' количество "ступенек"меньше n, то $\operatorname{rg} A < n$ (см. алгоритм отыскания ранга), и, значит, матрица A не имеет обратной. Иначе, продолжив метод Гаусса, приведем матрицу $(A\,E)$ к упрощенному виду $(E\,C)$. Тогда $C = A^{-1}$.

Пусть подматрица размера $k \times k$ матрицы A, расположенная на пересечении некоторой системы k столбцов и k строк, оказалась невырожденной. Тогда ясно, что эта система из k столбцов (строк) матрицы A линейно независима, в частности, $k \leqslant \operatorname{rg} A$. Оказывается, для подматриц размера $r \times r$, где $r = \operatorname{rg} A$, верно и обратное: в любой матрице ранга r найдется невырожденная подматрица порядка r.[34]

Теорема 4.4 (о базисном миноре)**.** *Пусть $A \in \mathbf{M}_{m\times n}$ и $\operatorname{rg} A = r$. Тогда квадратная подматрица порядка r, полученная на пересечении некоторой линейно независимой системы из r столбцов и некоторой линейно независимой системы из r строк, является невырожденной.*

$\triangleright$ Пусть для определенности система первых r строк $a_{1\bullet}$, $a_{2\bullet}$, $\ldots$, $a_{r\bullet}$ линейно независима (то есть базисная подсистема системы строк матрицы A), и система первых r столбцов $a_{\bullet 1}$, $a_{\bullet 2}$, $\ldots, a_{\bullet r}$ линейно независима. Докажем, что подматрица $A' \in \mathbf{M}_{r\times r}$, полученная в пересечении первых r строк и первых r столбцов, невырожденная.

Из теоремы 2.1 вытекает, что каждая из строк $a_{r+1\,\bullet}$, $\ldots$, $a_{m\,\bullet}$ является линейной комбинацией строк $a_{1\bullet}$, $a_{2\bullet}$, $\ldots$, $a_{r\bullet}$. Поэтому элементарными преобразованиями III типа можно строки $a_{r+1\,\bullet}, \ldots$ $\ldots, a_{m\,\bullet}$ последовательно сделать нулевыми (вычесть соответствующую линейную комбинацию первых r строк). После выполнения этих преобразований матрица имеет вид $\begin{pmatrix} A' & B \\ O & O \end{pmatrix}$. Но согласно предложению 4.4, система первых r столбцов осталась линейно независимой, то есть $\operatorname{rg} \begin{pmatrix} A' \\ O \end{pmatrix} = r$. При отбрасывании нулевых строк ранг не изменится, поэтому $\operatorname{rg} A' = r$. $\square$

[34] Иногда такую подматрицу называют *базисным минором.*

Задачи и упражнения

1. Докажите, что элементарное преобразование строк I типа можно получить последовательным выполнением нескольких элементарных преобразований строк II и III типа.

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

   Укажите в ней некоторую базисную подсистему столбцов, строк. Найдите невырожденную подматрицу порядка $\operatorname{rg} A$.

3. Дана верхнетреугольная матрица A, у которой на главной диагонали ненулевые элементы. Докажите, что A невырожденная (= обратимая), причем A^{-1} также верхнетреугольная.

4. а) Выразите ранг блочной матрицы $\begin{pmatrix} A & -3B \\ 2A & B \end{pmatrix}$ через ранги матриц $A \in \mathbf{M}_{m\times n}$ и $B \in \mathbf{M}_{m\times p}$.

   б) Докажите, что для любых матриц $A, B \in \mathbf{M}_{n\times n}$ ранг блочной матрицы $\begin{pmatrix} A & E_n \\ BA & B \end{pmatrix}$ равен n.

5. Докажите усиление предложения 3.4: Пусть $B \in \mathbf{M}_{m\times n}$ — произвольная матрица, а $A \in \mathbf{M}_{m'\times m}$ и $C \in \mathbf{M}_{n\times n'}$ — такие матрицы, что $\operatorname{rg} A = m$ и $\operatorname{rg} C = n$. Тогда $\operatorname{rg}(AB) = \operatorname{rg} B = \operatorname{rg}(BC)$.

6. Докажите, что матрицу $A \in \mathbf{M}_{m\times n}$ ранга r можно представить в виде произведения BC, где $B \in \mathbf{M}_{m\times r}$, $C \in \mathbf{M}_{r\times n}$.

7. а) Докажите, что элементарными преобразованиями строк данную матрицу можно привести к единственному упрощенному виду.

   б) Матрицы $A, B \in \mathbf{M}_{m\times n}$ могут быть получены друг из друга умножением слева на некоторые матрицы размера $m \times m$ тогда и только тогда, когда A и B проводятся элементарными преобразованиями строк к одному и тому же упрощенному виду.

8. Предложите алгоритм решения матричного уравнения $XA = B$ с невырожденной матрицей A.

§ 5. Системы линейных уравнений

Формы записи и критерий совместности

Рассмотрим систему m линейных уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \qquad\qquad \ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{2}$$

где a_{ij} и b_i — некоторые числа.

Система (2) записывается в виде одного матричного линейного уравнения $AX = b$, где $A \in \mathbf{M}_{m\times n}$ — *матрица коэффициентов*, $X \in \mathbf{M}_{n\times 1}$ — *столбец неизвестных*, $b \in \mathbf{M}_{m\times 1}$ — *столбец правых частей*:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Система (2) задается *расширенной матрицей* $(A\,|\,b) \in \mathbf{M}_{m\times(n+1)}$. Итак, систему (2) коротко записывают в виде $AX = b$ или $(A\,|\,b)$.

Также система (2) может быть переписана в *столбцовой* записи: $\sum\limits_{j=1}^{n} x_j a_{\bullet j} = b$. Получается следующая интерпретация (2): столбец b равен линейной комбинации столбцов матрицы A с коэффициентами $x_1, x_2, \ldots, x_n$. Таким образом, решить систему (2) — значит найти все разложения столбца правых частей b по столбцам матрицы A. Из столбцовой записи видно, что система (2) не изменится, если поменять местами i-й и j-й столбцы матрицы A и одновременно поменять местами x_i и x_j в столбце неизвестных X.

Определение. Столбец $X_0 \in \mathbf{M}_{n\times 1}$ называется *частным решением* (иногда просто *решением*) системы $AX = b$, если $AX_0 = b$.

Определение. Множество всех частных решений системы $AX = b$ называется *общим решением* системы.

Таким образом, общее решение — это подмножество $\{X | AX = b\}$ в множестве столбцов $\mathbf{M}_{n\times 1}$. Общее решение системы $AX = b$ для краткости будем обозначать $\mathrm{Sol}(A\,|\,b)$.[35]

Определение. Система $AX = b$ называется *совместной*, если $\mathrm{Sol}(A\,|\,b) \neq \varnothing$ (то есть имеется хотя бы одно частное решение).

Определение. Две совместные системы линейных уравнений называются *эквивалентными*, если они имеют одно и то же общее решение.

Предложение 5.1. *Система линейных уравнений $AX = b$ совместна $\Leftrightarrow b \in \langle a_{\bullet 1}, a_{\bullet 2}, \ldots, a_{\bullet n}\rangle$.*

$\triangleright$ Ясно из столбцовой записи системы линейных уравнений. $\square$

Теорема 5.1 (критерий Кронекера–Капелли)**.** *Система $AX = b$ совместна $\Leftrightarrow \operatorname{rg} A = \operatorname{rg}(A\,|\,b)$.*

$\triangleright$ Следует из основной теоремы о рангах 2.1 и предыдущего предложения. $\square$

Далее мы опишем структуру и алгоритм нахождения общего решения. [36]

---

[35]Система уравнений и ее общее решение имеет и другие интерпретации в линейной алгебре. Например, $\mathrm{Sol}(A\,|\,b)$ — это полный прообраз столбца b при линейном отображении $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, заданном матрицей A. В частности, $\mathrm{Sol}(A\,|\,O)$ — это *ядро* отображения φ.

[36]Проясним связь с предыдущим, рассмотрев частный случай. Пусть матрица A квадратная (то есть $m = n$) и невырожденная (= обратимая). Тогда уравнение $AX = b$ легко разрешить в матричном виде, домножив слева на A^{-1}: $A^{-1}AX = X = A^{-1}b$. Решение $X = A^{-1}b$ (подстановкой $A(A^{-1}b) = b$ убеждаемся, что найденный столбец в самом деле является решением) можно найти методом Гаусса — см. предложение 4.7. Ниже метод Гаусса используется для решения систем линейных уравнений с произвольной матрицей коэффициентов.

Однородные системы линейных уравнений, структура решения

Определение. Система $AX = b$ называется *однородной*, если $b = O$ (то есть столбец b нулевой).

Вместе с каждой системой $AX = b$ будем рассматривать *соответствующую* однородную систему $AX = O$. Ясно, что однородная система всегда совместна, так как нулевой столбец является ее решением.

Предложение 5.2. *Если $X_1, X_2, \dots, X_k$ — частные решения однородной системы $AX = O$, то линейная комбинация $\sum\limits_{i=1}^{k} \lambda_i X_i$ также является решением.*[37]

$\triangleright$ Сразу следует из матричного равенства (см. теорему 3.1) $A(\sum\limits_{i=1}^{k} \lambda_i X_i) = \sum\limits_{i=1}^{k} \lambda_i A X_i = O$. $\square$

Определение. Упорядоченную базисную подсистему столбцов в множестве $\mathrm{Sol}(A\,|\,O)$ называем *фундаментальной системой решений* однородной системы $AX = O$ (далее — ФСР системы $AX = O$).

Определение. Матрица $\Phi \in \mathbf{M}_{n \times s}$ называется *фундаментальной* матрицей однородной системы $AX = O$, если ее столбцы образуют ФСР.

Если однородная система линейных уравнений имеет только нулевое решение, то полагаем, что ФСР пуста.[38]

Предложение 5.3. *1) Для любой однородной системы $AX = O$ существует ФСР.*

---

[37] Для знакомых с понятием векторного пространства и подпространства: это предложение означает, что $\mathrm{Sol}(A\,|\,O)$ — подпространство в векторном пространстве $M_{n\times 1}$.

[38] Для знакомых с началами линейной алгебры: ФСР системы $AX = O$ — это базис в пространстве $\mathrm{Sol}(A\,|\,O)$. Если подпространство нулевое, то базис считаем пустым — в согласии с равенством $\langle \varnothing \rangle = O$.

2) Количество столбцов в ФСР системы $AX = O$ не зависит от выбора ФСР и равно равно $\mathrm{rg}(\mathrm{Sol}(A\,|\,O))$.

$\triangleright$ Вытекает из следствия из теоремы 2.1. $\square$

Из предыдущего вытекают следующие теоремы о структуре общего решения системы линейных уравнений.

