Лекция 8

Мы видели, что каждому линейному отображению $\varphi\colon V \to W$ при фиксированных базисах $e = \{e_1, \ldots, e_n\}$ в V и $f = \{f_1, \ldots, f_m\}$ в W можно сопоставить матрицу $A(\varphi, e, f)$ размера $m \times n$. В j-ом столбце матрицы $A(\varphi, e, f)$ стоят координаты вектора $\varphi(e_i)$ в базисе f.

Предложение 1. *Соответствие $\varphi \to A(\varphi, e, f)$ задает биекцию между линейными отображениями $V \to W$ и матрицами $m \times n$.*

Доказательство. Сперва докажем следующую полезную лемму.

Лемма 1. *Пусть есть два векторных пространства V и W. Пусть $e_1, \ldots, e_n$ – базис V и $w_1, \ldots, w_n$ – векторы из W. Тогда существует единственное линейное отображение $\psi\colon V \to W$ такое, что $\psi(e_i) = w_i$ для всех i.*

Доказательство леммы. Рассмотрим отображение $\psi\colon V \to W$, заданное по формуле

$$\psi\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n a_i w_i.$$

Ясно, что так как $e_1, \ldots, e_n$ – базис, то любой вектор из V представляется в виде $\sum_{i=1}^n a_i e_i$ единственным образом и потому отображение определено корректно. Также видно из определения, что $\psi(e_i) = w_i$ для всех i. Проверим, что ψ – линейное отображение. В самом деле, пусть $v_1 = \sum a_i e_i$ и $v_2 = \sum b_i e_i$. Тогда

$$\psi(v_1+v_2) = \psi\left(\sum a_i e_i + \sum b_i e_i\right) = \psi\left(\sum (a_i + b_i) e_i\right) = \sum (a_i+b_i) w_i = \sum a_i w_i + \sum b_i w_i = \psi(v_1)+\psi(v_2);$$

$$\psi(\lambda v_1) = \psi\left(\lambda \sum a_i e_i\right) = \psi\left(\sum \lambda a_i e_i\right) = \sum \lambda a_i w_i = \lambda \sum a_i w_i = \lambda(\psi(v_1)).$$

Итак, мы проверили, что ψ линейно. То есть мы предъявили линейное отображение, переводящее e_i в w_i для всех i. Осталось объяснить, что линейное отображение с такими свойствами единственно. Для этого заметим, что раз ψ линейно, то выполнено

$$\psi\left(\sum a_i e_i\right) = \sum (a_i \psi(e_i)) = \sum a_i w_i.$$

То есть если ψ – линейное отображение, переводящее e_i в w_i для всех i, то оно задается именно формулой

$$\psi\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n a_i w_i$$

и никак иначе. Лемма доказана. □

Предложение сразу следует из Леммы, так как задание матрицы равносильно заданию образов $\varphi(e_i)$, так как в i-ом столбце матрицы как раз и стоят координаты этого вектора в базисе f. Получается, что для любой матрицы существует единственное линейное отображение, которое имеет эту матрицу в заданных базисах. Предложение доказано. □

На прошлой лекции мы обсудили, что с точки зрения сложения и умножения на число множество матриц $m \times n$ образует векторное пространство размерности mn с базисом из матричных единиц (не путайте с единичной матрицей!) E_{ij}, $1 \le i \le m$, $1 \le j \le n$. Разберем несколько свойств умножения матриц.

Теорема 1 (Свойства умножения матриц)**.** *В следующих выражениях если имеет смысл одна из сторон равенства, то имеет смысл и вторая и тогда они равны.*

1) $(AB)C = A(BC)$*;*
2) $A(B + C) = AB + AC$*;*
3) $(A + B)C = AB + AC$*;*
4) $\lambda(AB) = (\lambda A)B = A(\lambda B)$*.*

Доказательство. Для каждого утверждения есть 2 доказательства. Одно из них более техничное, через выкладки с матрицами. Второе – более концептуальное через линейные отображения. Дадим все эти доказательства. Будем оброзначать доказательство I, если оно через метрицы и II – через отображения.

