

# Учебник по Эконометрике

## Лекция 2: Линейная Алгебра и Матрицы

Джон Стачурски

Лекции: Акшай Шенкер

Перевел: Алексей Кедо

8 октября 2020 г.

# Матрицы

Обычная  $N \times K$  матрица:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NK} \end{pmatrix}$$

Символ  $a_{nk}$  означает элемент, стоящий в  $n$ -ой строке  $k$ -ого столбца

$N \times K$  матрица также называется

- вектором строки, если  $N = 1$
- вектором столбца, если  $K = 1$

Если  $N = K$ , то матрицу  $A$  называют **квадратной**

Если  $A$  квадратная и  $a_{nk} = a_{kn}$  для любых  $k$  и  $n$ , то  $A$  называют **симметричной**

Часто элементы матрицы  $A$  представляют собой коэффициенты в системе линейных уравнений

$$\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1K}x_K = b_1 \\ \vdots \\ a_{N1}x_1 + a_{N2}x_2 + \cdots + a_{NK}x_K = b_N \end{array}$$

Для матрицы  $A$  применяются обозначения:

- $\text{row}_n \mathbf{A}$  означает  $n$ -ую строку  $\mathbf{A}$
- $\text{col}_k \mathbf{A}$  означает  $k$ -ый столбец  $\mathbf{A}$

Символы **0** и **1** представляют собой матрицы, все элементы которых равны нулю и единице соответственно

$$\begin{pmatrix} \textcolor{red}{a}_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & \textcolor{red}{a}_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & \textcolor{red}{a}_{NN} \end{pmatrix}$$
$$\mathbf{I} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

6/104



Сложение тоже выполняется поэлементно:

$$\begin{pmatrix} a_{11} & \cdots & a_{1K} \\ a_{21} & \cdots & a_{2K} \\ \vdots & \vdots & \vdots \\ a_{N1} & \cdots & a_{NK} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1K} \\ b_{21} & \cdots & b_{2K} \\ \vdots & \vdots & \vdots \\ b_{N1} & \cdots & b_{NK} \end{pmatrix} \\ := \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1K} + b_{1K} \\ a_{21} + b_{21} & \cdots & a_{2K} + b_{2K} \\ \vdots & \vdots & \vdots \\ a_{N1} + b_{N1} & \cdots & a_{NK} + b_{NK} \end{pmatrix}$$

Заметим, что матрицы должны быть одинакового размера



Умножение матриц:

Произведение  $\mathbf{AB}$ :  $i, j$ -ый элемент — скалярное умножение  $i$ -ой строки  $\mathbf{A}$  и  $j$ -ого столбца  $\mathbf{B}$

$$c_{ij} = \langle \text{row}_i \mathbf{A}, \text{col}_j \mathbf{B} \rangle = \sum_{k=1}^K a_{ik} b_{kj}$$

Для  $i = j = 1$ :

$$\begin{pmatrix} \textcolor{red}{a}_{11} & \cdots & a_{1K} \\ a_{21} & \cdots & a_{2K} \\ \vdots & \vdots & \vdots \\ a_{N1} & \cdots & a_{NK} \end{pmatrix} \begin{pmatrix} \textcolor{red}{b}_{11} & \cdots & b_{1J} \\ \textcolor{red}{b}_{21} & \cdots & b_{2J} \\ \vdots & \vdots & \vdots \\ \textcolor{red}{b}_{K1} & \cdots & b_{KJ} \end{pmatrix} = \begin{pmatrix} \textcolor{red}{c}_{11} & \cdots & c_{1J} \\ c_{21} & \cdots & c_{2J} \\ \vdots & \vdots & \vdots \\ c_{N1} & \cdots & c_{NJ} \end{pmatrix}$$

Здесь

$$c_{11} = \langle \text{row}_1(\mathbf{A}), \text{col}_1(\mathbf{B}) \rangle = \sum_{k=1}^K a_{1k} b_{k1}$$

Предположим, что размер  $\mathbf{A} = N \times K$ ,  $\mathbf{B} = J \times M$

- $\mathbf{AB}$  определена, только если  $K = J$
- Размер  $\mathbf{AB} = N \times M$

Запомните правило:

Произведение  $N \times K$  и  $K \times M$  равно  $N \times M$

Умножение не коммутативно:  $\mathbf{AB} \neq \mathbf{BA}$

Заметим, что произведение  $\mathbf{BA}$  определено, только если  $N = M$  соблюдается

**Факт.** (??) Для согласованных матриц  $A$ ,  $B$ ,  $C$  и скаляра  $\alpha$ , верно

1.  $A(BC) = (AB)C$ ,
2.  $A(B + C) = AB + AC$ ,
3.  $(A + B)C = AC + BC$ ,
4.  $A\alpha B = \alpha AB$ , и
5.  $AI = A$  и  $IA = A$ , где  $I$  — единичная матрица.

Здесь “согласованность” значит, что операция определена при заданных размерностях матриц

**$k$ -ая мощность** квадратной матрицы  $\mathbf{A}$  определяется как

$$\mathbf{A}^k := \underbrace{\mathbf{A} \cdots \mathbf{A}}_{k \text{ terms}}$$

Если матрица  $\mathbf{B}$  такая, что  $\mathbf{B}^2 = \mathbf{A}$ , то  $\mathbf{B}$  называется **квадратным корнем** матрицы  $\mathbf{A}$  и записывается как  $\sqrt{\mathbf{A}}$

Возьмем матрицу  $\mathbf{A}$  размерности  $N \times K$  и  $K \times 1$  вектор столбца  $\mathbf{x}$ , перемножение  $\mathbf{Ax}$ :

$$\mathbf{Ax} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NK} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{pmatrix}$$

$$= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{N2} \end{pmatrix} + \cdots + x_K \begin{pmatrix} a_{1K} \\ a_{2K} \\ \vdots \\ a_{NK} \end{pmatrix}$$

$$= \sum_{k=1}^K x_k \text{col}_k \mathbf{A}$$

## Матрицы как отображения

Можно размышлять о матрице  $\mathbf{A}$  размерна  $N \times K$  как об отображении из  $\mathbb{R}^K$  в  $\mathbb{R}^N$ :

$$\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$$

Такое отображение линейно

Как насчет примеров линейных функций, не использующих матрицы?

...на самом деле, таких не существует!

