# Линейная алгебра

## Матрицы и векторы. Первичное знакомство, основные операции.

Глеб Карпов

МНаД ФКН ВШЭ

# Course overview

# Matrix

- **Definition**

  Matrix - arranged array of numbers in a form of $n$ rows and $m$ columns.

$$A_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

# Matrix

- **Definition**

  Matrix - arranged array of numbers in a form of $n$ rows and $m$ columns.

  $$A_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

- We denote it usually as $A_{n\times m}$ or to emphasize the nature of numbers presented in a matrix we write $A \in \mathbb{R}^{n\times m}$.

# Matrix

- **Definition**

  Matrix - arranged array of numbers in a form of $n$ rows and $m$ columns.

  $$A_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

- We denote it usually as $A_{n\times m}$ or to emphasize the nature of numbers presented in a matrix we write $A \in \mathbb{R}^{n\times m}$.
- If $n = m$ we call matrix square, if $n \neq m$ we call it rectangular

# Basic operations: matrix addition

- **Definition**

  **Matrix addition** is only possible for matrices of the same size. The result is obtained by adding corresponding elements. If $A, B \in \mathbb{R}^{n \times m}$, then $C = A + B$, where $c_{ij} = a_{ij} + b_{ij}$

**Example:**

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

# Basic operatons: scalar multiplication

- **Definition**

  **Scalar multiplication** - each element of the matrix is multiplied by the given number. If $A \in \mathbb{R}^{n \times m}$ and $\alpha \in \mathbb{R}$, then $C = \alpha A$, where $c_{ij} = \alpha \cdot a_{ij}$

**Example:**

$$3 \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 \cdot 1 & 3 \cdot 2 \\ 3 \cdot 3 & 3 \cdot 4 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$$

**Another example - Linear combination:**

$$2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

# Vector

- Basically we start from thinking about vector as a

# Matrix-by-vector multiplication (matvec)

## General perception

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,m-1} & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2,m-1} & a_{2m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,m-1} & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$A \qquad \mathbf{x} \qquad \mathbf{y}$$

$$y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,m-1}x_{m-1} + a_{1m}x_m$$

$$\boxed{y_j = \sum_{k=1}^{m} a_{jk}x_k}$$

# Matrix-by-vector multiplication (matvec)

## Complexity and short glance at parallel computing

**Recall the general formula:**

$$\boxed{y_j = \sum_{k=1}^{m} a_{jk} x_k, \quad j = 1, 2, \dots, n}$$

**Operations analysis:**

- For computing one element $y_j$: $m$ multiplications(each $a_{jk} \cdot x_k$), $m-1$ additions (summing $m$ products)
- For the entire vector $\mathbf{y}$ (n elements): $n \cdot (2m-1)$ total operations.

**Time complexity:**

- For square matrix $n \times n$: $\mathcal{O}(2n^2 - n) = \mathcal{O}(n^2)$
- For rectangular matrix $n \times m$: $\mathcal{O}(nm)$

Natural parallelism

Computing each element $y_j$ is **independent** of other elements!

- **Row-wise parallelization:** each processor computes its own $y_j$

# Matrix-by-vector multiplication (matvec)

Guru perception

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\mathbf{y}}$$

# Matrix-by-vector multiplication (matvec)

Guru perception

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\mathbf{y}}$$

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4$$

# Matrix-by-vector multiplication (matvec)

## Guru perception

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\mathbf{y}}$$

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4$$
$$y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4$$

# Matrix-by-vector multiplication (matvec)

## Guru perception

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\mathbf{y}}$$

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 \\ y_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 \end{aligned}$$

# Matrix-by-vector multiplication (matvec)

## Guru perception

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\mathbf{y}}$$

$$\mathbf{y} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} + x_4 \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix}$$

# Matrix-by-vector product

## Visual comparison

$$\underbrace{\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 7 \\ 8 \end{bmatrix}}_{\mathbf{y}}$$

## Traditional View: Row-wise

**Calculations:**

- $y_1 = 3 \cdot 2 + 1 \cdot 1 = 7$
- $y_2 = 2 \cdot 2 + 4 \cdot 1 = 8$

## Guru View: Column Linear Combination

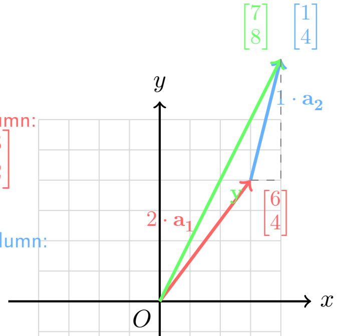
**Linear combination:**

First column:

$\mathbf{a_1} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

Second column:

$\mathbf{a_2} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$



# Matrix-by-matrix multiplication

## General perception

$$
\underset{A}{
\begin{array}{r}
\\ \\ i\text{-th row} \rightarrow \\ \\ \\
\end{array}
\left[\begin{array}{cccc}
a_{11} & a_{12} & \cdots & a_{1k} \\
\vdots & \vdots & \ddots & \vdots \\
a_{i1} & a_{i2} & \cdots & a_{ik} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nk}
\end{array}\right]}
\;
\underset{B}{
\begin{array}{c}
j\text{-th column} \\
\left[\begin{array}{ccccc}
b_{11} & \cdots & b_{1j} & \cdots & b_{1m} \\
b_{21} & \cdots & b_{2j} & \cdots & b_{2m} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
b_{k1} & \cdots & b_{kj} & \cdots & b_{km}
\end{array}\right]
\end{array}}
=
\underset{C}{
\left[\begin{array}{ccccc}
c_{11} & \cdots & c_{1j} & \cdots & c_{1m} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
c_{i1} & \cdots & c_{ij} & \cdots & c_{im} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
c_{n1} & \cdots & c_{nj} & \cdots & c_{nm}
\end{array}\right]}
$$

$$c_{ij} = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + \cdots + a_{ik} \cdot b_{kj}$$

$$\boxed{c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}}$$

# Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$

# Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)

# Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$, то $x = 0$

# Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$, то $x = 0$

# Нормы

Норма - это **количественная мера малости вектора** и обычно обозначается как $\|x\|$.

Норма должна удовлетворять определенным свойствам:

1. $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)
3. Если $\|x\| = 0$, то $x = 0$

Расстояние между двумя векторами определяется как

$$d(x, y) = \|x - y\|.$$

Наиболее широко используемой нормой является **Евклидова норма**:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2},$$

которая соответствует расстоянию в нашей реальной жизни. Если векторы имеют комплексные элементы, мы используем их модуль. Евклидова норма, или $2$-норма, является подклассом важного класса $p$-норм:

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p}.$$

# $p$-норма вектора

Существуют два очень важных частных случая. Бесконечность-норма, или норма Чебышева, определяется как максимальное абсолютное значение элемента вектора:

$$\|x\|_\infty = \max_i |x_i|$$

# $p$-норма вектора

Существуют два очень важных частных случая. Бесконечность-норма, или норма Чебышева, определяется как максимальное абсолютное значение элемента вектора:

$$\|x\|_\infty = \max_i |x_i|$$

$l_1$ норма (или **манхэттенское расстояние**) определяется как сумма модулей элементов вектора $x$:

$$\|x\|_1 = \sum_i |x_i|$$

# Скалярное произведение

## Определение

**Definition**

**Скалярное произведение** двух векторов $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ — это число, вычисляемое по формуле:

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

**Геометрический смысл:**

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| \cdot |\mathbf{v}| \cdot \cos\theta$$

где $\theta$ — угол между векторами $\mathbf{u}$ и $\mathbf{v}$.

# Обозначения скалярного произведения

## Различные способы записи

### 1. Через транспонирование

$$\mathbf{u}^T\mathbf{v}$$

**Матричная форма:**

$$\mathbf{u}^T\mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

**Результат:**

$$= u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

### 2. Через угловые скобки

$$\langle \mathbf{u}, \mathbf{v} \rangle$$

**Альтернативная запись:** - $\mathbf{u} \cdot \mathbf{v}$ — точечное произведение

**Обозначения эквивалентны:**

$$\mathbf{u}^T\mathbf{v} = \langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$$

# Основные свойства скалярного произведения

## Алгебраические свойства

1. **Коммутативность:**

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

## Геометрические свойства

# Основные свойства скалярного произведения

## Алгебраические свойства

1. **Коммутативность:**

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

2. **Линейность по первому аргументу:**

$$\langle \alpha \mathbf{u} + \beta \mathbf{w}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle$$

## Геометрические свойства

# Основные свойства скалярного произведения

## Алгебраические свойства

1. **Коммутативность:**

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

2. **Линейность по первому аргументу:**

$$\langle \alpha \mathbf{u} + \beta \mathbf{w}, \mathbf{v} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle$$

## Геометрические свойства

3. **Связь с длиной вектора:**

$$|\mathbf{u}| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\mathbf{u}^T \mathbf{u}}$$

# Основные свойства скалярного произведения

## Алгебраические свойства

1. **Коммутативность:**

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

2. **Линейность по первому аргументу:**

$$\langle \alpha\mathbf{u} + \beta\mathbf{w}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{w}, \mathbf{v} \rangle$$

## Геометрические свойства

3. **Связь с длиной вектора:**

$$|\mathbf{u}| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\mathbf{u}^T\mathbf{u}}$$

4. **Ортогональность:**

$$\mathbf{u} \perp \mathbf{v} \Leftrightarrow \langle \mathbf{u}, \mathbf{v} \rangle = 0$$