1) доказательство I. Для того, чтобы имела смысл левая часть нужно, чтобы матрица A была $m\times n$, матрица B чтобы была $n \times k$. Тогда их можно перемножить и получить матрицу AB размера $m \times k$. Чтобы ее можно было умножить на матрицу C нужно, чтобы C была $k\times r$. При этом матрица $(AB)C$ будет иметь размер $m \times r$. Несложно проследить ровно так же за выражением $A(BC)$ и увидеть, что оно имеет смысл также только при размерах $m \times n$, $n \times k$ и $k \times r$ матриц A, B и C соответственно, а итоговая матрица будет иметь размеры $m\times r$. Итак, условия на размеры и размеры итоговых матриц совпали.

Введем обозначения. $AB = D$, $DC = P$, $BC = F$, $AF = Q$. По определению произведения матриц имеем

$$p_{ij} = \sum_{l=1}^{k} d_{il}c_{lj}.$$

При этом

$$d_{il} = \sum_{t=1}^{n} a_{it}b_{tl}.$$

Получаем

$$p_{ij} = \sum_{l=1}^{k}\left(\left(\sum_{t=1}^{n} a_{it}b_{tl}\right)c_{lj}\right) = \sum_{l=1}^{k}\sum_{t=1}^{n} a_{it}b_{tl}c_{lj}.$$

Аналогично

$$q_{ij} = \sum_{t=1}^{n} a_{it}f_{tj}.$$

При этом

$$f_{tj} = \sum_{l=1}^{k} b_{tl}c_{lj}.$$

Получаем

$$q_{ij} = \sum_{t=1}^{n}\left(a_{it}\left(\sum_{l=1}^{k} b_{tl}c_{lj}\right)\right) = \sum_{t=1}^{n}\sum_{l=1}^{k} a_{it}b_{tl}c_{lj}.$$

Видим, что выражения для p_{ij} и q_{ij} отличаются только порядком суммирования, а значит, равны.

1) доказательство II. Рассмотрим 4 векторных пространства $\dim U = r$, $\dim V = k$, $\dim W = n$, $\dim L = m$. Фиксируем некоторые базисы в каждом из этих пространств, пусть эти базисы соответственно e, f, s, z. И рассмотрим отображения $\varphi\colon U \to V$, $\psi\colon V \to W$, $\alpha\colon W \to L$ такие, что $A(\varphi, e, f) = C$, $A(\psi, f, s) = B$, $A(\alpha, s, z) = A$. Тогда равенство $(AB)C = A(BC)$ равносильна равенству

$$(\alpha \circ \psi) \circ \varphi = \alpha \circ (\psi \circ \varphi).$$

Докажем последнее равенство (причем заметим, что для него не нужно даже, чтобы отображения были линейными и даже, чтобы U, V, W и L были пространствами). Чтобы доказать нужное равенство, применим обе части к некоторому элементу $u \in U$.

$$(\alpha \circ \psi) \circ \varphi(u) = (\alpha \circ \psi(\varphi(u)) = \alpha(\psi(\varphi(u))).$$

$$\alpha \circ (\psi \circ \varphi)(u) = \alpha((\psi \circ \varphi)(u)) = \alpha(\psi(\varphi(u))).$$

2) доказательство I. для того, чтобы имело смысл выражение $A(B + C)$, матрицы B и C должны быть одинакового размера $n \times k$, а матрица A – размера $m \times n$. итоговая матрица получится $m \times k$. Аналогично, выражение $AB + AC$ имеет смысл только когда матрицу A можно умножить на B и C и получатся матрицы одинаковых размеров. Это дает те же условия.