Теорема 5.2. *Общее решение* $\mathrm{Sol}(A\,|\,O)$ *однородной системы линейных уравнений имеет вид* $\left\{\sum\limits_{i=1}^{s}\lambda_i X_i \,|\, \lambda_i \in \mathbb{R}\right\}$, *где* $X_1, X_2, \ldots, X_s$ *— некоторая фиксированная ФСР.*

$\triangleright$ Сразу следует из предложения 5.2 и определения ФСР. $\square$

Таким образом, общее решение однородной системы линейных уравнений можно представить в следующем матричном виде: $\mathrm{Sol}(A\,|\,O) = \Phi\Lambda$, где $\Phi \in \mathbf{M}_{n\times s}$ — фундаментальная матрица, а $\Lambda = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_s\end{pmatrix}$ — столбец произвольных чисел.

Теорема 5.3. *Общее решение* $\mathrm{Sol}(A\,|\,b)$ *совместной системы* $AX = b$ *имеет вид* $X_0 + \overline{X}$, *где* X_0 *— некоторое частное решение системы* $\mathrm{Sol}(A\,|\,b)$, *а* $\overline{X}$ *пробегает* $\mathrm{Sol}(A\,|\,O)$.[39]

$\triangleright$ Пусть X — произвольный столбец высоты n. Положим $\overline{X} = X - X_0$, где X_0 — частное решение системы $(A\,|\,b)$. Имеем $AX = b \Leftrightarrow A(X_0 + \overline{X}) = b \Leftrightarrow AX_0 + A\overline{X} = b \Leftrightarrow b + A\overline{X} = b \Leftrightarrow A\overline{X} = O$. $\square$

Решение системы линейных уравнений методом Гаусса

Введем *элементарные преобразования* системы линейных уравнений трех типов, соответствующие элементарным преобразованиям

[39] В геометрических терминах: если $AX = b$ совместна, и $\mathrm{Sol}(A\,|\,O)$ — s-мерное подпространство, то $\mathrm{Sol}(A\,|\,b)$ можно получить сдвигом подпространства $\mathrm{Sol}(A\,|\,O)$ на вектор, то есть $\mathrm{Sol}(A\,|\,b)$ — *s-мерная плоскость.*

строк расширенной матрицы $(A\,|\,b)$: перестановка двух уравнений; умножение уравнения на $\lambda \neq 0$; прибавление к одному уравнению другого, умноженного на λ.

Предложение 5.4. *При элементарных преобразованиях строк расширенной матрицы* $(A\,|\,b)$ *общее решение* $\mathrm{Sol}(A\,|\,b)$ *не меняется.*

$\rhd$ Для элементарных преобразований I и II типа это очевидно. При элементарном преобразовании III типа $(A\,|\,b) \overset{III}{\to} (A'|b')$ появляется новое уравнение, которое является следствием системы $(A\,|\,b)$, поэтому $\mathrm{Sol}(A\,|\,b) \subset \mathrm{Sol}(A'\,|\,b')$. Рассматривая обратное элементарное преобразование, доказываем обратное включение $\mathrm{Sol}(A'\,|\,b') \subset \subset \mathrm{Sol}(A\,|\,b)$. $\square$

Пользуясь последним предложением, опишем

АЛГОРИТМ решения системы линейных уравнений.

Пусть система линейных уравнений задана расширенной матрицей $(A\,|\,b)$.

1) Приводим матрицу $(A\,|\,b)$ элементарными преобразованиями строк к ступенчатому виду (выполняем прямой ход метода Гаусса).

Уже после этой процедуры можно дать ответ на вопрос о совместности системы. Если последняя ненулевая строка в ступенчатом виде имеет вид $(0,\ldots,0|t)$, где $t \neq 0$, то соответствующее уравнение $0 = t$ противоречиво, и система несовместна. Иначе продолжим действия и увидим, что они приведут нас к непустому множеству решений.[40]

Пусть в полученном ступенчатом виде ровно r ненулевых строк (нулевые строки можно просто отбросить), и $j_1 < j_2 < \ldots < j_r \leqslant n$ — номера столбцов, содержащих ведущие элементы строк. Назовем *главными неизвестными* неизвестные $x_{j_1},\ldots,x_{j_r}$, а остальные неизвестные — *свободными*.

2) Приводим $(A\,|\,b)$ элементарными преобразованиями строк к упрощенному виду (выполняем обратный ход метода Гаусса).

[40]Эти наблюдения о совместности находятся в согласии с критерием Кронекера–Капелли: если после приведения к ступенчатому виду один из ведущих элементов строк расположен в последнем столбце, то $\mathrm{rg}(A\,|\,b) = \mathrm{rg}\,A + 1$, и система несовместна; иначе $\mathrm{rg}(A\,|\,b) = \mathrm{rg}\,A$, и система совместна. Фактически приводимый здесь алгоритм дает другое доказательство критерия Кронекера–Капелли.

3) (Изменение порядка неизвестных.) Переобозначим неизвестные так, чтобы $x'_1 = x_{j_1}$, $x'_2 = x_{j_2}$, $\ldots$, $x'_r = x_{j_r}$ стали главными, а $x'_{r+1}, x'_{r+2}, \ldots, x'_n$ — свободными, одновременно меняя местами столбцы матрицы. После этого единичная подматрица порядка r переместится в первые r столбцов, и расширенная матрица системы будет иметь (блочный) вид $\left(E_r\, R \,|\, \widetilde{b}\right)$, где $R \in \mathbf{M}_{r\times(n-r)}$ — некоторая матрица. Теперь систему можно записать в виде $(E\,R)X' = \widetilde{b}$, где $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$.

4) (результат) Оказывается, теперь общее решение системы может быть записано в виде[41]

$$X' = \begin{pmatrix} \widetilde{b} \\ O \end{pmatrix} + \begin{pmatrix} -R \\ E_{n-r} \end{pmatrix} \Lambda, \qquad (3)$$

где $\Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-r} \end{pmatrix}$ — столбец произвольных чисел. Докажем, что это действительно так. Перепишем систему в виде $EX'_{\text{гл.}} + RX'_{\text{св.}} = \widetilde{b}$ $\Leftrightarrow X'_{\text{гл.}} + RX'_{\text{св.}} = \widetilde{b}$, где $X'_{\text{гл.}} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_r \end{pmatrix}$ и $X'_{\text{св.}} = \begin{pmatrix} x'_{r+1} \\ \vdots \\ x'_n \end{pmatrix}$ — столбцы главных и свободных неизвестных.

Положим $X'_{\text{св.}} = O$, $X'_{\text{гл.}} = \widetilde{b}$, получаем верное равенство, значит, в качестве частного решения можно взять (блочный) столбец $\begin{pmatrix} \widetilde{b} \\ O \end{pmatrix}$.

Остается найти общее решение однородной системы $X'_{\text{гл.}} + RX'_{\text{св.}} = O$. Придавая свободным неизвестным произвольные значения, то есть полагая $X'_{\text{св.}} = \Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-r} \end{pmatrix}$, однозначно находим

[41] Отметим, что результат находится в соответствии с найденной в теореме 5.3 структурой общего решения.

$X'_{\text{гл.}} = -R\Lambda$. Отсюда $X' = \begin{pmatrix} X'_{\text{гл.}} \\ X'_{\text{св.}} \end{pmatrix} = \begin{pmatrix} -R\Lambda \\ \Lambda \end{pmatrix} = \begin{pmatrix} -R \\ E_{n-r} \end{pmatrix} \Lambda$. Нетрудно видеть, что $\Phi = \begin{pmatrix} -R \\ E_{n-r} \end{pmatrix}$ является фундаментальной матрицей системы $X'_{\text{гл.}} + RX'_{\text{св.}} = O$ (столбцы матрицы Φ образуют линейно независимую систему, так как Φ содержит подматрицу E_{n-r}).

5) (Возврат к исходному порядку неизвестных.) В матричном равенстве (3) надо переставить строки так, чтобы первый столбец стал столбцом $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Предложение 5.5. *Пусть* $A \in \mathbf{M}_{m\times n}$, $\operatorname{rg} A = r$. *Тогда*

1) число столбцов в (любой) ФСР системы $AX = O$ *равно* $n - r$.

2) $\operatorname{rg}(\operatorname{Sol}(A\,|\,O)) = n - r$.[42]

$\triangleright$ 1) Из описанного алгоритма вытекает, что одна из ФСР системы $AX = O$ содержит ровно $n - r$ столбцов.[43]

2) Следует из 1) и предложения 5.3. $\square$

Следствие. *Если* $n > m$ *(число неизвестных больше числа уравнений), то однородная система* $AX = O$ *имеет ненулевое решение.*

Двойственность

В предыдущем пункте мы научились по матрице $A \in \mathbf{M}_{m\times n}$ однородной системы $AX = O$ находить такую матрицу Φ, для которой $\operatorname{Sol}(A\,|\,O)$ является линейной оболочкой столбцов Φ. Оказывается матрица A^T играет аналогичную роль для Φ^T, иными словами, верна следующая теорема двойственности.

[42]Это утверждение можно переформулировать так: сумма размерностей ядра и образа линейного отображения $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, заданного матрицей A, равна n.

[43]Согласно предложению 5.3, отсюда следует, что и любая другая ФСР системы $AX = 0$ содержит ровно $n - r$ столбцов.

Предложение 5.6. *Пусть даны матрицы* $A = \begin{pmatrix} a_{1\,\bullet} \\ \vdots \\ a_{m\,\bullet} \end{pmatrix} \in \mathbf{M}_{m\times n}$ *и* $B = \begin{pmatrix} b_{1\,\bullet} \\ \vdots \\ b_{p\,\bullet} \end{pmatrix} \in \mathbf{M}_{p\times n}$. *Тогда* $\operatorname{Sol}(A\,|\,O) = \langle b_{1\,\bullet}^T, \dots, b_{p\,\bullet}^T\rangle \Leftrightarrow$ $\Leftrightarrow \operatorname{Sol}(B\,|\,O) = \langle a_{1\,\bullet}^T, \dots, a_{m\,\bullet}^T\rangle$.[44]

$\triangleright$ Докажем $\boxed{\Rightarrow}$ (обратное следствие доказывается так же).

Пусть $\operatorname{Sol}(A\,|\,O) = \langle b_{1\,\bullet}^T, \dots, b_{p\,\bullet}^T\rangle$. Так как $b_{j\bullet}^T$ — решения системы $AX = O$, то $a_{i\,\bullet} b_{j\bullet}^T = 0$ для всех $i = 1, \dots, m$, $j = 1, \dots, p$. Транспонирование этих равенств дает $b_{j\bullet} a_{i\,\bullet}^T = 0$ для всех $i = 1, \dots, m$, $j = 1, \dots, p$. Это означает, что столбцы $a_{i\,\bullet}^T$ являются решениями системы $BX = O$: $a_{1\,\bullet}^T \dots a_{m\,\bullet}^T \in \operatorname{Sol}(B\,|\,O)$. Тогда по предложению 5.2 $\langle a_{1\,\bullet}^T \dots a_{m\,\bullet}^T\rangle \subset \operatorname{Sol}(B\,|\,O)$.

