# ЛЕКЦИЯ 7

# МАТРИЦЫ И ОТОБРАЖЕНИЯ

# ПРОИЗВЕДЕНИЕ МАТРИЦ

# ТРАНСПОНИРОВАНИЕ МАТРИЦ

# РАНГ ПРОИЗВЕДЕНИЯ МАТРИЦ

# МАТРИЦЫ И ОТОБРАЖЕНИЯ

Пусть $\mathbb{R}^n$ и $\mathbb{R}^m$ — векторные пространства столбцов высоты $n$ и $m$ соответственно. Пусть, далее, $A = (a_{ij})$ — матрица размера $m \times n$.

Определим отображение

$$\varphi_A : \mathbb{R}^n \to \mathbb{R}^m,$$

полагая для любого

$$X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n,$$

что

$$\varphi_A(X) = x_1 A^{(1)} + x_2 A^{(2)} + \cdots + x_n A^{(n)},$$

где $A^{(1)}, A^{(2)}, \ldots, A^{(n)}$ — столбцы матрицы $A$. Так как они имеют высоту $m$, то в правой части стоит вектор-столбец

$$Y = [y_1, y_2, \ldots, y_m] \in \mathbb{R}^m.$$

Более подробно это выражение переписывается в виде

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \ldots, m.$$

Если

$$X = X' + X'' = [x_1' + x_1'', x_2' + x_2'', \ldots, x_n' + x_n''],$$

то

$$\varphi_A(X' + X'') = \sum_{i=1}^{n} (x_i' + x_i'') A^{(i)} = \sum_{i=1}^{n} x_i' A^{(i)} + \sum_{i=1}^{n} x_i'' A^{(i)} = \\ = \varphi_A(X') + \varphi_A(X'').$$

Аналогично,

$$\varphi_A(\lambda X) = \sum_{i=1}^{n} \lambda x_i A^{(i)} = \lambda \sum_{i=1}^{n} A^{(i)} = \lambda \varphi_A(X), \quad \lambda \in \mathbb{R}.$$

Таким образом, для введенного отображения $\varphi_A$ всегда выполняются два свойства:

(1) $\varphi(X' + X'') = \varphi(X') + \varphi(X'')$ для всех $X', X'' \in \mathbb{R}^n$;

(2) $\varphi(\lambda X) \lambda \varphi(X)$ для всех $X \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Определение 1. Отображение $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, обладающее свойствами (1) и (2), называется *линейным отображением* из $\mathbb{R}^n$ в $\mathbb{R}^m$.

Как мы знаем, $\mathbb{R}^n = \langle E^{(1)}, \ldots, E^{(n)} \rangle$ — линейная оболочка стандартных базисных столбцов, так что

$$X = [x_1, x_2, \ldots, x_n] = \sum_{j=1}^{n} x_j E^{(j)}.$$

Из свойств линейности имеем

$$\varphi(X) = \varphi\left(\sum_{j=1}^{n} x_j E^{(j)}\right) = \sum_{j=1}^{n} x_j \varphi(E^{(j)}).$$

Это соотношение показывает, что отображение $\varphi$ полностью определяется своими значениями на базисных векторах-столбцах.

Положив

$$\varphi(E^{(j)}) = [a_{1j}, a_{2j}, \ldots, a_{ml}] = A^{(j)} \in \mathbb{R}^m,$$

видим, что задание $\varphi$ равносильно заданию прямоугольной матрицы $A = (a_{ij})$ размера $m \times n$ со столбцами $A^{(1)}, \ldots, A^{(n)}$. Значит, можно положить $\varphi = \varphi_A$.

Будем называть такую соответствующую отображению матрицу *матрицей линейного отображения* $\varphi_A$.

Пусть $\varphi_A$ и $\varphi_{A'}$ — два линейных отображения $\mathbb{R}^n \to \mathbb{R}^m$ с матрицами $A = (a_{ij})$ и $A' = (a'_{ij})$. Тогда равенство $\varphi_A = \varphi_{A'}$ равносильно совпадению значений

$$\varphi_A(X) = \varphi_{A'}(X) \text{ для всех } X \in \mathbb{R}^n.$$

В частности,

$$A'^{(j)} = \varphi_{A'}(E^{(j)}) = \varphi_A(E^{(j)}) = A^{(j)}, \quad 1 \leqslant j \leqslant n,$$

откуда $a'_{ij} = a_{ij}$ и $A' = A$.