При этом, если $B + C = D$, $AD = P$, $AB = F$, $AC = R$, $F + R = Q$ то

$$p_{ij} = \sum_{l=1}^{n} a_{il}d_{lj} = \sum_{l=1}^{n} a_{il}(b_{lj} + c_{lj}) = \sum_{l=1}^{n}(a_{il}b_{lj} + a_{il}c_{lj}) = \sum_{l=1}^{n} a_{il}b_{lj} + \sum_{l=1}^{n} a_{il}c_{lj} = f_{ij} + r_{ij} = q_{ij}.$$

2) доказательство II. Рассмотрим 3 векторных пространства $\dim U = k$, $\dim V = n$ и $\dim W = m$ и фиксируем базисы e, f и s в них. Рассмотрим отображения $\varphi\colon U \to V$, $\xi\colon U \to V$ и $\psi\colon V \to W$ и такие, что $A(\varphi, e, f) = B$, $A(\xi, e, f) = C$ и $A(\psi, f, s) = A$. Тогда равенство с матрицами эквивалентно равенству

$$\psi \circ (\varphi + \xi) = \psi \circ \varphi + \psi \circ \xi.$$

Чтобы доказать последнее равенство, применим обе части к некоторому вектору $u \in U$. Получим

$$\psi\circ(\varphi+\xi)(u) = \psi((\varphi+\xi)(u)) = \psi(\varphi(u)+\xi(u)) = \psi(\varphi(u))+\psi(\xi(u)) = \psi\circ\varphi(u)+\psi\circ\xi(u) = (\psi\circ\varphi+\psi\circ\xi)(u).$$

Заметьте, что здесь мы воспользовались линейностью ψ.

3) доказывается ровно так же, как и 2).

4) Это самый простой пункт. Докажите в качестве упражнения (желательно 2-мя способами). □

У нас есть следующие операции на матрицах (не всегда они возможны с данными матрицами): сложение, умножение на число, умножение, вычитание, транспонирование. Нужно договориться о том, какой приоритет (порядок действий) у данных операций, если не стоит скобок. Договариваемся так: максимальный приоритет у транспонирования, затем умножение на чило, затем умножение, затем одинаковый приоритет у сложения и вычитания. (Можно считать, что у умножения и умножения на число приоритет одинаков.) Надо понимать, что это вопрос договоренности и никакого утверждения в этом нет. Если в вашем выражении нужно сделать действия в другом порядке, запишите его используя скобки.

Определение 1. Пусть у нас есть бинарная операция (например, умножение). Она называется обобщенно ассоциативной, если произведение любого количества сомножителей не зависит от расстановки скобок в этом произведении.

Теорема 2. *Ассоциативная операция всегда обобщенно ассоциативна.*

Доказательство. Нам надо доказать, что произведение (с помощью нашей операции, но далее будем для удобства говорить об умножении) любых k множителей не зависит от расстановки скобок. Докажем это по индукции по k.

База индукции. $k = 3$. В этом случае обобщенная ассоциативность совпадает с ассоциативностью.

Шаг индукции. Пусть для $k < n$ утверждение верно. Докажем его для $k = n$. Введем *стандартную расстановку скобок*:

$$(\ldots(x_1x_2)x_2)x_4)\ldots)x_n.$$

При такой расстановке скобок операции делаются слева-направо. Нам нужно взять провольную расстановку скобок и доказать, что результат для нее такой же, как и для стандартной. Рассмотрим некоторую расстановку скобок в произведении $x_1\ldots x_n$. Пусть последнее умножение перемножает результаты умножения в скобках A и B. (Возможно скобка A или B состоит только из одного x_i.) Есть 2 случая:

Случай 1: B состоит только из x_n. Тогда по предположению индукции внутри скобки A расстановку скобок можно считать стандартной. Тогда во всем выражении имеем стандартную расстановку скобок.

Случай 2: B состоит не только из x_n. Тогда внутри B по предположению индукции расстановку скобок можно считать стандартной. Тогда $B = C \cdot x_n$. То есть все выражение имеет вид $A(Cx_n)$. Применяя ассоциативность, получаем

$$A(Cx_n) = (AC)x_n.$$

Теперь мы попали в случай 1. □

Замечание 1. Данная теорема применима к умножению матриц и к сложению матриц. Более того формально говоря ее стоило доказывать для сложения и умножения чисел, но формально в школе этого обычно не делают, хотя мы все и привыкли, что она верна.