Множеству линейных функций из  $\mathbb{R}^K$  в  $\mathbb{R}^N$  инъективно соответствует множество матриц  $N \times K$ :

**Теорема.** (??) Пусть  $T$  является функцией из  $\mathbb{R}^K$  в  $\mathbb{R}^N$ .  
Следующее эквивалентно:

1.  $T$  линейна.
2. Существует матрица  $A$  размера  $N \times K$ , такая что  $Tx = Ax$  для всех  $x \in \mathbb{R}^K$ .



**Доказательство.** (1.  $\implies$  2.)

Пусть  $T: \mathbb{R}^K \rightarrow \mathbb{R}^N$  линейна

Мы собираемся построить матрицу  $\mathbf{A}$  размерности  $N \times K$ , такую что

$$T\mathbf{x} = \mathbf{A}\mathbf{x}, \quad \forall \mathbf{x} \in \mathbb{R}^K$$

Как обычно, пусть  $\mathbf{e}_k$  —  $k$ -ый канонический базисный вектор в  $\mathbb{R}^K$

Определим матрицу  $\mathbf{A}$  как  $\text{col}_k(\mathbf{A}) = T\mathbf{e}_k$ . Возьмем любой  $\mathbf{x} = (x_1, \dots, x_K) \in \mathbb{R}^K$ . По линейности, получается

$$T\mathbf{x} = T \left[ \sum_{k=1}^K x_k \mathbf{e}_k \right] = \sum_{k=1}^K x_k T\mathbf{e}_k = \sum_{k=1}^K x_k \text{col}_k(\mathbf{A}) = \mathbf{A}\mathbf{x}$$

**Доказательство.** ( $2 \implies 1$ ) Возьмем матрицу  $\mathbf{A}$  размера  $N \times K$  и пусть  $T$  определяется

$$T: \mathbb{R}^K \rightarrow \mathbb{R}^N, \quad T\mathbf{x} = \mathbf{A}\mathbf{x}$$

Возьмем любые  $\mathbf{x}, \mathbf{y}$  в  $\mathbb{R}^K$ , и любые скаляры  $\alpha$  и  $\beta$

Правила матричной арифметики говорят нам, что

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) := \mathbf{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{A}\mathbf{x} + \beta\mathbf{A}\mathbf{y} =: \alpha T\mathbf{x} + \beta T\mathbf{y}$$

Рассмотрим возможность решения системы линейных уравнений, например,  $\mathbf{Ax} = \mathbf{b}$

Существование: можем ли мы найти  $\mathbf{x}$ , удовлетворяющий уравнению при любых заданных  $\mathbf{b}$ ?

- является ли линейное отображение  $T\mathbf{x} = \mathbf{Ax}$  сюръекцией?
- эквивалентно, равен ли  $\text{rng } T$  всему  $\mathbb{R}^N$ ?

## Пространство столбцов

Диапазон  $T$  — все вектора вида  $T\mathbf{x} = \mathbf{Ax}$ , где  $\mathbf{x}$  варьируется в  $\mathbb{R}^K$

Для  $\mathbf{x} \in \mathbb{R}^K$  получается

$$\mathbf{Ax} = \sum_{k=1}^K x_k \text{col}_k \mathbf{A}$$

Таким образом,  $\text{rng } T$  равен **пространству столбцов**  $\mathbf{A}$  — линейная оболочка столбцов  $\mathbf{A}$

$$\text{colspace } \mathbf{A} := \text{span}\{\text{col}_1 \mathbf{A}, \dots, \text{col}_K \mathbf{A}\}$$

В итоге,

$$\text{colspace } \mathbf{A} = \text{rng } T = \{\mathbf{Ax} : \mathbf{x} \in \mathbb{R}^K\}$$

# Ранг

## Эквивалентные вопросы

- Насколько велик диапазон линейного отображения  $Tx = Ax$ ?
- Насколько велико пространство столбцов  $A$ ?

Очевидной мерой размера линейного подпространства является его размерность

Размерность  $\text{colspace } A$  известна как **ранг**  $A$

$$\text{rank } A := \dim \text{colspace } A$$

Так как  $\text{colspace } A$  — линейная оболочка  $K$  векторов, получается

$$\text{rank } A = \dim \text{colspace } A \leq K$$

**A** имеет **полный ранг системы столбцов**, если

$$\text{rank } A = \text{количество столбцов } A$$

**Факт.** (??) Для любой матрицы **A**, следующие утверждения эквивалентны:

1. **A** является полным рангом системы столбцов
2. Столбцы **A** линейно независимы
3. Если  $Ax = 0$ , то  $x = 0$

## Квадратные матрицы и обратимость

Рассмотрим случай с матрицей  $\mathbf{A}$  размера  $N \times N$

Мы ищем условия, при которых для каждого  $\mathbf{b} \in \mathbb{R}^N$ , существует только один  $\mathbf{x} \in \mathbb{R}^N$ , такой что  $\mathbf{Ax} = \mathbf{b}$

Пусть  $T$  является линейным отображением  $T\mathbf{x} = \mathbf{Ax}$

- Когда каждая точка  $\mathbf{b} \in \mathbb{R}^N$  имеет только один прообраз в  $T$ ?
- Эквивалентно, когда  $T$  — биекция?

Напомним, линейные биекции называются несингулярными функциями.

**Факт.** (??) Для матрицы  $\mathbf{A}$  размера  $N \times N$  следующее эквивалентно:

1. Столбцы  $\mathbf{A}$  линейно независимы.
2. Столбцы  $\mathbf{A}$  формируют базис  $\mathbb{R}^N$ .
3.  $\text{rank } \mathbf{A} = N$ .
4.  $\text{colspace } \mathbf{A} = \mathbb{R}^N$ .
5.  $\mathbf{Ax} = \mathbf{Ay} \implies \mathbf{x} = \mathbf{y}$ .
6.  $\mathbf{Ax} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$ .
7. Для каждого  $\mathbf{b} \in \mathbb{R}^N$ , уравнение  $\mathbf{Ax} = \mathbf{b}$  имеет решение.
8. Для каждого  $\mathbf{b} \in \mathbb{R}^N$ , уравнение  $\mathbf{Ax} = \mathbf{b}$  имеет единственное решение.