Резюмируем полученные результаты:

**Теорема 1.** *Между линейными отображениями $\mathbb{R}^n$ в $\mathbb{R}^m$ и матрица размера $m \times n$ существует взаимно однозначное соответствие.*

Обратим внимание на специальный случай $m = 1$, когда линейное отображение $\mathbb{R}^n \to \mathbb{R}$, обычно называемое *линейной функцией* от $n$ переменных, задается $n$ скалярами $a_1, a_2, \ldots, a_n$:

$$\varphi(X) = \varphi(x_1, x_2, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n.$$

Линейные функции и произвольные линейные отображения $\mathbb{R}^n$ в $\mathbb{R}^m$ при фиксированных $n$ и $m$ можно складывать и умножать на скаляры.

В самом деле, пусть $\varphi_A, \varphi_B : \mathbb{R}^n \to \mathbb{R}^m$ — два линейных отображения. Отображение

$$\varphi = \alpha\varphi_A + \beta\varphi_B : \mathbb{R}^n \to \mathbb{R}^m, \quad \alpha, \beta \in \mathbb{R},$$

определяется своими значениями

$$\varphi(X) = \alpha\varphi_A(X) + \beta\varphi_B(X).$$

В правой части стоит обычная линейная комбинация векторов-столбцов.

Так как

$$\begin{aligned}\varphi(X' + X'') &= \alpha\varphi_A(X' + X'') + \beta\varphi_B(X' + X'') = \\ &= \alpha(\varphi_A(X') + \varphi_A(X'')) + \beta(\varphi_B(X') + \varphi_B(X'')) = \\ &= (\alpha\varphi_A(X') + \beta\varphi_B(X')) + (\alpha\varphi_A(X'') + \beta\varphi_B(X'')) = \varphi(X') + \varphi(X'')\end{aligned}$$

и

$$\begin{aligned}\varphi(\lambda X) &= \alpha\varphi_A(\lambda X) + \beta\varphi_B(\lambda X) = \alpha\lambda\varphi_A(X) + \beta\lambda\varphi_B(X) = \\ &= \lambda(\alpha\varphi_A(X) + \beta\varphi_B(X)) = \lambda\varphi(X),\end{aligned}$$

то $\varphi$ — линейное отображение. Значит, можно говорить о его матрице $C$: $\varphi = \varphi_C$.

# ПРОИЗВЕДЕНИЕ МАТРИЦ

Пусть $\varphi_B : \mathbb{R}^n \to \mathbb{R}^s$, $\varphi_A : \mathbb{R}^s \to \mathbb{R}^m$ — линейные отображения, $\varphi_C = \varphi_A \circ \varphi_B$ — их композиция.

Проверим сначала, что $\varphi = \varphi_A \circ \varphi_B$ — линейное отображение:

$$\begin{aligned}\varphi(X' + X'') = \varphi_A(\varphi_B(X' + X'')) = \varphi_A(\varphi_B(X') + \varphi_B(X'')) = \\ = \varphi_A(\varphi_B(X')) + \varphi_A(\varphi_B(X'')) = \varphi(X') + \varphi(X'');\end{aligned}$$

и

$$\varphi(\lambda X) = \varphi_A(\varphi_B(\lambda X) = \varphi_A(\lambda\varphi_B(X)) = \lambda\varphi_A(\varphi_B(X)) = \lambda\varphi(X).$$

Таким образом, с отображением $\varphi$ ассоциируется определенная матрица $C$.

Чтобы найти ее, предположим, что при отображении $\varphi_B$ столбик $[x_1, \ldots, x_n]$ переходит в столбик $[y_1, \ldots, y_s]$, который, в свою очередь, при отображении $\varphi_A$ переходит в столбик $[z_1, \ldots, z_m]$.