Также (по предложению 5.5 и теореме 2.1) $n - \operatorname{rg} A = \operatorname{rg}(\operatorname{Sol}(A\,|\,O)) = \operatorname{rg}(b_{1\,\bullet}^T, \dots, b_{p\,\bullet}^T) = \operatorname{rg} B$, откуда $\operatorname{rg}(\operatorname{Sol}(B\,|\,O)) = n - \operatorname{rg} B = \operatorname{rg} A = \operatorname{rg}(a_{1\,\bullet}^T \dots a_{m\,\bullet}^T) = \operatorname{rg}\langle a_{1\,\bullet}^T \dots a_{m\,\bullet}^T\rangle$.

Из доказанного равенства рангов и включения $\langle a_{1\,\bullet}^T \dots a_{m\,\bullet}^T\rangle \subset \operatorname{Sol}(B\,|\,O)$ получаем (снова по теореме 2.1) $\langle a_{1\,\bullet}^T \dots a_{m\,\bullet}^T\rangle = \operatorname{Sol}(B\,|\,O)$. $\square$

Из предыдущего предложения мы получаем

<u>АЛГОРИТМ</u> восстановления однородной системы по множеству ее решений.

Пусть дана система столбцов $X_1, \dots, X_s \in \mathbf{M}_{n\times 1}$. Найдем некоторую матрицу[45] A такую, что $\operatorname{Sol}(A\,|\,O) = \langle X_1, \dots, X_s\rangle$.

Достаточно записать матрицу Φ, состоящую из столбцов $X_1, \dots$ $\dots, X_s$, найти фундаментальную матрицу Ψ системы $\Phi^T X = O$, и взять $A = \Psi^T$.

---

[44] Приведем следующую трактовку этого предложения. Пусть в пространстве $\mathbf{M}_{n\times 1}$ введена структура евклидова пространства со "стандартным"скалярным произведением $(X, Y) = X^T Y$. Тогда подпространства $\operatorname{Sol}(A\,|\,O)$ и $\langle a_{1\,\bullet}^T, \dots, a_{m\,\bullet}^T\rangle$ — ортогональные дополнения друг для друга.

[45] Вообще, таких матриц может быть бесконечно много.

Задачи и упражнения

1. Решите систему $\begin{cases} 2x_3 + x_4 - x_5 = 2, \\ -4x_3 - 2x_4 + 2x_5 = -4, \\ 3x_1 - 6x_2 + 5x_3 + x_4 + 2x_5 = 5, \\ -x_1 + 2x_2 + 5x_3 + 3x_4 - 7x_5 = 2. \end{cases}$

2. Найдите хотя бы одну однородную систему, имеющую множество решений $\langle \begin{pmatrix} 0 \\ -3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \\ 2 \\ -1 \end{pmatrix} \rangle$.

3. Решите матричное уравнение $AXB = C$, где $C \in \mathbf{M}_{m\times n}$ — произвольная матрица, а $A \in \mathbf{M}_{m\times m}$ и $B \in \mathbf{M}_{n\times n}$ — невырожденные матрицы.

4. Пусть $\Phi \in \mathbf{M}_{n\times s}$ — некоторая фундаментальная матрица для системы $AX = O$. Докажите, что множество всех фундаментальных матриц имеет вид ΦS, где S пробегает все невырожденные матрицы порядка s.

5. Докажите, что две совместные системы эквивалентны тогда и только тогда, когда каждое уравнение одной из них является линейной комбинацией уравнений другой системы.

6. Пусть $A \in \mathbf{M}_{m\times n}$. Пусть $\Phi \in \mathbf{M}_{n\times s}$ — фундаментальная матрица системы $\mathrm{Sol}(A\,|\,O)$. Система строк матрицы Φ с номерами $j_1, j_2, \ldots, j_s$ является базисной для Φ тогда и только тогда, когда линейно независима система столбцов матрицы A, полученная после вычеркивания столбцов с номерами $j_1, j_2, \ldots, j_s$.

§ 6. Определитель (детерминант)

Детерминант квадратной матрицы может быть определен различными эквивалентными путями. Как мы увидим ниже, детерминант матрицы является линейной и кососимметричной функцией строк или столбцов матрицы.[46] Для вычисления определителя можно действовать следующим образом: элементарными преобразованиями строк и столбцов приводят (отслеживая изменение определителя) матрицу к матрице верхнетреугольного или нижнетреугольного вида, для которой определитель равен произведению чисел, стоящих на главной диагонали.

Индуктивное определение

Подматрицу размера $(m-1)\times(n-1)$, полученную из матрицы $A \in \mathbf{M}_{m\times n}$ вычеркиванием строки $a_{i\bullet}$ и столбца $a_{\bullet j}$, назовем *дополнительной подматрицей*, отвечающей элементу a_{ij}. Дополнительную подматрицу, отвечающую элементу a_{ij}, будем обозначать A_{ij}.

Каждой квадратной матрице A поставим в соответствие число, называемое *определителем* или *детерминантом* матрицы A. Определитель обозначают $|A|$ или $\det A$.

Для матрицы $A=(a_{11})$ порядка 1 положим $|A|=a_{11}$.

Предполагая, что детерминант уже определен для квадратных матриц порядка $n-1$, зададим $|A|$ для матрицы $A=(a_{ij})$ порядка n следующей формулой *разложения по первому столбцу*:

$$|A| = \sum_{i=1}^{n}(-1)^{i+1}a_{i1}|A_{i1}|. \tag{4}$$

Основные свойства

Теорема 6.1. *Если $A=(a_{ij})$ — верхнетреугольная матрица порядка n, то $|A|=a_{11}a_{22}\dots a_{nn}$.*

$\triangleright$ Применим индукцию по n. Для матриц порядка 1 утверждение верно. Пусть утверждение верно для матриц порядка $n-1$.

[46] Как можно заметить, это и другие свойства определителя находятся в согласии с его следующей геометрической интерпретацией. Пусть в некотором базисе $\mathbf{e}_1,\dots,\mathbf{e}_n$ n-мерного векторного пространства n векторов $\mathbf{a}_1,\dots,\mathbf{a}_n$ заданы координатными столбцами, которые записаны в матрицу A размера $n\times n$. Тогда определитель матрицы A — это отношение ориентированных объемов параллелепипедов $\dfrac{V_\pm(\mathbf{a}_1,\dots,\mathbf{a}_n)}{V_\pm(\mathbf{e}_1,\dots,\mathbf{e}_n)}$.

Все элементы первого столбца матрицы A, кроме возможно a_{11}, равны 0, поэтому $|A| = a_{11}|A_{11}|$. Но A_{11} — верхнетреугольная матрица порядка $n-1$ с диагональными элементами $a_{22}, \ldots, a_{nn}$, значит, $|A_{11}| = a_{22} \ldots a_{nn}$. $\square$

Следствие. $|\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)| = \lambda_1\lambda_2 \ldots \lambda_n$, *в частности,* $|E| = 1$; $|D_i(\lambda)| = \lambda$.

Теорема 6.2 (линейность по строкам). *Пусть даны квадратные матрицы* $A = \begin{pmatrix} a_{1\bullet} \\ \vdots \\ a_{n\bullet} \end{pmatrix}$, $A' = \begin{pmatrix} a'_{1\bullet} \\ \vdots \\ a'_{n\bullet} \end{pmatrix}$, $A'' = \begin{pmatrix} a''_{1\bullet} \\ \vdots \\ a''_{n\bullet} \end{pmatrix}$, *которые отличаются только в одной строке, то есть для некоторого номера* k *выполнено* $a_{i\bullet} = a'_{i\bullet} = a''_{i\bullet}$ *при* $i \neq k$. *Тогда*

1) если $a_{k\bullet} = a'_{k\bullet} + a''_{k\bullet}$, *то* $|A| = |A'| + |A''|$;

2) если $a_{k\bullet} = \lambda a'_{k\bullet}$, *то* $|A| = \lambda|A'|$.

$\triangleright$ Применим индукцию по n. Для матриц порядка 1 утверждение верно. Предполагая, что утверждение верно для матриц порядка $n-1$, докажем его для матриц A, A', A'' порядка n.

1) Запишем разложение $|A|$ по первому столбцу, выделив отдельно k-е слагаемое: $|A| = \sum\limits_{i\neq k}(-1)^{i+1}a_{i1}|A_{i1}| + (-1)^{k+1}a_{k1}|A_{k1}|$. При $i \neq k$ матрицы A_{i1}, A'_{i1} и A''_{i1}, отличаются в одной строке, причем соответствующая строка матрицы A_{i1} равна сумме строк матриц A'_{i1} и A''_{i1}. Согласно предположению индукции $|A_{i1}| = |A'_{i1}| + |A''_{i1}|$ при $i \neq k$. Учитывая, что $A_{k1} = A'_{k1} = A''_{k1}$ и $a_{i1} = a'_{i1} = a''_{i1}$ при $i \neq k$, имеем

$$|A| = \sum_{i\neq k}(-1)^{i+1}a_{i1}(|A'_{i1}| + |A''_{i1}|) + (-1)^{k+1}(a'_{k1} + a''_{k1})|A_{k1}| =$$

$$= \left(\sum_{i\neq k}(-1)^{i+1}a_{i1}|A'_{i1}| + (-1)^{k+1}a'_{k1}|A_{k1}|\right) +$$

$$+ \left(\sum_{i\neq k}(-1)^{i+1}a_{i1}|A''_{i1}| + (-1)^{k+1}a''_{k1}|A_{k1}|\right) =$$

$$= \left(\sum_{i \neq k} (-1)^{i+1} a'_{i1} |A'_{i1}| + (-1)^{k+1} a'_{k1} |A'_{k1}| \right) +$$

$$+ \left(\sum_{i \neq k} (-1)^{i+1} a''_{i1} |A''_{i1}| + (-1)^{k+1} a''_{k1} |A''_{k1}| \right) = |A'| + |A''|.$$

2) Аналогично 1), $|A| = \sum\limits_{i \neq k} (-1)^{i+1} a_{i1} |A_{i1}| + (-1)^{k+1} a_{k1} |A_{k1}| =$

$$= \sum_{i \neq k} (-1)^{i+1} a'_{i1} (\lambda |A'_{i1}|) + (-1)^{k+1} (\lambda a'_{k1}) |A'_{k1}| =$$

$$= \lambda \left(\sum_{i \neq k} (-1)^{i+1} a'_{i1} |A'_{i1}| + (-1)^{k+1} a'_{k1} |A'_{k1}| \right) = \lambda |A'|. \ \square$$

Следствие 1. *При выполнении элементарного преобразования строк $\widetilde{D}_i(\lambda)$ II типа определитель умножается на λ.*

Следствие 2. *Определитель матрицы с нулевой строкой равен 0.*

$\triangleright$ В утверждении 2) теоремы достаточно положить $\lambda = 0$. $\square$

Далее докажем, что определитель изменит знак, если поменять местами две строки матрицы. Вначале рассмотрим случай двух соседних строк.

Лемма. *Пусть квадратные матрицы A и A' таковы, что (для некоторого номера k) $\widetilde{P}_{k\,k+1}(A) = A'$. Тогда $|A'| = -|A|$.*

$\triangleright$ Для матриц порядка 1 нечего доказывать. Предполагая, что утверждение верно для матриц порядка $n-1$, докажем его для матриц порядка n.