Если хоть одно — и, следовательно, все — из этих условий не выполняются, то **A** называется **сингулярной**



**Теорема.** (??) Для несингулярной  $\mathbf{A}$  следующие утверждения верны:

1. Существует квадратная матрица  $\mathbf{B}$ , такая что  $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ , где  $\mathbf{I}$  — единичная матрица. Матрица  $\mathbf{B}$  называется **обратной**  $\mathbf{A}$ , и записывается как  $\mathbf{A}^{-1}$ .
2. Для каждого  $\mathbf{b} \in \mathbb{R}^N$ , единственное решение задачи  $\mathbf{Ax} = \mathbf{b}$  записывается как

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

**Пример.** Рассмотрим линейный спрос для  $N$  товаров

$$q_n = \sum_{k=1}^N a_{nk} p_k + b_n, \quad n = 1, \dots, N$$

где  $q_n$  и  $p_n$  — количество и цена  $n$ -ого товара

Мы хотим вычислить обратную функцию спроса, которая показывает зависимость цены от количества

Запишем систему в матричной форме:  $\mathbf{q} = \mathbf{A}\mathbf{p} + \mathbf{b}$

Если столбцы  $\mathbf{A}$  линейно независимы, то система обратима — единственное решение существует для любых фиксированных  $\mathbf{q}$  и  $\mathbf{b}$ :

$$\mathbf{p} = \mathbf{A}^{-1}(\mathbf{q} - \mathbf{b})$$

**Факт.** (??) Пусть  $\mathbf{A}$  и  $\mathbf{B}$  — квадратные матрицы размера  $N \times N$ . Если  $\mathbf{B}$  является левой или правой обратной матрицы  $\mathbf{A}$ , то  $\mathbf{A}$  несингулярна и  $\mathbf{B}$  — ее обратная матрица.

**Факт.** (??) Если  $\mathbf{A}$  и  $\mathbf{B}$  несингулярны и  $\alpha \neq 0$ , то

1.  $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ ,
2.  $(\alpha \mathbf{A})^{-1} = \alpha^{-1} \mathbf{A}^{-1}$ , и
3.  $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$

## Определитель

Определитель — уникальное число для любой квадратной матрицы  $A$

Пусть  $S(N)$  является множеством всех биекций из  $\{1, \dots, N\}$  в саму себя

Для  $\pi \in S(N)$  определим **функцию знака**  $\pi$

$$\text{sgn}(\pi) := \prod_{m < n} \frac{\pi(m) - \pi(n)}{m - n}$$

**Определитель** матрицы  $A$  размера  $N \times N$  — это

$$\det A := \sum_{\pi \in S(N)} \text{sgn}(\pi) \prod_{n=1}^N a_{\pi(n)n}$$

**Факт.** (??) Если  $\mathbf{I}$  — единичная матрица размера  $N \times N$ ,  $\mathbf{A}$  и  $\mathbf{B}$  являются матрицами размера  $N \times N$  и  $\alpha \in \mathbb{R}$ , то

1.  $\det \mathbf{I} = 1$ ,
2.  $\mathbf{A}$  несингулярна тогда и только тогда, когда  $\det \mathbf{A} \neq 0$ ,
3.  $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$ ,
4.  $\det(\alpha \mathbf{A}) = \alpha^N \det(\mathbf{A})$ , и
5.  $\det(\mathbf{A}^{-1}) = (\det(\mathbf{A}))^{-1}$ .

В случае матрицы размера  $2 \times 2$  можно показать, что

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$



## Доказательство. (Факт ??)

Возьмем кважратную матрицу  $\mathbf{A}$  и предположим, что правая обратная матрица  $\mathbf{B}$  существует:

$$\mathbf{AB} = \mathbf{I}$$

Тогда обе матрицы  $\mathbf{A}$  и  $\mathbf{B}$  несингулярны, так как по правилам факта ??:

$$\det(\mathbf{A}) \det(\mathbf{B}) = 1$$

Оба определителя  $\det \mathbf{A}$  и  $\det \mathbf{B}$  ненулевые и, следовательно, обе матрицы несингулярные. Далее,  $\mathbf{AB} = \mathbf{I}$ , значит

$$\mathbf{BAB} = \mathbf{B}$$

Умножим на  $\mathbf{B}^{-1}$ , получается  $\mathbf{BA} = \mathbf{I}$

## Диагональные и треугольные матрицы

Квадратная матрица называется **нижней треугольной**, если каждый элемент строго выше главной диагонали равен нулю

Пример.

$$\mathbf{L} := \begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 6 & 1 \end{pmatrix}$$

Квадратная матрица называется **верхней треугольной**, если каждый элемент строго ниже главной диагонали равен нулю

Пример.

$$\mathbf{U} := \begin{pmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 1 \end{pmatrix}$$

Квадратная матрица называется **треугольной**, если она верхняя или нижняя треугольная

**Факт.** (??) Если  $\mathbf{A} = (a_{mn})$  треугольная, то  $\det \mathbf{A} = \prod_{n=1}^N a_{nn}$ .

Связанные линейные уравнения просто решить

Пример.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

становится

$$\begin{aligned} x_1 &= b_1 \\ 2x_1 + 5x_2 &= b_2 \\ 3x_1 + 6x_2 + x_3 &= b_3 \end{aligned}$$

Верхнее уравнение включает только  $x_1$ , так что его можно решить напрямую

Подставьте это значение во второе уравнение и решите для  $x_2$  и т.д.

Рассмотрим квадратную матрицу  $\mathbf{A}$  размера  $N \times N$

$N$  элементов  $a_{nn}$  называются **главной диагональю**

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix}$$

Квадратная матрица  $\mathbf{D}$  называется **диагональной**, если все значения вне главной диагонали равны нулю

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_N \end{pmatrix}$$

Часто записывается как

$$\mathbf{D} = \text{diag}(d_1, \dots, d_N)$$

Диагональные системы очень легко решить

Пример.

$$\begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

ЭКВИВАЛЕНТНО

$$d_1 x_1 = b_1$$

$$d_2 x_2 = b_2$$

$$d_3 x_3 = b_3$$

**Факт.** (??) Если  $\mathbf{C} = \text{diag}(c_1, \dots, c_N)$  и  $\mathbf{D} = \text{diag}(d_1, \dots, d_N)$ ,  
то

1.  $\mathbf{C} + \mathbf{D} = \text{diag}(c_1 + d_1, \dots, c_N + d_N)$
2.  $\mathbf{CD} = \text{diag}(c_1 d_1, \dots, c_N d_N)$
3.  $\mathbf{D}^k = \text{diag}(d_1^k, \dots, d_N^k)$  для любых  $k \in \mathbb{N}$
4.  $d_n \geq 0$  для всех  $n \implies \mathbf{D}^{1/2}$  существует и равняется

$$\text{diag}(\sqrt{d_1}, \dots, \sqrt{d_N})$$

5.  $d_n \neq 0$  для всех  $n \implies \mathbf{D}$  несингулярна и

$$\mathbf{D}^{-1} = \text{diag}(d_1^{-1}, \dots, d_N^{-1})$$

Доказательства: Проверьте 1 и 2 напрямую. Части 3-5 следуют из 1 и 2.