Запишем это в явном виде как сумму:

$$z_i = \sum_{k=1}^{s} a_{ik} y_k = \sum_{k=1}^{s} a_{ik} \sum_{j=1}^{n} b_{kj} x_j = \sum_{j=1}^{n} \left( \sum_{k=1}^{s} a_{ik} b_{kj} \right) x_j.$$

С другой стороны,

$$z_i = \sum_{j=1}^{n} c_{ij} x_j, \quad i = 1, 2, \ldots, m.$$

Сравнивая полученные выражения, мы приходим к соотношениям

$$c_{ij} = \sum_{k=1}^{s} a_{ik} b_{kj}, \quad i = 1, \ldots, m, \ j = 1, \ldots, n.$$

Будем говорить, что матрица $C$ получается в результате *умножения* матрицы $A$ на матрицу $B$. Принято писать

$$C = AB.$$

Таким образом, произведением прямоугольной матрицы $(a_{ik})$ размера $m \times s$ и прямоугольной матрицы $(b_{kj})$ размера $s \times n$ называется прямоугольная матрица $(c_{ij})$ размера $m \times n$ с элементами $c_{ij}$, задающимися формулой, выписанной выше.

Нами доказана

**Теорема 2.** *Произведение $\varphi_A \varphi_B$ двух линейных отображений с матрицами $A$ и $B$ является линейным отображением с матрицей $C = AB$,*

$$\varphi_A \circ \varphi_B = \varphi_{AB}.$$

Мы можем забыть о линейных отображениях и находить произведение $AB$ двух произвольных матриц $A, B$, имея при этом в виду, что такое произведение имеет смысл только в том случае, когда число столбцов в матрице $A$ совпадает с числом строк в матрице $B$.

Именно при этом условии работает правило умножения $i$-й строки $A_{(i)}$ на $j$-й столбец $B^{(j)}$, согласно которому

$$c_{ij} = (a_{i1}, \ldots, a_{is})[b_{1j}, \ldots, b_{sj}] = A_{(i)} B^{(j)}.$$

Число строк матрицы $AB$ равно числу строк матрицы $A$, а число столбцов — числу столбцов матрицы $B$.

В частности, произведение квадратных матриц одинаковых порядков всегда определено, но даже в этом случае, вообще говоря, $AB \neq BA$, как показывает простой пример

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

**Следствие 1.** *Умножение матриц ассоциативно:*

$$A(BC) = (AB)C.$$

*Доказательство.* Действительно, произведение матриц соответствует композиции линейных отображений, а композиция всегда ассоциативна. □

Обратим внимание еще и на так называемые законы дистрибутивности

$$(A+B)C = AC + BC, \quad D(A+B) = DA + DB,$$

где $A, B, C, D$ — произвольные матрицы размеров, соответственно, $m \times s$, $m \times s$, $s \times n$, $n \times m$.

Действительно, полагая

$$A = (a_{ij}), B = (b_{ij}), C = (c_{ij}),$$

мы получим для любых $i, j$ равенство

$$\sum_{k=1}^{n}(a_{ik} + b_{ik})c_{kj} = \sum_{k=1}^{n} a_{ik}c_{kj} + \sum_{k=1}^{n} b_{ik}c_{kj},$$

левая часть которого дает $i, j$-й элемент матрицы $(A+B)C$, а правая часть — $i, j$-й элемент матрицы $AC + BC$.

Второй закон дистрибутивности проверяется совершенно аналогично.

## ТРАНСПОНИРОВАНИЕ МАТРИЦ

Будем говорить, что матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ и } A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

размеров $m \times n$ и $n \times m$ соответственно получаются друг из друга *транспонированием* — заменой строк на столбцы, а столбцов на строки.

Непосредственно видно, что

$$(A^T)^T = A, \quad (A+B)^T = A^T + B^T, \quad (\lambda A^T) = \lambda A^T.$$

Транспонирование произведения матриц подчиняется более интересной закономерности.

Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1s} \\ a_{21} & a_{22} & \dots & a_{2s} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{ms} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{s1} & b_{s2} & \dots & b_{sn} \end{pmatrix},$$

и

$$A^T = (a'_{ki}), \quad B^T = (b'_{jk}),$$

где

$$a'_{ki} = a_{ik}, \quad b'_{jk} = b_{kj}.$$

Вычислим

$$C = AB = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix}$$

и

$$D = B^T A^T = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1m} \\ d_{21} & d_{22} & \dots & d_{2m} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & d_{nm} \end{pmatrix}$$

с помощью полученных нами формул:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \qquad d_{ji} = \sum_{k=1}^{n} b'_{jk} a'_{ki} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Отсюда видно, что $d_{ji} = c_{ij}$ для всех $i = 1, \dots, m$ и $j = 1, \dots, n$. Значит, $C^T = D$, или, как было исходно,

$$(AB)^T = B^T A^T.$$

Понятно, что если имелось несколько матриц $A_1, A_2, \dots, A_k$, для которых было определено последовательное произведение, то

$$(A_1 A_2 \dots A_k)^T = A_k^T \dots A_2^T A_1^T.$$

Благодаря тому, что ранги матриц по строкам и столбцам совпадают и равны рангу матрицы,

$$\operatorname{rank} A^T = \operatorname{rank} A.$$

## РАНГ ПРОИЗВЕДЕНИЯ МАТРИЦ

Пусть $A$ и $B$ — произвольные матрицы размеров $m \times s$ и $s \times n$. Что можно сказать о величине ранга $\operatorname{rank} AB$?

**Теорема 3.** *Справедливо неравенство*

$$\operatorname{rank} AB \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\}.$$

*Доказательство.* Для строк $C_{(i)}$ и столбцов $C^{(i)}$ матрицы $C = AB$ мы имеем выражения

$$C_{(i)} = A_{(i)}B, \quad C^{(j)} = AB^{(j)}.$$

Интерпретируя теперь ранг матрицы $A$ как

$$r_1 = \operatorname{rank} A = \dim\langle A_{(1)}, A_{(2)}, \ldots, A_{(m)}\rangle,$$

считаем без ограничения общности базисными строки $A_{(1)}, \ldots, A_{(r_1)}$, поскольку необходимая перестановка строк в $A$ будет сопровождаться точно такой же перестановкой строк матрицы $C$, а это преобразование никак не меняет ни один из рангов.

Итак,

$$A_{(k)} = \sum_{i=1}^{r_1} \lambda_{ki} A_{(i)}, \quad r_1 < k \leqslant m,$$

откуда (используя дистрибутивность)

$$C_{(k)} = A_{(k)}B = \left(\sum_{i=1}^{r_1} \lambda_{ki} A_{(i)}\right) B = \sum_{i=1}^{r_1} \lambda_{ki}(A_{(i)}B) = \sum_{i=1}^{r_1} \lambda_{ki} C_{(i)}.$$

Таким образом,

$$\langle C_{(1)}, \ldots, C_{(m)} \rangle = \langle C_{(1)}, \ldots, C_{(r_1)} \rangle.$$

Следовательно,

$$\operatorname{rank} C = \dim \langle C_{(1)}, \ldots, C_{(m)} \rangle \leqslant r_1 = \operatorname{rank} A.$$

Аналогично, интерпретируя ранг матрицы $B$ как

$$r_2 = \operatorname{rank} B = \dim \langle B^{(1)}, B^{(2)}, \ldots, B^{(n)} \rangle$$

и считая без ограничения общности базисными столбцы $B^{(1)}, \ldots, B^{(r_2)}$, будем иметь

$$B^{(k)} = \sum_{j=1}^{r_2} \mu_{kj} B^{(j)},$$

$$C^{(k)} = AB^{(k)} = A \left( \sum_{j=1}^{r_2} \mu_{kj} B^{(j)} \right) \sum_{j=1}^{r_2} \mu_{kj} AB^{(j)} = \sum_{j=1}^{r_2} \mu_{kj} C^{(j)},$$
$$r_2 < k \leqslant n,$$

откуда

$$\operatorname{rank} C = \dim \langle C^{(1)}, \ldots, C^{(n)} \rangle \leqslant r_2 = \operatorname{rank} B.$$

$\square$