Положим $A = \begin{pmatrix} a_{1\bullet} \\ \vdots \\ a_{n\bullet} \end{pmatrix}$, $A' = \begin{pmatrix} a'_{1\bullet} \\ \vdots \\ a'_{n\bullet} \end{pmatrix}$, где $a'_{i\bullet} = a_{i\bullet}$ при $i \neq k, k+1$, и $a'_{k\bullet} = a_{k+1\bullet}$, $a'_{k+1\bullet} = a_{k\bullet}$.

Запишем разложение $|A|$ по первому столбцу, выделив отдельно k-е и $(k+1)$-е слагаемые: $|A| = \sum\limits_{i\neq k,k+1}(-1)^{i+1}a_{i1}|A_{i1}| + (-1)^{k+1}a_{k1}|A_{k1}| + (-1)^{k+2}a_{k+1\,1}|A_{k+1\,1}|$. При $i \neq k, k+1$ матрицы A_{i1} и A'_{i1} получаются друг из друга перестановкой двух соседних строк, значит, по предположению индукции $|A_{i1}| = -|A'_{i1}|$. Учитывая, что $A_{k1} = A'_{k+1\,1}$, $A_{k+1\,1} = A'_{k1}$, $a_{k1} = a'_{k+1\,1}$, $a_{k+1\,1} = a'_{k1}$ и $a_{i1} = a'_{i1}$ при $i \neq k, k+1$, имеем $|A| = \sum\limits_{i\neq k,k+1}(-1)^{i+1}a'_{i1}(-|A'_{i1}|) + (-1)^{k+1}a'_{k+1\,1}|A'_{k+1\,1}| + (-1)^{k+2}a'_{k\,1}|A'_{k\,1}| = -\sum\limits_{i\neq k,k+1}(-1)^{i+1}a'_{i1}|A'_{i1}| - (-1)^{k+2}a'_{k+1\,1}|A'_{k+1\,1}| - (-1)^{k+1}a'_{k\,1}|A'_{k\,1}| = -|A'|$. □

Предложение 6.1. *Пусть матрица A' получается из матрицы A применением элементарного преобразования строк II типа. Тогда $|A'| = -|A|$.*

▷ Положим $A' = \widetilde{P}_{kl}(A)$. Пусть для определенности $k < l$. Тогда $\widetilde{P}_{kl}$ можно получить, последовательно выполняя $\widetilde{P}_{k\,k+1}$, $\widetilde{P}_{k+1\,k+2}$, … …, $\widetilde{P}_{l-2\,l-1}$, $\widetilde{P}_{l-1\,l}$, $\widetilde{P}_{l-2\,l-1}$, …, $\widetilde{P}_{k+1\,k+2}$, $\widetilde{P}_{k\,k+1}$. Таким образом, $\widetilde{P}_{kl}$ можно осуществить за $(2(l-k)-1)$ обменов соседних строк. Но из леммы следует, что в результате нечетного числа обменов соседних строк $|A|$ поменяется на $-|A|$. □

Следствие 1. *Определитель матрицы, содержащей две равные строки, равен 0.*[47]

Следствие 2. $|P_{ij}| = -1$.

Предложение 6.2. *Пусть матрица A' получается из матрицы A применением элементарного преобразования строк III типа. Тогда $|A'| = |A|$.*

▷ Пусть матрица A' получается из матрицы $A = \begin{pmatrix} a_{1\bullet} \\ \vdots \\ a_{n\bullet} \end{pmatrix}$ преоб-

[47] Это утверждение перестает быть следствием для определителя с элементами из поля характеристики 2.

разованием $\widetilde{T}_{kl}(\lambda)$, то есть $A' = \begin{pmatrix} a'_{1\bullet} \\ \vdots \\ a'_{n\bullet} \end{pmatrix}$, где $a'_{i\bullet} = a_{i\bullet}$ при $i \neq k$, и $a'_{k\bullet} = a_{k\bullet} + \lambda a_{l\bullet}$. Рассмотрим матрицу $A'' = \begin{pmatrix} a''_{1\bullet} \\ \vdots \\ a''_{n\bullet} \end{pmatrix}$, которая получается из A заменой k-й строки на l-ю, то есть $a''_{i\bullet} = a_{i\bullet}$ при $i \neq k$, и $a''_{k\bullet} = a_{l\bullet}$. По теореме 6.2 имеем $|A'| = |A| + \lambda|A''|$, но по следствию 1 из предложения 6.1, $|A''| = 0$. $\square$

Следствие. $|T_{ij}(\lambda)| = 1$.

Теорема 6.3 (критерий невырожденности-2). *Для матрицы $A \in M_{n\times n}$ эквивалентны следующие условия: A невырожденная (= обратимая) $\Leftrightarrow$ $|A| \neq 0$.*

$\triangleright$ Как мы показали в предыдущих утверждениях, условие $|A| = 0$ не изменяется при элементарных преобразованиях строк. Если A невырожденная, то по теореме 4.3 A приводится элементарными преобразованиями строк к единичной матрице, поэтому $|A| \neq 0$. Если же A вырождена, то она приводится элементарными преобразованиями строк к некоторой матрице A' упрощенного вида, содержащей строку нулей. Поэтому (см. следствие 2 из теоремы 6.2) $|A'| = 0 \Rightarrow |A| = 0$. $\square$

Следующая теорема — об определителе произведения матриц.[48] Вначале докажем лемму (являющуюся частным случаем теоремы).

Лемма. *Пусть $S, A \in \mathbf{M}_{n\times n}$, причем S — элементарная матрица. Тогда $|SA| = |S| \cdot |A|$.*

$\triangleright$ Как нам известно (предложение 4.2), SA — это матрица, полученная из A соответствующим элементарным преобразованием

[48] Эта теорема имеет следующую трактовку. Рассмотрим линейные преобразования $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ и $\psi : \mathbb{R}^n \to \mathbb{R}^n$, заданные (в некотором базисе) матрицами A и B. Тогда $|A|$ — коэффициент изменения ориентированного объема при выполнении φ, $|B|$ — соответствующий коэффициент для ψ, $|AB|$ — соответствующий коэффициент для $\varphi\psi$.

строк. Согласно теореме 6.2 и предложениям 6.1, 6.2, имеем: если $S = P_{ij}$, то $|S| = -1$, $|SA| = |\widetilde{P}_{ij}(A)| = -|A|$; если $S = D_i(\lambda)$, то $|S| = \lambda$, $|SA| = |\widetilde{D}_i(\lambda)(A)| = \lambda|A|$; если $S = T_{ij}(\lambda)$, то $|S| = 1$, $|SA| = |\widetilde{T}_{ij}(\lambda)(A)| = |A|$. Как видим, во всех трех случаях равенство $|SA| = |S| \cdot |A|$ выполнено. $\square$

Теорема 6.4 (произведение определителей). $\forall\ A, B \in \mathbf{M}_{n\times n}$ *выполнено* $|AB| = |A| \cdot |B|$.

$\triangleright$ 1) Если $|A| = 0$, то по теореме 6.3 $\operatorname{rg} A < n \Rightarrow$ (по предложению 3.3) $\operatorname{rg}(AB) < n \Rightarrow$ (по теореме 6.3) $|AB| = 0$.

2) Если $|A| \neq 0$, то A можно представить в виде произведения нескольких элементарных матриц (теорема 4.3): $A = S_1 S_2 \dots S_k$. Тогда из леммы следует, что $|A| = |S_1| \cdot |S_2 \dots S_k| = |S_1| \cdot |S_2| \cdot |S_3 \dots \dots S_k| = \dots = |S_1| \cdot |S_2| \cdot |S_3| \cdot \dots \cdot |S_k|$. Также $|AB| = |S_1 S_2 \dots S_k B| = |S_1| \cdot |S_2 \dots S_k B| = |S_1| \cdot |S_2| \cdot |S_3 \dots S_k B| = \dots = |S_1| \cdot |S_2| \cdot |S_3| \cdot \dots \cdot |S_k| \cdot |B| = |A| \cdot |B|$. $\square$

Следствие. *Если A — невырожденная матрица порядка n, то* $|A^{-1}| = |A|^{-1}$.

$\triangleright$ $AA^{-1} = E \Rightarrow |A||A^{-1}| = 1$. $\square$

Теорема 6.5 (транспонирование определителя). $\forall\ A \in \mathbf{M}_{n\times n}$ *выполнено* $|A^T| = |A|$.

$\triangleright$ 1) Если $|A| = 0$, то $\operatorname{rg} A < n \Rightarrow \operatorname{rg}(A^T) < n \Rightarrow |A^T| = 0$.

2) Нетрудно видеть (см. предложение 4.1), что равенство $|S^T| = |S|$ верно для элементарной матрицы S. Если $|A| \neq 0$, то A можно представить в виде произведения нескольких элементарных матриц: $A = S_1 S_2 \dots S_k$. Тогда $|A| = |S_1 S_2 \dots S_k| = |S_1| \cdot |S_2| \dots |S_k|$; $|A^T| = |S_k^T S_{k-1}^T \dots S_1^T| = |S_k^T| \cdot |S_{k-1}^T| \dots |S_1^T| = |S_k| \cdot |S_{k-1}| \cdot \dots \cdot |S_1|$. $\square$

Следствие. *Определитель нижнетреугольной матрицы равен произведению диагональных элементов.*

$\triangleright$ Следует из теоремы 6.1. $\square$

Предыдущая теорема устанавливает "равноправие"между строками и столбцами, что позволяет во многих доказанных утверждениях строки заменить на столбцы. Объединим полученные результаты в следующую теорему.

Теорема 6.6. *При выполнении элементарного преобразования строк и столбцов*
I типа — определитель меняет знак;
II типа (умножение строки или столбца на λ) — определитель умножается на λ;
III типа — определитель не изменяется.

$\triangleright$ Следует из теорем 6.5, 6.2 и предложений 6.1, 6.2. $\square$

Формула, аналогичная (4), может быть записана и для любого столбца или строки.

Теорема 6.7 (разложение по любому столбцу или строке)**.** *Для матрицы $A = (a_{ij}) \in \mathbf{M}_{n\times n}$ выполнены равенства:*
1) $|A| = \sum\limits_{i=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}|$ *для каждого $j = 1, 2, \ldots, n$;*
2) $|A| = \sum\limits_{j=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}|$ *для каждого $i = 1, 2, \ldots, n$.*

$\triangleright$ 1) Применим $j-1$ элементарных преобразований столбцов I типа, последовательно поменяв столбцы с номерами j и $j-1$, $j-1$ и $j-2$, и т. д., 2 и 1. Таким образом, j-й столбец переместился на место 1-го; при этом определитель умножился на $(-1)^{j-1}$. Теперь нужное равенство следует из формулы (4).

2) Воспользуемся теоремой 6.5. Транспонируем матрицу (при этом соответствующие дополнительные подматрицы тоже транспонируются) и применим равенство, доказанное в первом пункте. $\square$

Явное разложение определителя

Пусть $(i_1, i_2, \ldots, i_n)$ — *перестановка* чисел $a_1 < a_2 < \ldots < a_n$ (т. е. числа $a_1, a_2, \ldots, a_n$, записанные в некотором порядке). Через $S(a_1, a_2, \ldots, a_n)$ обозначим множество всех перестановок чисел $a_1, a_2, \ldots, a_n$.