## След, транспонирование, симметрия

**След** матрицы определяется как

$$\text{trace} \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} = \sum_{n=1}^N a_{nn}$$

**Факт.** (??) Если  $\mathbf{A}$  и  $\mathbf{B}$  являются квадратными матрицами и  $\alpha, \beta \in \mathbb{R}$ , то

$$\text{trace}(\alpha \mathbf{A} + \beta \mathbf{B}) = \alpha \text{trace}(\mathbf{A}) + \beta \text{trace}(\mathbf{B})$$

Если  $\mathbf{A}$  размера  $N \times M$  и  $\mathbf{B}$  размера  $M \times N$ , то  $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$

**Транспонирование** превращает матрицу  $\mathbf{A}$  размера  $N \times K$  в матрицу  $\mathbf{A}^T$  размера  $K \times N$ , определяемую как

$$\text{col}_n(\mathbf{A}') = \text{row}_n(\mathbf{A})$$

**Пример.** Если

$$\mathbf{A} := \begin{pmatrix} 10 & 40 \\ 20 & 50 \\ 30 & 60 \end{pmatrix}, \text{ то } \mathbf{A}' = \begin{pmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \end{pmatrix}$$

Если

$$\mathbf{B} := \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}, \text{ то } \mathbf{B}' := \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

Квадратная матрица  $A$  называется **симметричной**, если  $A^T = A$

- эквивалентно,  $a_{nk} = a_{kn}$  для всех  $k$  и  $n$

Матрицы  $A^T A$  и  $A A^T$  всегда корректно определены и симметричны

**Факт.** (??) Для согласующихся матриц  $\mathbf{A}$  и  $\mathbf{B}$  транспонирование удовлетворяет

1.  $(\mathbf{A}^T)^T = \mathbf{A}$ ,
2.  $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$ ,
3.  $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$ , и
4.  $(c\mathbf{A})^T = c\mathbf{A}^T$  для любых постоянных  $c$ .

**Факт.** (??) Для любой квадратной матрицы  $\mathbf{A}$ , верно

1.  $\text{trace}(\mathbf{A}) = \text{trace}(\mathbf{A}^T)$  и
2.  $\det(\mathbf{A}^T) = \det(\mathbf{A})$ .
3. Если  $\mathbf{A}$  несингулярна, то  $\mathbf{A}^T$  тоже несингулярна, и  $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ .

Если  $\mathbf{a}$  и  $\mathbf{b}$  — векторы размера  $N \times 1$ , умножение матриц  $\mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a}$  равняется  $\sum_{n=1}^N a_n b_n$ ,

- так же, как скалярное произведение  $\langle \mathbf{a}, \mathbf{b} \rangle$

## Собственные значения и собственные векторы

Возьмем матрицу  $\mathbf{A}$  размера  $N \times N$

Обычно  $\mathbf{A}$  отображает  $\mathbf{x}$  в какое-то произвольное новое место  $\mathbf{Ax}$

Но иногда  $\mathbf{x}$  будет только масштабироваться:

$$\mathbf{Ax} = \lambda \mathbf{x} \quad \text{для некоторого скаляра } \lambda \quad (1)$$

Если (1) выполняется и  $\mathbf{x}$  — ненулевой, то

1.  $\mathbf{x}$  называется **собственным вектором**  $\mathbf{A}$  и  $\lambda$  называется **собственным значением**
2.  $(\mathbf{x}, \lambda)$  называется **собственной парой**

Ясно, что  $(\mathbf{x}, \lambda)$  — собственная пара  $\mathbf{A} \implies (\alpha \mathbf{x}, \lambda)$  — собственная пара  $\mathbf{A}$  для любых ненулевых  $\alpha$

Пример. Пусть

$$\mathbf{A} := \begin{pmatrix} 1 & -1 \\ 3 & 5 \end{pmatrix}$$

Тогда

$$\lambda = 2 \quad \text{и} \quad \mathbf{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

формируют собственную пару, потому что  $\mathbf{x} \neq \mathbf{0}$  и

$$\mathbf{Ax} = \begin{pmatrix} 1 & -1 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \lambda \mathbf{x}$$

Пример. Рассмотрим матрицу

$$\mathbf{R} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Вызывает вращение против часовой стрелки в любой точке на  $90^\circ$

Значит ни одна точка  $\mathbf{x}$  не масштабируется

Значит не существует пары  $\lambda \in \mathbb{R}$  и  $\mathbf{x} \neq \mathbf{0}$ , такой что

$$\mathbf{R}\mathbf{x} = \lambda\mathbf{x}$$

- Другими словами, не существует вещественных собственных пар



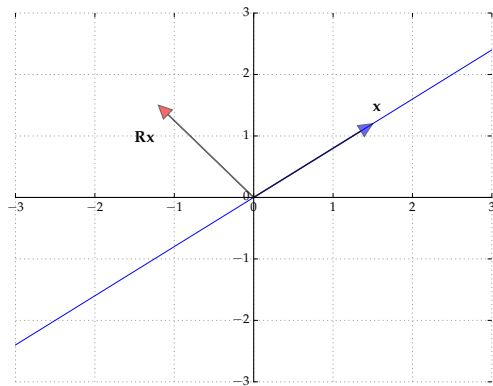


Рис.: Матрица  $R$  поворачивает точки на  $90^\circ$

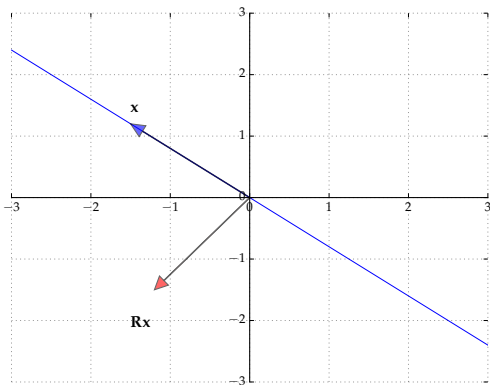


Рис.: Матрица  $R$  поворачивает точки на  $90^\circ$

Но  $\mathbf{R}\mathbf{x} = \lambda\mathbf{x}$  может выполняться, если мы допускаем комплексные числа

Пример.