Все предыдущие равенства показывают, что для умножения матриц выполнены некоторые свойства, к которым мы привыкли для чисел. Но есть и отличия. Например, для чисел выполенено равенство $AB = BA$. Однако для того, чтобы имела смысл левая часть нужно, чтобы матрица A имела размеры $m\times n$, а матрица B – размеры $n\times k$. При этом, если $k \neq m$, то правая часть не имеет смысла. Пусть даже $k = m$. Тогда AB – матрица $m \times m$, а матрица BA имеет размер $n \times n$. Если $m \neq n$, то размеры этих матриц разные. Ограничимся ситуацией, когда $m = n = k$. Тогда матрицы AB и BA определены и имеют одинаковый размер $n \times n$. Однако, даже и в этом случае может быть $AB \neq BA$. Например, пусть

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Тогда
$$AB = \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} \neq \begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix} = BA.$$

Замечание 2. Не стоит думать, что матрицы AB и BA никогда не совпадают. Они могут совпадать (конечно же только для случая, когда A и B квадратные), а могут не совпадать.

Упомянем несколько свойств операци, содержащие транспонирование. (Здесь снова если имеет смысл одна из частей, то и вторая и они равны.)
1) $(A^T)^T = A$;
2) $(A \pm B)^T = A^T \pm B^T$;
3) $(AB)^T = B^T A^T$.
Пожалуй, в доказательстве нуждается лишь свойство 3 (остальные уже были).

Доказательство свойства 3. Пусть A имеет размер $m \times n$. Чтобы имела смысл левая часть нужно, чтобы B была $n \times k$. При этом $(AB)^T$ будет $k \times n$. Смотрим на правую часть. Пусть A имеет размер $m \times n$, тогда A^T имеет размер $n \times m$. Чтобы множно было умножить B^T на A^T нужно, чтобы B^T было $k \times n$, то есть B имеет размеры $n \times k$. Итоговая же матрица снова $k \times m$.

Обозначим $AB = C$, $C^T = D$, $A^T = P$, $B^T = Q$, $QP = R$. Тогда
$$d_{ij} = c_{ji} = \sum_{l=1}^{n} a_{jl} b_{li}.$$
$$r_{ij} = \sum_{l=1}^{k} q_{il} p_{lj} = \sum_{l=1}^{k} b_{li} a_{jl}.$$
Так как выражения совпали, получаем заявленное равенство. □

Определение 2. Для каждого n рассмотрим следующую квадратную матрицу $n \times n$
$$E = E_n = \begin{pmatrix}1 & 0 & 0 & \dots & 0\\ 0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & 1\end{pmatrix}.$$
Такая матрица называется *единичной матрицей.* По *главной диагонали*, то есть сверху слева вправо вниз стоят единицы, а остальные элементы матрицы – нули.

Лемма 2. *Для любой матрицы A размера $m \times n$ выполнено $AE = A$. Для любой матрицы B размера $n \times k$ выполнено $EB = B$.*

Доказательство. Для доказательства удобно использовать символ Кронекера
$$\delta_{ij} = \begin{cases}1, & \text{при } i = j;\\ 0, & \text{при } i \neq j.\end{cases}$$
Элемент матрицы E на месте (i, j) равен δ_{ij}.

Пусть $AE = C$. Тогда
$$c_{ij} = \sum_{l=1}^{n} a_{il} \delta_{lj} = 0 + \ldots + 0 + a_{ij} + 0 + \ldots + 0.$$
То есть $C = A$.

Второе равенство доказывается аналогично. □

Определение 3. Пусть A – матрица $m \times n$. Матрица B размера $n \times m$ называется *левой обратной* к матрице A, если $BA = E$.

Матрица B размера $n \times m$ называется *правой обратной* к матрице A, если $AB = E$.

Матрица B размера $n \times n$ называется *(двустороней) обратной* к матрице A, если она и левая и правая обратная к A.

На следующей лекции мы получим критерии существования левой и правой обратной и докажем, что двустороняя обратная бывает только у квадратных матриц и то не у всех.