Будем говорить, что пара i_k, i_l, где $1 \leqslant k < l \leqslant n$, является *инверсией*, если $i_k > i_l$. Таким образом, каждой перестановке $(i_1, i_2, \ldots, i_n)$ соответствует некоторое число инверсий; обозначим его $N(i_1, i_2, \ldots, i_n)$. Если $N(i_1, i_2, \ldots, i_n)$ — четное число, то перестановка $(i_1, i_2, \ldots, i_n)$ называется *четной*, в противном случае — *нечетной*.

Теорема 6.8. *Для матрицы* $A = (a_{ij}) \in \mathbf{M}_{n\times n}$ *выполнено равенство*

$$|A| = \sum_{\substack{(i_1,i_2,\ldots,i_n)\in\\ S(1,2,\ldots,n)}} (-1)^{N(i_1,i_2,\ldots,i_n)} a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}. \tag{5}$$

$\triangleright$ Применим индукцию по n. База индукции тривиальна.

Перестановки $(i_1, i_2, \ldots, i_n)$ из $S(1, 2, \ldots, n)$, для которых i_1 равно фиксированному $i \in \{1, 2, \ldots, n\}$, имеют вид $(i, i_2, \ldots, i_n)$, где $(i_2, \ldots$ $\ldots, i_n)$ — перестановка из $S(1, 2, \ldots, i-1, i+1, \ldots, n)$. Заметим, что $N(i, i_2, \ldots, i_n) = N(i_2, \ldots, i_n) + i - 1$, так как $i_1 = i$ входит ровно в $i-1$ инверсий. Отсюда сумма в правой части (5) равна

$$\sum_{i=1}^{n} \sum_{\substack{(i,i_2,\ldots,i_n)\in\\ S(1,2,\ldots,n)}} (-1)^{N(i,i_2,\ldots,i_n)} a_{i1} a_{i_2 2} \ldots a_{i_n n} = \sum_{i=1}^{n} a_{i1} S_{i1}, \text{ где}$$

$$S_{i1} = \sum_{\substack{(i_2,\ldots,i_n)\in\\ S(1,\ldots,i-1,i+1,\ldots,n)}} (-1)^{N(i_2,\ldots,i_n)+i-1} a_{i_2 2} \ldots a_{i_n n} =$$

$$= (-1)^{i+1} \sum_{\substack{(i_2,\ldots,i_n)\in\\ S(1,2,\ldots,i-1,i+1,\ldots,n)}} (-1)^{N(i_2,\ldots,i_n)} a_{i_2 2} \ldots a_{i_n n}.$$

Как видим, сумма S_{i1} равна (по предположению индукции) $(-1)^{i+1}|A_{i1}|$ (в подматрице A_{i1} строки нумеруются числами $1, 2, \ldots$ $\ldots, i-1, i+1, \ldots, n$, а столбцы — числами $2, 3, \ldots, n$), поэтому (5) следует из (4). $\square$

Формулы с использованием определителя

Для матрицы $A \in M_{n\times n}$ и столбца $t = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}$ через $A_k(t)$ обозначим матрицу, полученную из A заменой k-го столбца на столбец t.

Заметим, что по теореме 6.7 сумма $\sum\limits_{i=1}^{n}(-1)^{i+k}t_i|A_{ik}|$ равна определителю $|A_k(t)|$.

Предложение 6.3 (правило Крамера). *Пусть $A \in \mathbf{M}_{n\times n}$ — невырожденная матрица, и столбец $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ является (единственным) решением системы линейных уравнений $AX = b$. Тогда $x_k = \dfrac{|A_k(b)|}{|A|}$ (для $k = 1, 2, \ldots, n$).*[49]

$\triangleright$ Домножим i-е уравнение системы на $(-1)^{i+k}|A_{ik}|$, и сложим все получившиеся уравнения. Получим уравнение $\sum\limits_{j=1}^{n} c_j x_j = d$, где $c_j = \sum\limits_{i=1}^{n}(-1)^{i+k}|A_{ik}|a_{ij} = |A_k(a_{\bullet j})|$, $d = \sum\limits_{i=1}^{n}(-1)^{i+k}|A_{ik}|b_i = |A_k(b)|$. При $j \neq k$ матрица $A_k(a_{\bullet j})$ имеет два равных столбца (j-й и k-й), поэтому $c_j = |A_k(a_{\bullet j})| = 0$. Поскольку $A_k(a_{\bullet k}) = A$, имеем $|A|x_k = |A_k(b)|$. $\square$

Предложение 6.4 (формула обратной матрицы). *Пусть $A = (a_{ij}) \in \mathbf{M}_{n\times n}$ — невырожденная матрица, $A^{-1} = (x_{ij})$. Тогда $x_{ij} = \dfrac{(-1)^{i+j}|A_{ji}|}{|A|}$ (для $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$).*

$\triangleright$ Из равенства $AA^{-1} = E$ следует, что столбец $x_{\bullet j}$ обратной матрицы является решением системы $Ax_{\bullet j} = e_j$, где в столбце e_j все элементы, за исключением единицы на j-м месте — нули. По правилу Крамера $x_{ij} = \dfrac{|A_i(e_j)|}{|A|}$. Раскладывая $|A_i(e_j)|$ по i-му столбцу, получаем, что $|A_i(e_j)| = (-1)^{i+j}|A_{ji}|$. $\square$

[49] Правило Крамера показывает, чему равны коэффициенты в разложении по вектора по базису в терминах ориентированных объемов.

Задачи и упражнения

1. Вычислите определитель матрицы порядка n:

а) $\begin{vmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 1 & 2 & \dots & 2 \\ 2 & 2 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & 2 & \dots & 1 \end{vmatrix}$; б) $\begin{vmatrix} 2 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ 0 & 1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 2 & 1 \\ 0 & 0 & \dots & 1 & 2 \end{vmatrix}$.

2. Даны матрицы $A \in M_{m\times m}$, $B \in M_{n\times n}$, $C \in M_{m\times n}$. Докажите, что определитель блочной матрицы с "углом нулей" $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ равен $|A| \cdot |B|$.

3. Как изменится определитель матрицы $A \in \mathbf{M}_{n\times n}$, если а) переставить ее строки в обратном порядке; б) "отразить"матрицу симметрично относительно ее центра; в) "повернуть"матрицу на 90° относительно центра; г) отразить симметрично относительно побочной диагонали?

4. а) Выразите $|-A|$ через $|A|$ (в зависимости от порядка n матрицы A). б) Докажите, что кососимметрическая матрица нечетного порядка всегда вырожденная.

5. Докажите, что (при $n \geqslant 2$) в явном разложении определителя количества слагаемых со знаком "$+$"и со знаком "$-$"равны.

6. Пусть все элементы квадратной матрица A — целые числа, причем все элементы на главной диагонали нечетные, а вне главной диагонали четные. Докажите, что матрица A невырожденная.

7. Дана обратимая матрица A, все элементы которой — целые числа. Докажите, что все элементы матрицы A^{-1} — целые числа $\Leftrightarrow |A| = \pm 1$.

Приложение

Логика

В логических рассуждениях иногда для краткости используются следующие знаки.

$\Rightarrow$ — знак следствия (импликации); например, запись $A \Rightarrow B$ означает, что из утверждения A следует утверждение B.

$\Leftrightarrow$ — знак эквивалентности (равносильности), например, запись $A \Leftrightarrow B$ означает, что утверждение A верно тогда и только тогда, когда верно утверждение B.

$\forall$ — квантор всеобщности; он заменяет слова "для любого "при любых"и т.д.

$\exists$ — квантор существования; он заменяет слово "существует".

Иногда мы будем пользоваться *принципом математической индукции*, который заключается в следующем.

Пусть имеется последовательность утверждений T_1, T_2, T_3, $\ldots$, про которую известно, что

1) T_1 верно (база индукции);

2) из того, что T_n верно, вытекает, что T_{n+1} верно (для $n = 1, 2, 3, \ldots$) (переход или шаг индукции).

Тогда вся последовательность состоит из верных утверждений.

Возможна вариация условия 2): из предположения, что утверждения T_1, T_2, $\ldots$, T_n верны, вытекает, что T_{n+1} верно (для $n = 1, 2, 3, \ldots$).

Множества

Множество — это совокупность некоторых объектов, эти объекты называются *элементами* данного множества.

Приняты следующие обозначения:

Включение $a \in A$ означает, что элемент a принадлежит множеству A; $a \notin A$ — элемент a не принадлежит множеству A.

Множества, содержащие хотя бы один элемент, называются *непустыми*; *пустое* множество (то есть множество, не содержащее элементов) обозначается $\varnothing$.

Если в множестве бесконечное число элементов, то множество называется *бесконечным*, в противном случае оно называется *конечным.* Конечное множество может задаваться перечислением своих элементов, скажем, $A = \{a_1, a_2, \ldots, a_n\}$.

Приняты обозначения $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ соответственно для множеств натуральных, целых, рациональных и вещественных (действительных) чисел.

Говорят, что множество B является *подмножеством* множества A (иногда говорят, что B *вложено* в A), если $x \in B \Rightarrow x \in A$. Обозначение: $B \subset A$ или $B \subseteq A$. Заметим, что $\varnothing \subseteq A$ для любого множества A. Множества A и B называются *равными*, если одновременно $B \subseteq A$ и $A \subseteq B$. Подмножество можно выделить некоторым условием. Для подмножества элементов множества A, удовлетворяющих условию $\mathcal{X}$, примем обозначение $\{a \in A \,|\, \mathcal{X}\}$. Например, подмножество четных (целых) чисел можно задать записью $\{n \in \mathbb{Z} \,|\, \exists k \in \mathbb{Z}\, n = 2k\}$ или более коротко $\{2k \,|\, k \in \mathbb{Z}\}$.

Над множествами определяются следующие операции.

Объединение множеств A и B — это множество элементов x, для которых имеет место хотя бы одно из двух включений: $x \in A$, $x \in B$. Обозначение: $A \bigcup B$. Можно определить объединение $\bigcup\limits_{i \in I} A_i$ (i пробегает множество индексов I) как множество элементов, принадлежащих хотя бы одному из множеств A_i, $i \in I$.

Пересечение множеств A и B — это множество элементов x, для которых имеют место оба включения: $x \in A$, $x \in B$. Обозначение: $A \bigcap B$. Иначе говоря, $A \bigcap B = \{x \in A \,|\, x \in B\}$ (или $A \bigcap B = \{x \in B \,|\, x \in A\}$. В случае $A \bigcap B = \varnothing$ говорят, что множества A и B *непересекающиеся.* Можно определить пересечение $\bigcap\limits_{i \in I} A_i$ как множество элементов, принадлежащих каждому из множеств A_i, $i \in I$.

Разностью множеств A и B называется множество $\{x \in A \,|\, x \notin B\}$. Обозначение: $A \setminus B$.