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} i \\ 1 \end{pmatrix} = i \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

То есть

$$\mathbf{R}\mathbf{x} = \lambda\mathbf{x} \quad \text{для} \quad \lambda := i \quad \text{и} \quad \mathbf{x} := \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

Тогда  $(\mathbf{x}, \lambda)$  является собственной парой при условии, что мы допускаем комплексные числа

**Факт.** (??) для любой квадратной матрицы  $\mathbf{A}$

$\lambda$  является собственным значением  $\mathbf{A} \iff \det(\mathbf{A} - \lambda \mathbf{I}) = 0$

**Доказательство.** Пусть  $\mathbf{A}$  — матрица размера  $N \times N$  и  $\mathbf{I}$  — единичная матрица размера  $N \times N$

Получается

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \iff \mathbf{A} - \lambda \mathbf{I} \text{ сингулярно}$$

$$\iff \exists \mathbf{x} \neq \mathbf{0} \text{ s.t. } (\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$$

$$\iff \exists \mathbf{x} \neq \mathbf{0} \text{ s.t. } \mathbf{A}\mathbf{x} = \lambda \mathbf{x}$$

$$\iff \lambda \text{ является собственным значением } \mathbf{A}$$

Пример. В случае матрицы  $2 \times 2$ ,

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \mathbf{A} - \lambda \mathbf{I} = \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix}$$

$$\therefore \det(\mathbf{A} - \lambda \mathbf{I}) = (a - \lambda)(d - \lambda) - bc$$

$$= \lambda^2 - (a + d)\lambda + (ad - bc)$$

Значит собственные значения  $\mathbf{A}$  являются двумя корнями уравнения

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

Эквивалентно,

$$\lambda^2 - \text{trace}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0$$



**Факт.** (??) Возьмем матрицу  $\mathbf{A}$  размера  $N \times N$  с собственными значениями  $\lambda_1, \dots, \lambda_N$ . Получается

1.  $\det(\mathbf{A}) = \prod_{n=1}^N \lambda_n$
2.  $\text{trace}(\mathbf{A}) = \sum_{n=1}^N \lambda_n$
3. если  $\mathbf{A}$  симметрична, то  $\lambda_n \in \mathbb{R}$  для всех  $n$
4. если  $\mathbf{A}$  несингулярна, то  
eigenvalues of  $\mathbf{A}^{-1} = 1/\lambda_1, \dots, 1/\lambda_N$
5. если  $\mathbf{A} = \text{diag}(d_1, \dots, d_N)$ , то  $\lambda_n = d_n$  для всех  $n$

Значит матрица  $\mathbf{A}$  несингулярна  $\iff$  все ее собственные значения ненулевые





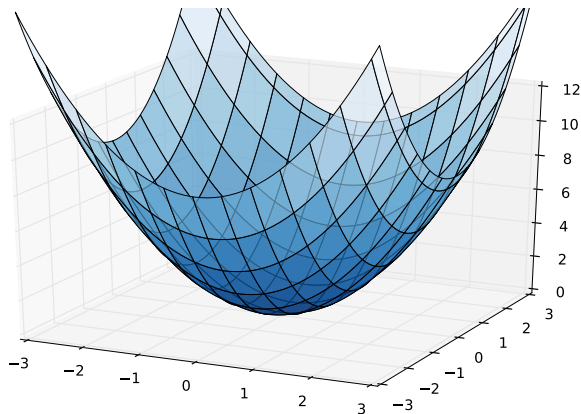


Рис.: Квадратичная функция  $Q(\mathbf{x}) = x_1^2 + x_2^2$

Внимание:

- График для  $Q(\mathbf{x}) = \|\mathbf{x}\|^2 = x_1^2 + x_2^2$  лежит всюду выше нуля

Матрица  $\mathbf{A}$  с квадратичной формой с указанным выше свойством  $Q(\mathbf{x}) \geq 0$  называется *положительно определенной*

В более общем смысле, симметричная матрица  $\mathbf{A}$  размера  $N \times N$  называется

- **неотрицательно определенной**, если  $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$  для всех  $\mathbf{x} \in \mathbb{R}^N$ ,
- **положительно определенной**, если  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  для всех  $\mathbf{x} \in \mathbb{R}^N$  with  $\mathbf{x} \neq \mathbf{0}$ ,
- **неположительно определенной**, если  $\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0$  для всех  $\mathbf{x} \in \mathbb{R}^N$ , и
- **отрицательно определенной**, если  $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$  для всех  $\mathbf{x} \in \mathbb{R}^N$  with  $\mathbf{x} \neq \mathbf{0}$ .

Если  $\mathbf{A}$  не подходит ни к одной из этих категорий, то  $\mathbf{A}$  называется **неопределенной**

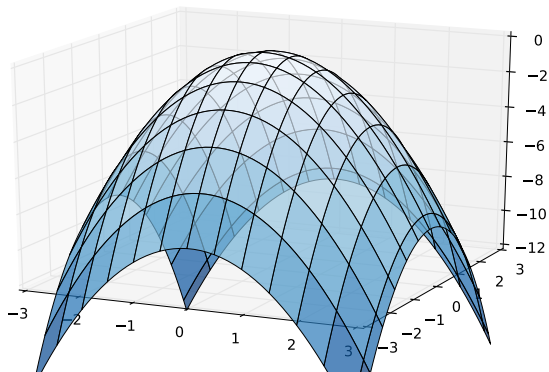


Рис.: Квадратичная функция  $Q(x) = -x_1^2 - x_2^2$

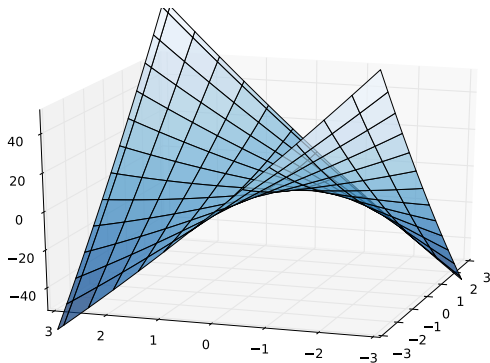


Рис.: Квадратичная функция  $Q(x) = x_1^2/2 + 8x_1x_2 + x_2^2/2$

Когда матрица  $\mathbf{A}$  диагональная:

$$\mathbf{A} = \text{diag}(d_1, \dots, d_N) \quad \text{подразумевает} \quad Q(\mathbf{x}) = d_1 x_1^2 + \dots + d_N x_N^2$$

Диагональная матрица положительно определена, тогда и только тогда, когда все диагональные элементы положительны

**Факт.** (??) Пусть  $A$  — некоторая симметричная матрица.  
Матрица  $A$

1. положительно определена тогда и только тогда, когда все ее собственные значения положительны
2. отрицательно определена тогда и только тогда, когда все ее собственные значения отрицательны

...аналогично для неположительно и неотрицательно определенных

**Факт.** (??) Если  $A$  положительно определена, то  $A$  несингулярна,  $\det A > 0$

Необходимое (но не достаточное) условие для каждого вида определенности:

**Факт.** (??) Если  $A$  положительно определена, то каждый элемент  $a_{nn}$  на главной диагонали положительный, то же самое для неотрицательной, неположительной и отрицательной.