Декартовым произведением множеств A и B называется множество упорядоченных пар $\{(a, b) \,|\, a \in A, b \in B\}$. Обозначение: $A \times B$. Аналогично определяется декартово произведение $A_1 \times A_2 \times \ldots \times$

$\times A_n$ как множество n-ок $\{(a_1, a_2, \ldots, a_n) \,|\, a_1 \in A_1, \ldots, a_n \in A_n\}$. Если $A_1 = A_2 = \ldots = A_n = A$, то декартово произведение $A_1 \times A_2 \times \times \ldots \times A_n$ обозначается также A^n.

Знак суммирования

Сумму $a_1 + a_2 + \ldots + a_n$ коротко обозначают $\sum\limits_{i=1}^{n} a_i$. Сумму чисел a_i, где i пробегает конечное множество индексов I, обозначают $\sum\limits_{i \in I} a_i$. Отметим следующие свойства знака суммирования.

(Линейность знака суммирования) *Для любых чисел* $a_1, a_2, \ldots \ldots, a_n, b_1, b_2, \ldots, b_n$ *выполнены равенства:*
1) $\sum\limits_{i=1}^{n} (a_i + b_i) = \sum\limits_{i=1}^{n} a_i + \sum\limits_{i=1}^{n} b_i$; *2)* $\sum\limits_{i=1}^{n} \lambda a_i = \lambda \sum\limits_{i=1}^{n} a_i$.

(Изменение порядка при двойном суммировании) *Для* mn *действительных чисел* $a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{m1}, \ldots, a_{mn}$ *выполнено равенство* $\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} a_{ij} = \sum\limits_{j=1}^{n} \sum\limits_{i=1}^{m} a_{ij}$.

$\triangleright$ Запишем числа в таблицу (матрицу). Вычислим сумму всех элементов матрицы, просуммировав элементы по столбцам, а затем по строкам. При этом получим левую часть равенства. Просуммировав вначале по строкам, а затем по столбцам, получим правую часть равенства. $\square$

Понятия отображения и преобразования

Пусть X и Y — произвольные множества. Говорят, что задано *отображение* f множества X в множество Y (или отображение из X в Y), если каждому элементу $x \in X$ поставлен в соответствие единственный элемент $y \in Y$. Обозначения для отображения: $f : X \to Y$ или $X \xrightarrow{f} Y$.

Образом подмножества $X' \subset X$ при отображении $f : X \to Y$ называется множество $\{f(x) \,|\, x \in X'\}$. Обозначение: $f(X')$. Образ $f(X)$ всего множества X при отображении $f : X \to Y$ называется также *образом отображения* f. Помимо $f(X)$ для образа отображения используется обозначение $\operatorname{Im} f$.

Отображение $f : X \to X$ множества X в себя называется также *преобразованием* множества X. Преобразование $f : X \to X$ называется *тождественным* преобразованием, если $\forall x \in X \ f(x) = x$. Обозначение: I_X.

Пусть даны некоторые множества X, Y, Z и отображения $f : X \to Y$, $g : Y \to Z$. Отображение $h : X \to Z$, заданное $\forall x \in X$ равенством $h(x) = g(f(x))$, называется *композицией отображений* g и f или *произведением* g на f. Композиция обозначается $g \circ f$ или gf.

Для произведения отображений выполнено следующее свойство ассоциативности.

Пусть даны некоторые множества X, Y, Z, T и отображения $f : X \to Y$, $g : Y \to Z$, $h : Z \to T$. Тогда $h(gf) = (hg)f$.

$\triangleright$ Пользуясь многократно определением композиции отображений, $\forall x \in X$ имеем: $(h(gf))(x) = h((gf)(x)) = h(g(f(x)))$. С другой стороны, $((hg)f)(x) = (hg)(f(x)) = h(g(f(x)))$. $\square$

Отсюда можно получить следующее обобщение:

Пусть даны множества X_1, X_2, $\ldots$, X_{n+1} и отображения $f_i : X_i \to X_{i+1}$, $i = 1, 2, \ldots, n$. Тогда произведение $f_n f_{n-1} \ldots f_1$ определено и не зависит от порядка выполнения композиций (не зависит от расстановки скобок).

$\triangleright$ Используем индукцию по n. База: $n = 3$ (см. предыдущее утверждение). Пусть утверждение верно для композиции $3, 4, \ldots, n-1$ отображений. Пусть в произведении $f_n f_{n-1} \ldots f_1$ Скобки расставлены двум способами: $g = (f_n f_{n-1} \ldots f_{k+1})(f_k f_{k-1} \ldots f_1)$ и $h = (f_n f_{n-1} \ldots f_{l+1})(f_l f_{l-1} \ldots f_1)$ (в каждом из способов мы отметили последнюю операцию композиции, в остальном операции выполняются в произвольном порядке).

Так как $k < n$, то в силу предположения индукции отображение $(f_k f_{k-1} \ldots f_1) : X_1 \to X_{k+1}$ не зависит от порядка выполнения композиции (то есть скобки в этом выражении можно расставлять как угодно, оно от этого не изменится). Такое же замечание справедливо и для отображений $(f_n f_{n-1} \dots f_{k+1}) : X_{k+1} \to X_{n+1}$, $(f_l f_{l-1} \dots f_1) : X_1 \to X_{l+1}$, $(f_n f_{n-1} \dots f_{l+1}) : X_{l+1} \to X_{n+1}$.

Таким образом, если $k = l$, то сразу получаем $g = h$.

Пусть для определенности $k < l$.

Тогда $g = ((f_n f_{n-1} \dots f_{l+1})(f_l f_{l-1} \dots f_{k+1}))(f_k f_{k-1} \dots f_1)$, $h = (f_n f_{n-1} \dots f_{l+1})((f_l f_{l-1} \dots f_{k+1})(f_k f_{k-1} \dots f_1))$. Из предыдущего утверждения, примененного к трем отображениям $(f_k f_{l-1} \dots \dots f_1) : X_1 \to X_{k+1}$, $(f_l f_{l-1} \dots f_{k+1}) : X_{k+1} \to X_{l+1}$, $(f_n f_{n-1} \dots \dots f_{l+1}) : X_{l+1} \to X_{n+1}$, получаем, что $g = h$. □

Доказанное свойство произведения отображений позволяет для любого $n \in \mathbb{N}$ корректно определить n-ю степень преобразования $f : X \to X$ как $f^n = \underbrace{ff \dots f}_{n \text{ букв } f}$. Из предыдущего ясно, что $\forall\, m, n \in \mathbb{N}$ верны равенства $f^{m+n} = f^m f^n$ и $f^{mn} = (f^m)^n$. Для произведения отображений роль "единицы"играет тождественное преобразование.

Отображение $g : Y \to X$ называется *левым (правым) обратным* для отображения $f : X \to Y$, если $gf = I_X$ $(fg = I_Y)$. Не для всякого отображения $f : X \to Y$ найдется левое обратное отображение. Если оно найдется, то будем говорить, что f обратимо слева. Аналогично определяются отображения, обратимые справа. Отображения, являющиеся одновременно обратимыми справа и слева, будем называть *обратимыми.*

Для обратимого отображения $f : X \to Y$ существует единственное левое обратное отображение и единственное правое обратное отображение, причем они равны.

▷ Пусть $g : Y \to X$ — некоторое левое обратное отображение для f, $h : Y \to X$ — некоторое правое обратное отображение для f. Тогда $g = gI_Y = g(fh)$, что равно $(gf)h = I_X h = h$. Итак, каждое левое обратное отображение g совпадает с h, и значит оно единственно. Аналогично, каждое правое обратное отображение h совпадает с g, и, значит, оно единственно. □

Для обратимого отображения $f : X \to Y$ его единственное левое (оно же и единственное правое) обратное отображение будем называть *обратным отображением* и обозначать f^{-1}.

Пусть даны некоторые множества X, Y, Z и обратимые отображения $f : X \to Y$ и $g : Y \to Z$. Тогда

1) f^{-1} обратимо, причем $(f^{-1})^{-1} = f$;

2) gf обратимо, причем $(gf)^{-1} = f^{-1}g^{-1}$.

$\triangleright$ 1) Равенства $ff^{-1} = I_Y$, $f^{-1}f = I_X$ означают, что f — обратное отображение для f^{-1}.

2) Нужное утверждение вытекает из равенств
$(f^{-1}g^{-1})(gf) = f^{-1}(g^{-1}g)f = f^{-1}I_Y f = f^{-1}f = I_X$,
$(gf)(f^{-1}g^{-1}) = g(ff^{-1})g^{-1} = gI_Y g^{-1} = gg^{-1} = I_Z$. $\square$

Следствие:

Пусть даны множества X_1, X_2, $\ldots$, X_{n+1} и обратимые отображения $f_i : X_i \to X_{i+1}$, $i = 1, 2, \ldots, n$. Тогда произведение $f_n f_{n-1} \ldots f_1$ обратимо, причем $(f_n f_{n-1} \ldots f_1)^{-1} = f_1^{-1} f_2^{-1} \ldots f_n^{-1}$.

Если $f : X \to X$ — обратимое преобразование, то можно определить n-ю степень и для неположительных целых n: положим $f^0 = I_X$, $f^{-m} = (f^{-1})^m$, если $m \in \mathbb{N}$. Нетрудно проверить, что для всех $m, n \in \mathbb{Z}$ верны равенства $f^{m+n} = f^m f^n$ и $f^{mn} = (f^m)^n$.

Отображение $f : X \to Y$ называется

1) *инъективным* (или *инъекцией*), если из равенства $f(x) = f(x')$ вытекает, что $x = x'$;

2) *сюръективным* (или *сюръекцией*), если $f(X) = Y$;

3) *биективным* (или *биекцией*, или взаимно-однозначным соответствием), если оно одновременно является и сюръекцией, и инъекцией.

Нетрудно установить, что

1) $f : X \to Y$ инъективно $\Leftrightarrow$ f обратимо слева;

2) $f : X \to Y$ сюръективно $\Leftrightarrow$ f обратимо справа;

3) $f : X \to Y$ биективно $\Leftrightarrow$ f обратимо.

Ответы, указания и решения

§ 1

1. Ответ: $\begin{pmatrix} -5 & 4 & 3 \\ 11 & -23 & 20 \end{pmatrix}$.

2. Например, $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $A_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.[50] Тогда произвольная матрица $\begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix}$ равна $\sum\limits_{i=1}^{4} \lambda_i A_i$.

3. Нетрудно видеть,[51] что строка (x_1, x_2, x_3, x_4) с условием $x_1 + x_2 + x_3 + x_4 = 0$ равна $x_2(-1, 1, 0, 0) + x_3(-1, 0, 1, 0) + x_4(-1, 0, 0, 1)$.

4. Линейная комбинация линейных комбинаций столбцов из $\mathcal{A}$ равна линейной комбинации столбцов из $\mathcal{A}$.[52]

5. Можно использовать разложение $A = \dfrac{A + A^T}{2} + \dfrac{A - A^T}{2}$.[53]

§ 2

[50] Это стандартный базис в $\mathbf{M}_{2\times 2}$.