**Теорема.** (??) Пусть  $S$  — подпространство в  $\mathbb{R}^N$ . Если  $P = \text{proj } S$ , то

$$\mathbf{P} = \mathbf{B}(\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \quad (2)$$

для каждой матрицы  $\mathbf{B}$ , такой что столбцы  $\mathbf{B}$  формируют базис  $S$

Смотрите упражнение ?? для доказательства

- Матрица  $\mathbf{M} = \mathbf{I} - \mathbf{P}$  обозначает остаточную проекцию (смотрите страницу ??)

Пример. Вспомним пример ?? на странице ??

Мы выяснили, что проекция  $\mathbf{y} \in \mathbb{R}^N$  на  $\text{span}\{\mathbf{1}\}$  — это  $\bar{y}\mathbf{1}$

Тот же результат с помощью теоремы (??):

- Так как  $\mathbf{1}$  является базисом  $\text{span}\{\mathbf{1}\}$ :

$$\mathbf{P} = \text{proj } \text{span}\{\mathbf{1}\} \implies \mathbf{P} = \mathbf{1}(\mathbf{1}^\top \mathbf{1})^{-1} \mathbf{1}^\top = \frac{1}{N} \mathbf{1} \mathbf{1}^\top$$

- Значит,  $\mathbf{P}\mathbf{y} = \bar{y}\mathbf{1}$ , как и ожидалось
- Соответствующая остаточная проекция

$$\mathbf{M}_c = \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top$$





**Факт.** (??) Пусть  $S$  является линейным подпространством в  $\mathbb{R}^N$ . Если  $\mathbf{P} = \text{proj } S$  и  $\mathbf{M}$  — остаточная проекция, то

1.  $\text{rank } \mathbf{P} = \text{trace } \mathbf{P} = \dim S$  и
2.  $\text{rank } \mathbf{M} = \text{trace } \mathbf{M} = N - \dim S$

### Доказательство.

- Ранг линейного отображения — это размерность его диапазона. Когда  $\mathbf{P} = \text{proj } S$ , Диапазон  $\mathbf{P}$  равняется  $S$
- Чтобы показать, что  $\text{trace } \mathbf{P} = \dim S$  также соблюдается, используем факт ?? —  $\text{trace } \mathbf{P} = \dim S$ ,
- Это следует из того, что  $\text{trace } \mathbf{M} = N - \dim S$ , потому что

$$\text{trace } \mathbf{M} = \text{trace}(\mathbf{I} - \mathbf{P}) = \text{trace } \mathbf{I} - \text{trace } \mathbf{P} = N - \dim S$$

## Переопределенные системы уравнений

Рассмотрим систему уравнений в виде  $\mathbf{Ax} = \mathbf{b}$ , когда:

- Матрица  $\mathbf{A}$  размера  $N \times K$  имеет полный ранг столбцов
- Вектор  $\mathbf{x}$  имеет размер  $K \times 1$
- Вектор  $\mathbf{b}$  имеет размер  $N \times 1$
- $K \leq N$

Принимая как данность  $\mathbf{A}$  и  $\mathbf{b}$ , мы ищем  $\mathbf{x} \in \mathbb{R}^K$ , такой что  $\mathbf{Ax} = \mathbf{b}$

Если  $K = N$ , то система имеет ровно одно решение

Когда  $N > K$ , система уравнений считается  
**переопределенной**:

- количество уравнений  $>$  количества неизвестных
- количество ограничений  $>$  количества степеней свободы

Возможно, не удастся найти  $\mathbf{b}$ , который удовлетворяет всем  $N$  уравнениям



73/104

Когда  $K < N$ , случай с  $\mathbf{b} \in \text{colspace } \mathbf{A}$  является “очень редким”, потому что:

- точка  $\mathbf{b}$  является произвольной точкой в  $\mathbb{R}^N$
- пространство  $\text{colspace } \mathbf{A}$  имеет размерность  $K$
- подпространства  $\mathbb{R}^N$  с размерностями  $K$  имеют “нулевую меру Лебега” – “шанс” того, что  $\mathbf{b}$  лежит в этом подпространстве, крошечный

Стандартный подход: признать, что точного решения может не существовать

Следует сосредоточиться на поиске  $\mathbf{x} \in \mathbb{R}^K$ , чтобы  $\mathbf{Ax}$  оказалось настолько близко к  $\mathbf{b}$ , насколько возможно

- близки по обычной Евклидовой норме

Задача минимизации называется **методом наименьших квадратов**

$$\hat{\mathbf{x}} := \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^K} \|\mathbf{b} - \mathbf{Ax}\| \quad (3)$$

При условии, что  $\mathbf{A}$  размера  $N \times K$  с  $K \leq N$  и  $\mathbf{b}$  размера  $N \times 1$ , мы можем использовать теорему ортогональной проекции для решения (3)

**Теорема.** (??) Если  $\mathbf{A}$  имеет полный ранг столбцов, то (3) имеет единственное решение

$$\hat{\mathbf{x}} := (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (4)$$

## Доказательство.

Пусть:

- $\mathbf{A}$  и  $\mathbf{b}$  будут как в формулировке теоремы
- $\hat{\mathbf{x}}$  как в (4) и
- $S := \text{colspace } \mathbf{A}$

При условии полного ранга столбцов, столбцы  $\mathbf{A}$  формируют базис для  $S$ . Применяем теорему ??, ортогональная проекция  $\mathbf{b}$  на  $S$ :

$$\mathbf{Pb} := \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \mathbf{A} \hat{\mathbf{x}} \quad (5)$$

Поскольку теорема об ортогональной проекции дает единственный минимизатор в терминах ближайшей точки в  $S$  к  $\mathbf{b}$ ,

$$\|\mathbf{b} - \mathbf{A} \hat{\mathbf{x}}\| < \|\mathbf{b} - \mathbf{y}\| \quad \text{для всех } \mathbf{y} \in S, \mathbf{y} \neq \mathbf{A} \hat{\mathbf{x}} \quad (6)$$

**Доказательство.** (прод.) Возьмем любой  $x \in \mathbb{R}^K$ , такой что  $x \neq \hat{x}$

Получается  $Ax \in S$

К тому же, так как  $x \neq \hat{x}$  и  $A$  имеют полный ранг столбцов, должно быть  $Ax \neq A\hat{x}$  (упр. ??)

Значит

$$\|b - A\hat{x}\| < \|b - Ax\| \quad \text{для всех } x \in \mathbb{R}^K, x \neq \hat{x}$$

Другими словами,  $\hat{x}$  является единственным решением (3)

В (4) матрица  $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$  называется **псевдообратной** матрицы  $\mathbf{A}$

Если  $K = N$ , то решение методом наименьших квадратов  $\hat{\mathbf{x}}$  в (4) сокращается до:

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b}$$

Что произойдет, если столбцы  $\mathbf{A}$  не будут линейно независимыми?

- множество  $\text{colspace } \mathbf{A}$  все еще линейное подпространство и теорема ортогональной проекции все еще дает нам ближайшую точку  $\mathbf{Pb}$  к  $\mathbf{b}$  в  $\text{colspace } \mathbf{A}$
- так как  $\mathbf{Pb} \in \text{colspace } \mathbf{A}$ , все еще существует вектор  $\hat{\mathbf{x}}$ , такой что  $\mathbf{Pb} = \mathbf{A}\hat{\mathbf{x}}$
- но таких векторов существует бесконечное множество

Смотрите упражнение ??



## QR-разложение

**QR-разложение** данной матрицы **A** — это произведение вида **QR**

- первая матрица имеет ортонормированные столбцы и
- вторая является верхней треугольной

Приложения включают задачи наименьших квадратов и вычисление собственных значений



Возьмем разложение  $\mathbf{A} = \mathbf{QR}$ , решение методом наименьших квадратов  $\hat{\mathbf{x}}$ , определенное в (4), может также записываться как:

$$\hat{\mathbf{x}} = \mathbf{R}^{-1} \mathbf{Q}^T \mathbf{b}$$

Смотрите упр. ??

Умножение на  $\mathbf{R}$ :

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^T \mathbf{b}$$

# Диагонализация и спектральная теория

Если  $f: A \rightarrow A$  и  $g: B \rightarrow B$ , то говорят, что  $g$  **топологически сопряжен** с  $f$  везде, где существует непрерывная биекция  $\tau: B \rightarrow A$ , такая что

$$f = \tau \circ g \circ \tau^{-1}$$

Может быть полезно, если  $g$  проще  $f$

Говорят, что квадратная матрица **A** **подобна** матрице **B**, если существует обратимая матрица **P**, такая что  $\mathbf{A} = \mathbf{PBP}^{-1}$

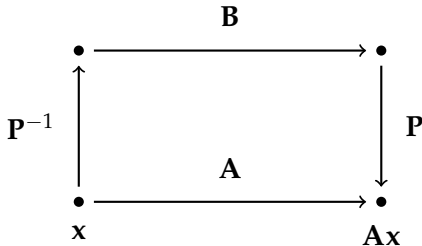


Рис.: **A** подобна **B**

Если  $A$  подобна диагональной матрице, то  $A$  называется  
**диагонализируемой**

Нас интересует подобие простым матрицам, диагональные  
матрицы — самый простой вид

**Факт.** (??) Если  $\mathbf{A}$  подобна  $\mathbf{B}$ , то  $\mathbf{A}^t$  подобна  $\mathbf{B}^t$  для всех  $t \in \mathbb{N}$

**Пример.** Мы хотим вычислить  $\mathbf{A}^t$  для некоторых данных  $t \in \mathbb{N}$

Если  $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$  для некоторых  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$ , то по факту ?? и факту ??, получается

$$\mathbf{A}^t = \mathbf{P} \text{diag}(\lambda_1^t, \dots, \lambda_N^t) \mathbf{P}^{-1}$$

## Диагонализация и собственные значения

**Факт.** (??) Если  $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$  для некоторых  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$ , то  $(\text{col}_n \mathbf{P}, \lambda_n)$  является собственной парой  $\mathbf{A}$  для любого  $n$

**Доказательство.** Заметим, что  $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$  значит, что  $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{\Lambda}$

Equating the  $n$ th column on each side gives Рассмотрим столбец  $n$  с каждой стороны:

$$\mathbf{A}\mathbf{p}_n = \lambda_n \mathbf{p}_n$$

Где  $\mathbf{p}_n := \text{col}_n \mathbf{P}$

Заметим, что  $\mathbf{p}_n$  — ненулевой вектор, так как  $\mathbf{P}$  обратима



Но когда  $\mathbf{A}$  диагонализируема?

**Факт.** (3.3.7) Матрица  $\mathbf{A}$  размера  $N \times N$  диагонализируема тогда и только тогда, когда она имеет  $N$  линейно независимых собственных векторов

В некоторых случаях, мы можем получить еще более простое разложение матрицы, если матрица  $\mathbf{P}$  имеет ортогональные столбцы

Такие матрицы называются **ортогональными матрицами**

**Факт.** (??) Если  $\mathbf{Q}$  и  $\mathbf{P}$  — ортогональные матрицы размера  $N \times N$ , то

1.  $\mathbf{Q}^T$  ортогональна и  $\mathbf{Q}^{-1} = \mathbf{Q}^T$ ,
2.  $\mathbf{QP}$  ортогонально, и
3.  $\det \mathbf{Q} \in \{-1, 1\}$ .