[51] Вообще, эта задача — частный случай задачи решения однородной системы линейных уравнений — см. § 5.

[52] Переформулировка этой задачи: линейная оболочка всегда является векторным пространством.

[53] Переформулировка этой задачи: векторное пространство $\mathbf{M}_{n\times n}$ раскладывается в прямую сумму подпространств $\mathbf{M}^{+}_{n\times n}$ и $\mathbf{M}^{-}_{n\times n}$ симметрических и кососимметрических матриц.

1. Пусть A_1, A_2, ..., A_k — линейно зависимая система. Тогда прибавив к разложению B по векторам A_1, A_2, ..., A_k нетривиальную линейную комбинацию, равную O, получим другое разложение B.

2. Ответ: нет. Достаточно применить теорему об оценке ранга суммы для матриц $C = A + B$ и $D = -B$.

3. Найдутся r фиксированных столбцов a_1, ..., a_r, по которым раскладываются все столбцы данной матрицы A. Пользуясь этими разложениями, можно получить разложение $A = \sum\limits_{i=1}^{r} A_i$, где все столбцы матрицы A_i пропорциональны a_i.

4. Ответ: $\operatorname{rg} \mathbf{M}^+_{n\times n} = \dfrac{n(n+1)}{2}$, $\operatorname{rg} \mathbf{M}^-_{n\times n} = \dfrac{n(n-1)}{2}$.

   Достаточно указать базисные подсистемы в $\mathbf{M}^+_{n\times n}$ и $\mathbf{M}^-_{n\times n}$. Пусть $E(k,l)$ — матрица, в которой элемент на пересечении k-й строки и l-го столбца равен 1, а остальные равны 0. Тогда матрицы $E(i,j) + E_(j,i)$, $1 \leqslant i \leqslant j \leqslant n$, образуют базисную подсистему в $\mathbf{M}^+_{n\times n}$, а матрицы $E(i,j) - E_(j,i)$, $1 \leqslant i < j \leqslant n$, образуют базисную подсистему в $\mathbf{M}^-_{n\times n}$. Можно решить задачу и в духе § 5, задавая $\mathbf{M}^+_{n\times n}$ и $\mathbf{M}^-_{n\times n}$ системами линейных уравнений.

5. Если линейно независимая подсистема $A_1, \ldots, A_k$ не является базисной для $\mathcal{A}$, то найдется $A_{k+1} \in \mathcal{A} \setminus \langle A_1, \ldots, A_k \rangle$. Тогда можно "нарастить" линейно независимую подсистему до $A_1, \ldots, A_k, A_{k+1}$. Далее можно продолжать процедуру.

§ 3

1. Ответ: $\begin{pmatrix} -13 & 4 \\ 4 & -10 \end{pmatrix}$.

2. Воспользуйтесь равенством $(SAS^{-1})^n = SA^nS^{-1}$.

3. Для матриц $A = (a_{ij})$ и $B = (b_{ji})$ каждое из выражений $\mathrm{tr}(AB)$ $\mathrm{tr}(BA)$ равно двойной сумме $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} a_{ij}b_{ji}$.

4. Воспользуйтесь равенством $(AB)^T = B^T A^T$.

5. а) В качестве матрицы X возьмите матрицы вида $E(i,j)$ (на пересечении i-й строки и j-го столбца единица, а остальные элементы равны 0).

   б) Используйте то, что при умножении на $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ слева i-я строка умножается на λ_i, а при умножении справа — j-й столбец умножается на λ_j.

6. Используйте существование многочлена, который принимает заданные значения в данных n точках.

7. а) и б) следуют из правила умножения матриц.

   в) Ответ: $\begin{pmatrix} 1 & 10 & 45 & 120 \\ 0 & 1 & 10 & 45 \\ 0 & 0 & 1 & 10 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

   Можно воспользоваться формулой бинома Ньютона (она применима, поскольку матрицы A и E перестановочны) и равенством $A^4 = O$.

8. Проверьте, что обратной матрицей является матрица $\sum\limits_{k=0}^{m-1} A^k$.[54]

9. Равенство $BA = E_n$ противоречит предложению 3.3 об оценке ранга произведения матриц.

§ 4

[54] Вообще, если квадратная матрица A такова, что ряд из матриц $\sum\limits_{k=0}^{\infty} A^k$, сходится, сумма этого ряда является обратной матрицей для матрицы $E - A$. Здесь можно провести параллель с разложением по Тейлору функции $(1-x)^{-1}$.

1. $\begin{pmatrix} \vdots \\ a \\ \vdots \\ b \\ \vdots \end{pmatrix} \xrightarrow{III} \begin{pmatrix} \vdots \\ a+b \\ \vdots \\ b \\ \vdots \end{pmatrix} \xrightarrow{III} \begin{pmatrix} \vdots \\ a+b \\ \vdots \\ -a \\ \vdots \end{pmatrix} \xrightarrow{III} \begin{pmatrix} \vdots \\ b \\ \vdots \\ -a \\ \vdots \end{pmatrix} \xrightarrow{II} \begin{pmatrix} \vdots \\ b \\ \vdots \\ a \\ \vdots \end{pmatrix}.$

2. Выполним цепочку элементарных преобразований строк, приводящую к ступенчатому виду, например:
$\begin{pmatrix} 0 & 0 & 2 & 1 & -1 & 2 \\ 0 & 0 & -4 & -2 & 2 & -4 \\ 3 & -6 & 5 & 1 & 2 & 5 \\ -1 & 2 & 5 & 3 & -7 & 2 \end{pmatrix} \xrightarrow{\widetilde{T}_{21}(2)} \xrightarrow{\widetilde{T}_{34}(3)} \begin{pmatrix} 0 & 0 & 2 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 20 & 10 & -19 & 11 \\ -1 & 2 & 5 & 3 & -7 & 2 \end{pmatrix} \xrightarrow{\widetilde{P}_{12}} \xrightarrow{\widetilde{P}_{14}} \xrightarrow{\widetilde{D}_1(-1)}$
$\begin{pmatrix} 1 & -2 & -5 & -3 & 7 & -2 \\ 0 & 0 & 2 & 1 & -1 & 2 \\ 0 & 0 & 20 & 10 & -19 & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\widetilde{T}_{32}(-10)} \begin{pmatrix} 1 & -2 & -5 & -3 & 7 & -2 \\ 0 & 0 & 2 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & -9 & -9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$
Теперь ясно, что $\operatorname{rg} A = 3$, и в системе столбцов матрицы A одна из базисных подсистем — столбцы с номерами 1, 3, 5. По теореме о базисном миноре, можно отыскать невырожденную подматрицу порядка 3 даже в матрице $B = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -4 & 2 \\ 3 & 5 & 2 \\ -1 & 5 & -7 \end{pmatrix}$, образованной 1-м, 3-м и 5-м столбцами матрицы A. Остается найти некоторую базисную подсистему строк матрицы B. Для удобства, можно привести к ступенчатому виду матрицу B^T элементарными преобразованиями строк:
$\begin{pmatrix} 0 & 0 & 3 & -1 \\ 2 & -4 & 5 & 5 \\ -1 & 2 & 2 & -7 \end{pmatrix} \xrightarrow{\widetilde{T}_{23}(2)} \xrightarrow{\widetilde{P}_{13}} \begin{pmatrix} -1 & 2 & 2 & -7 \\ 0 & 0 & 9 & -9 \\ 0 & 0 & 3 & -1 \end{pmatrix} \xrightarrow{\widetilde{D}(1/9)} \xrightarrow{\widetilde{T}_{32}(-3)} \begin{pmatrix} -1 & 2 & 2 & -7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$
В системе столбцов матрицы B^T одна из базисных подсистем — столбцы с номерами 1, 3, 4. Значит, в матрице B, а следовательно и в матрице A, одна из базисных подсистем строк — строки с номерами 1, 3, 4.

   Таким образом, можем выделить одну из возможных невырожденных подматриц порядка 3 в матрице A:
$$\begin{pmatrix} \mathbf{0} & 0 & \mathbf{2} & 1 & \mathbf{-1} & 2 \\ 0 & 0 & -4 & -2 & 2 & -4 \\ \mathbf{3} & -6 & \mathbf{5} & 1 & \mathbf{2} & 5 \\ \mathbf{-1} & 2 & \mathbf{5} & 3 & \mathbf{-7} & 2 \end{pmatrix}.$$

3. Используйте алгоритм отыскания обратной матрицы.

4. а) Ответ: $\operatorname{rg} A + \operatorname{rg} B$.

Используя "блочные"элементарные преобразования строк (одно блочное элементарное преобразование строк равносильно m обычным), приведите матрицу к виду $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$.

б) Согласно предложению 3.2, элементарными преобразованиями строк данная матрица приводится к виду $\begin{pmatrix} A & E \\ O & O \end{pmatrix}$.

5. Для доказательства равенства $\mathrm{rg}(AB) = \mathrm{rg}\, B$ достаточно доказать неравенство $\mathrm{rg}\, AB \geqslant \mathrm{rg}\, B$ (обратное неравенство выполнено всегда). Можно выделить в матрице A "базисный минор"K порядка m, тогда в матрице AB есть подматрица KB ранга $\mathrm{rg}\, B$, следовательно $\mathrm{rg}\, AB \geqslant \mathrm{rg}\, B$.

6. В качестве столбцов матрицы B можно взять базисную подсистему столбцов матрицы A, и подобрать соответствующую матрицу C.

7. а) Можно использовать соображение из начала пункта "Элементарные преобразования и ранг": равенство нулю линейной комбинации столбцов с фиксированными коэффициентами сохраняется в процессе элементарных преобразований строк.

   б) Используйте пункт а), предложение 4.2 и тот факт, что обратимая матрица представляется в виде произведения нескольких элементарных матриц.

8. Матричное уравнение $XA = B$ эквивалентно уравнению $A^T X^T = B^T$.

§ 5

1. Запишем расширенную матрицу системы $\left(\begin{array}{ccccc|c} 0 & 0 & 2 & 1 & -1 & 2 \\ 0 & 0 & -4 & -2 & 2 & -4 \\ 3 & -6 & 5 & 1 & 2 & 5 \\ -1 & 2 & 5 & 3 & -7 & 2 \end{array}\right)$. Выполнив прямой ход метода Гаусса, получим (см. решение задачи 1 из § 4) $\left(\begin{array}{ccccc|c} 1 & -2 & -5 & -3 & 7 & -2 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right)$.

Далее, выполняем обратный ход метода Гаусса:
$\begin{pmatrix} 1 & -2 & -5 & -3 & 7 & | & -2 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & | & 1 \\ 0 & 0 & 0 & 0 & 1 & | & 1 \end{pmatrix} \xrightarrow{\widetilde{T}_{23}(-\frac{1}{2})} \xrightarrow{\widetilde{T}_{13}(-7)} \begin{pmatrix} 1 & -2 & -5 & -3 & 0 & | & -9 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & | & \frac{3}{2} \\ 0 & 0 & 0 & 0 & 1 & | & 1 \end{pmatrix} \xrightarrow{\widetilde{T}_{12}(5)}$
$\begin{pmatrix} 1 & -2 & 0 & -\frac{1}{2} & 0 & | & -\frac{3}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 0 & | & 3/2 \\ 0 & 0 & 0 & 0 & 1 & | & 1 \end{pmatrix}$.