Если  $A = Q\Lambda Q^{-1}$  и  $Q$  имеют ортогональные столбцы, то

$$A = Q\Lambda Q^T$$

Ясно, что  $A$  должна быть симметричной. Следующая теорема показывает, что симметрия  $A$  также достаточна

**Теорема.** (??) Если  $A$  симметрична, то  $A$  может быть диагонализирована как  $A = Q\Lambda Q^T$ , где  $Q$  — ортогональная матрица и  $\Lambda$  — диагональная матрица, сформированная из собственных значений  $A$

Вышеупомянутая теорема была версией **теоремы о спектральном разложении**

$\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$  называется **симметричным разложением на собственные значения** матрицы  $\mathbf{A}$  – воздействие  $\mathbf{A}$  на вектор  $\mathbf{x}$  размера  $N \times 1$ :

$$\mathbf{A}\mathbf{x} = \sum_{n=1}^N \lambda_n (\mathbf{u}_n^T \mathbf{x}) \mathbf{u}_n$$

где  $\lambda_n$  является  $n$ -ым собственным значением  $\mathbf{A}$  и  $\mathbf{u}_n = \text{col}_n \mathbf{Q}$

Сравните с  $\mathbf{x} = \sum_{n=1}^N (\mathbf{u}_n^T \mathbf{x}) \mathbf{u}_n$

**Факт.** (??) Если  $\mathbf{A}$  неотрицательно определена, то  $\sqrt{\mathbf{A}}$  существует и равняется  $\mathbf{Q}\sqrt{\mathbf{\Lambda}}\mathbf{Q}^T$ . Матрица  $\sqrt{\mathbf{\Lambda}}$  равняется  $\text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_N})$

**Факт.** (??) Если  $\mathbf{A}$  положительно определена, то существует несингулярная, верхняя треугольная матрица  $\mathbf{R}$ , такая что  $\mathbf{A} = \mathbf{R}^T \mathbf{R}$

Такое разложение называется **разложением Холецкого**

**Доказательство.** (Разложение Холецкого) Мы можем написать:

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top = \mathbf{Q}\sqrt{\mathbf{\Lambda}}\sqrt{\mathbf{\Lambda}}\mathbf{Q}^\top = (\sqrt{\mathbf{\Lambda}}\mathbf{Q}^\top)^\top\sqrt{\mathbf{\Lambda}}\mathbf{Q}^\top$$

Затем применим QR-разложение к  $\sqrt{\Lambda}\mathbf{Q}^T$ :

$$\sqrt{\Lambda}\mathbf{Q}^\top = \tilde{\mathbf{Q}}\mathbf{R}$$

где  $\mathbf{R}$  несингулярная и верхняя треугольная, и  $\tilde{\mathbf{Q}}$  имеет ортонормированные столбцы

Так как столбцы  $\tilde{\mathbf{Q}}$  ортонормированные,

$$\mathbf{A} = (\tilde{\mathbf{Q}}\mathbf{R})^\top \tilde{\mathbf{Q}}\mathbf{R} = \mathbf{R}^\top \tilde{\mathbf{Q}}^\top \tilde{\mathbf{Q}}\mathbf{R} = \mathbf{R}^\top \mathbf{R}$$

# Нормы и непрерывность

Возьмем векторную последовательность  $\{\mathbf{x}_n\}$  в  $\mathbb{R}^K$  и любую точку  $\mathbf{x} \in \mathbb{R}^K$ . Говорят, что  $\{\mathbf{x}_n\}$  **сходится** к  $\mathbf{x}$  и записывается  $\mathbf{x}_n \rightarrow \mathbf{x}$ , если для любых  $\epsilon > 0$ , существует  $N \in \mathbb{N}$ , такое что  $\|\mathbf{x}_n - \mathbf{x}\| < \epsilon$  при  $n \geq N$

Эквивалентно, последовательность действительных чисел  $z_n := \|x_n - x\|$  сходится к нулю в  $\mathbb{R}$  при  $n \rightarrow \infty$

**Факт.** (3.3.11) Имеют место следующие результаты:

1. если  $\mathbf{x}_n \rightarrow \mathbf{x}$  и  $\mathbf{y}_n \rightarrow \mathbf{y}$ , то  $\mathbf{x}_n + \mathbf{y}_n \rightarrow \mathbf{x} + \mathbf{y}$ .
2. если  $\mathbf{x}_n \rightarrow \mathbf{x}$  и  $\alpha \in \mathbb{R}$ , то  $\alpha \mathbf{x}_n \rightarrow \alpha \mathbf{x}$ .
3.  $\mathbf{x}_n \rightarrow \mathbf{x}$  тогда и только тогда, когда  $\mathbf{a}^T \mathbf{x}_n \rightarrow \mathbf{a}^T \mathbf{x}$  для всех  $\mathbf{a} \in \mathbb{R}^K$ .



Мы хотим распространить понятие сходимости на матрицы

**Норма матрицы**  $\mathbf{A}$  размера  $N \times K$ :

$$\|\mathbf{A}\| := \max \left\{ \|\mathbf{Ax}\| : \mathbf{x} \in \mathbb{R}^K, \|\mathbf{x}\| = 1 \right\} \quad (7)$$

Значение нормы матрицы в общем случае найти непросто

Однако, матричная норма ведет себя как векторная норма

**Факт.** (3.3.12) Для любых согласующихся матриц  $\mathbf{A}$  и  $\mathbf{B}$ ,  
матричная норма удовлетворяет

1.  $\|\mathbf{A}\| \geq 0$  и  $\|\mathbf{A}\| = 0$  тогда и только тогда, когда все записи  $\mathbf{A}$  нулевые,
2.  $\|\alpha \mathbf{A}\| = |\alpha| \|\mathbf{A}\|$  для любых скаляров  $\alpha$ ,
3.  $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ , и
4.  $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$ .

**Факт.** (??) Для любых матриц  $\mathbf{A}$  размера  $J \times K$  с элементами  $a_{jk}$ , выходит

$$\|\mathbf{A}\| \leq \sqrt{JK} \max_{jk} |a_{jk}|$$

если каждый элемент  $\mathbf{A}$  близок к нулю, то  $\|\mathbf{A}\|$  тоже близка к нулю

## Ряд Неймана

В дальнейшем мы изучаем динамические системы вида

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{b}$$

Существует ли “стационарный” вектор  $x \in \mathbb{R}^N$  в том смысле, что  $x_t = x$  значит, что  $x_{t+1} = x$ ?

Мы ищем  $x \in \mathbb{R}^N$ , являющийся решением системы уравнений

$$\mathbf{x} = \mathbf{Ax} + \mathbf{b} \quad (\mathbf{A} \text{ is } N \times N \text{ и } \mathbf{b} \text{ is } N \times 1) \quad (8)$$



Когда выполняется условие леммы о рядах Неймана, (8) имеет единственное решение

$$\bar{\mathbf{x}} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} = \sum_{i=0}^{\infty} \mathbf{A}^i \mathbf{b}$$

Чтобы проверить это решение, используем **спектральный радиус**  $\mathbf{A}$ :

$$\rho(\mathbf{A}) := \max\{|\lambda| : \lambda \text{ является собственным значением } \mathbf{A}\}$$

$|\lambda|$  — это **модуль** возможно комплексного числа  $\lambda$