Введем новый порядок неизвестных: x_1, x_3, x_5, x_2, x_4. В новом порядке матрица системы имеет вид: $\begin{pmatrix} 1 & 0 & 0 & -2 & -\frac{1}{2} & | & -\frac{3}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & | & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 0 & | & 1 \end{pmatrix}$, и система имеет решение: $\begin{pmatrix} x_1 \\ x_3 \\ x_5 \\ x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ \frac{3}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 & \frac{1}{2} \\ 0 & -\frac{1}{2} \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$.

Возвращаясь к изначальному порядку неизвестных, получаем ответ: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ 0 \\ \frac{3}{2} \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 & \frac{1}{2} \\ 1 & 0 \\ 0 & -\frac{1}{2} \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$.

2. Запишем данные столбцы в матрицу Φ и найдем фундаментальную матрицу Ψ систему $(\Phi^T \,|\, O)$: $\Phi^T = \begin{pmatrix} 0 & -3 & 2 & 1 \\ 2 & 1 & 0 & -1 \\ 4 & -1 & 2 & -1 \end{pmatrix} \xrightarrow{\widetilde{P}_{12}}$ $\xrightarrow{\widetilde{T}_{31}(2)} \begin{pmatrix} 2 & 1 & 0 & -1 \\ 0 & -3 & 2 & 1 \\ 0 & -3 & 2 & 1 \end{pmatrix} \xrightarrow{\widetilde{T}_{32}(-1)} \xrightarrow{\widetilde{D}_2(-\frac{1}{3})} \xrightarrow{\widetilde{T}_{12}(-1)} \xrightarrow{\widetilde{D}_1(\frac{1}{2})} \begin{pmatrix} 1 & 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & -\frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Отсюда $\Psi = \begin{pmatrix} -\frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$, и система $(\Psi^T \,|\, O) = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 1 & 0 & | & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 1 & | & 0 \end{pmatrix}$ или $\begin{pmatrix} -1 & 2 & 3 & 0 & | & 0 \\ 1 & 1 & 0 & 3 & | & 0 \end{pmatrix}$ имеет заданное решение.

3. Домножая равенство $AXB = C$ на A^{-1} слева и на B^{-1} справа, находим $X = A^{-1}CB^{-1}$. Легко проверить, что найденная матрица удовлетворяет данному уравнению.

4. Указание: столбцы любой фундаментальной матрицы Φ' — линейные комбинации столбцов Φ.[55]

[55] Матрицу S можно интерпретировать как матрицу перехода от базиса к базису в векторном пространстве $\mathrm{Sol}(A \,|\, O)$.

5. Если каждое уравнение системы $(A' \mid b')$ является линейной комбинацией уравнений системы $(A \mid b)$, то из $AX = b$ следует $A'X = b'$. Аналогично, из $A'X = b'$ следует $AX = b$. Таким образом, системы равносильны. Наоборот, пусть две системы $(A \mid b)$ и $(A' \mid b')$ имеют одно и то же непустое множество решений $X_0 + \mathrm{Sol}(A \mid O)$. Тогда при добавлении к строкам матрицы $(A \mid b)$ строк матрицы $(A' \mid b')$ ранг не изменится (он равен $n - \mathrm{rg}(\mathrm{Sol}(A \mid O))$. Далее можно воспользоваться основной теоремой о рангах.

6. Одно из возможных решений основано на следующем соображении: условие линейной независимости системы строк матрицы Φ с номерами $j_1, j_2, \ldots, j_s$ означает, что неизвестные x_{j_1}, …, x_{j_s} могут (независимо друг от друга) принимать любые значения, то есть их можно взять за свободные неизвестные.

§ 6

1. а) Ответ: $(-1)^{n-1}(2n-1)$.

   Прибавим к первой строке все остальные, получим $(2n-1)\begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ 2 & 1 & 2 & \ldots & 2 \\ 2 & 2 & 1 & \ldots & 2 \\ & \ldots & \ldots & \ldots & \\ 2 & 2 & \ldots & 1 & 2 \\ 2 & 2 & \ldots & 2 & 1 \end{vmatrix}$. В последнем определителе теперь из каждой строки (кроме первой) вычтем удвоенную первую. Получим верхнетреугольную матрицу с числами $1, -1, -1, \ldots, -1$ по диагонали.

   б) Ответ: $n+1$.

   Обозначая данный определитель через K_n, из разложения по первому столбцу можно получить реккурентную формулу $K_{n+1} = 2K_n - K_{n-1}$, из которой следует, что последовательность $\{K_n\}$ — арифметическая прогрессия.

2. Можно привести данную матрицу к верхнетреугольному виду, используя только элементарные преобразования последних n строк и первых m столбцов.

3. а) Переставить строки матрицы в обратном порядке можно последовательно меняя местами первую строку с n-й, вторую — с $(n-1)$-й, и т.д. — всего $\left[\frac{n}{2}\right]$ элементарных преобразований строк I типа. Таким образом, определитель умножится на $(-1)^{\left[\frac{n}{2}\right]}$.

   б) Чтобы отразить матрицу симметрично относительно ее центра, можно сначала переставить строки в обратном порядке (из а) следует, что при этом определитель умножится на $(-1)^{\left[\frac{n}{2}\right]}$), а потом переставить столбцы в обратном порядке (при этом определитель еще раз умножится на $(-1)^{\left[\frac{n}{2}\right]}$). Таким образом, определитель не изменится.

   в) "Поворот"матрицы на 90° можно получить, выполнив последовательно преобразование из пункта а) и транспонирование.

   г) Отражение относительно главной диагонали можно получить, последовательно выполнив преобразования из пунктов в) и а).

4. а) Ответ: $|-A| = (-1)^n|A|$.

   б) Следует из а) и теоремы 6.5.

5. Можно доказать индукцией по n, пользуясь разложением по первому столбцу.

6. В явном разложении определителя $|A|$ будет одно нечетное слагаемое, а все остальные четные.

7. Заметим, что определитель матрицы с целыми числами равен целому числу. В одну сторону нужное утверждение дает следствие из теоремы 6.4. Чтобы доказать обратное утверждение, можно воспользоваться формулой обратной матрицы (см. предложение 6.4).

Литература

Учебники

1. Беклемишев Д. В. Курс аналитической геометрии и линейной алгебры. — 11-е изд., испр. — М.: Физматлит, 2006.

2. Босс. В. Лекции по математике: линейная алгебра. — М.: КомКнига, 2005.

3. Винберг Э. Б. Курс алгебры. — М.: “Факториал Пресс”, 2002.

4. Кострикин А. И. Введение в алгебру. Ч. I. Основы алгебры. Ч. II. Линейная алгебра — М.: МЦНМО, 2009.

5. Умнов А. Е. Аналитическая геометрия и линейная алгебра. — М.О.: Издание ЗАО “Оптимизационные системы и технологии”, 2004.

6. Чехлов В. И. Лекции по аналитической геометрии и линейной алгебре. — М.: МФТИ, 2000.

Задачники

1. Беклемишева Л. А., Беклемишев Д. В., Петрович А. Ю., Чубаров И. А. Сборник задач по аналитической геометрии и линейной алгебре. — 3-е изд. — СПб.: Лань, 2008.

2. Сборник задач по алгебре / под. ред. А. И. Кострикина. — М.: Физматлит, 2005.

3. Прасолов В. В. Задачи и теоремы линейной алгебры. — М.: Наука. Физматлит, 1996.

4. Сборник задач по аналитической геометрии и линейной алгебре / под. ред. Ю. М. Смирнова. — Изд. 2-е, перераб. и доп. — М.: Логос, 2005.

Список используемых обозначений

$\Rightarrow$	знак следствия
$\Leftrightarrow$	знак эквивалентности (равносильности)
$\forall$	квантор всеобщности
$\exists$	квантор существования
$a \in A$	a принадлежит множеству A
$a \notin A$	a не принадлежит множеству A
$\varnothing$	пустое множество
$A = \{a_1, a_2, \ldots, a_n\}$	множество с элементами $a_1, a_2, \ldots, a_n$
$\mathbb{N}$	множество натуральных чисел
$\mathbb{Z}$	множество целых чисел
$\mathbb{R}$	множество действительных чисел
$B \subset A$	B является подмножеством множества A
$B \subseteq A$	B является подмножеством множества A
$\{a \in A \,\vert\, \mathcal{X}\}$	подмножество A, заданное условием $\mathcal{X}$
$\bigcup$	объединение (множеств)
$\bigcap$	пересечение (множеств)
$A \setminus B$	разность множеств A и B
$\times$	декартово произведение (множеств)
$\sum$	знак суммирования
$f : X \to Y$	отображение из X в Y
$\mathbf{M}_{m\times n}$	множество матриц размера $m \times n$
$\mathbf{M}^{+}_{n\times n}$	симметрические матрицы $n \times n$
$\mathbf{M}^{-}_{n\times n}$	кососимметрические матрицы $n \times n$
(a_{ij})	матрица с элементами a_{ij}
$a_{i\bullet}$	стока с номером i матрицы (a_{ij})
$a_{\bullet j}$	столбец с номером j матрицы (a_{ij})
$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$	диагональная матрица
E	единичная матрица
E_n	единичная матрица порядка n
$A + B$	сумма матриц A и B
λA	произведение матрицы A на число λ

O	нулевая матрица (некоторого размера)
$-A$	матрица, противоположная матрице A
A^T	матрица, транспонированная к матрице A
AB	произведение матриц A и B
A^{-1}	матрица, обратная к матрице A
$\langle A_1, A_2, \ldots, A_k \rangle$	линейная оболочка системы $A_1, A_2, \ldots, A_k$
$\langle \mathcal{A} \rangle$	линейная оболочка системы $\mathcal{A}$
$\operatorname{rg} \mathcal{A}$	ранг системы столбцов (строк, матриц) $\mathcal{A}$
$\operatorname{rg} A$	ранг матрицы A
$\operatorname{rg}_h A$	строчный ранг матрицы A $(= \operatorname{rg} A)$
$\operatorname{rg}_v A$	столбцовый ранг матрицы A $(= \operatorname{rg} A)$
$\operatorname{tr} A$	след матрицы A
$\widetilde{P}_{ij}$	элементарное преобразование строк I типа
$\widetilde{D}_i(\lambda)$	элементарное преобразование строк II типа
$\widetilde{T}_{ij}(\lambda)$	элементарное преобразование строк III типа
P_{ij}	элементарная матрица I типа
$D_i(\lambda)$	элементарная матрица II типа
$T_{ij}(\lambda)$	элементарная матрица III типа
$AX = b$	матричная запись системы
$(A \mid b)$	расширенная матрицы системы
$\operatorname{Sol}(A \mid b)$	общее решение системы
$\lvert A \rvert$	определитель (детерминант) матрицы A
$\det A$	определитель (детерминант) матрицы A

Предметный указатель