Линейная алгебра

Матрицы и векторы. Первичное знакомство, основные операции.

Глеб Карпов

Прикладные нейросетевые технологии, ФКН ВШЭ

Матрица

- **Definition**

  Матрица — упорядоченный массив чисел в виде n строк и m столбцов.

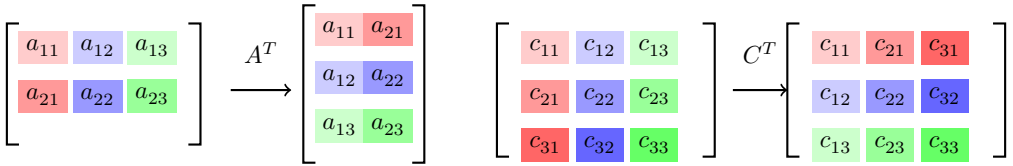
$$A_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

- Обычно обозначаем как $A_{n\times m}$ или, чтобы подчеркнуть природу чисел в матрице, пишем $A \in \mathbb{R}^{n\times m}$.
- Если $n = m$, то матрицу называют *квадратной*, если $n \neq m$ — *прямоугольной*

Основные операции: транспонирование

Definition

Транспонирование матрицы — это операция, при которой строки и столбцы меняются местами. Если $A \in \mathbb{R}^{n \times m}$, то $B = A^T$, $B \in \mathbb{R}^{m \times n}$, где $b_{ij} = a_{ji}$



$A_{2\times3}$ $B_{3\times2} = A^T$ $C_{3\times3}$ $D_{3\times3} = C^T$

Основные операции: сложение матриц

- **Definition**

  Сложение матриц возможно только для матриц одинакового размера. Результат получается сложением соответствующих элементов. Если $A, B \in \mathbb{R}^{n \times m}$, то $C = A + B$, где $c_{ij} = a_{ij} + b_{ij}$

Пример:

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

Основные операции: умножение на скаляр

-

> **Definition**
>
> **Умножение на скаляр** — каждый элемент матрицы умножается на заданное число. Если $A \in \mathbb{R}^{n \times m}$ и $\alpha \in \mathbb{R}$, то $C = \alpha A$, где $c_{ij} = \alpha \cdot a_{ij}$

Пример:

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

Комбинация операций:

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

Вектор

В самом простом представлении будем трактовать **вектор** как частный случай матрицы:

Вектор-столбец

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

Размерность: $n \times 1$ (матрица с одним столбцом)

Вектор-строка

$$\mathbf{y}^T = \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix} \in \mathbb{R}^{1 \times m}$$

Размерность: $1 \times m$ (матрица с одной строкой)

Обозначения

- **Векторы:** обычно обозначаются строчными буквами x, v или $\mathbf{u}$
- **Матрицы:** обычно обозначаются заглавными буквами A, B, C
- **По умолчанию:** вектор считается **вектором-столбцом**
- **Транспонирование:** $\mathbf{x}^\top$ превращает столбец в строку

Умножение матрицы на вектор (matvec)

Обычный подход

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

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

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

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

Умножение матрицы на вектор (matvec)

Вычислительная сложность и немного о параллельных вычислениях

Вспомним общую формулу:

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

Анализ операций:

- Для вычисления одного элемента y_j: m умножений (каждое $a_{jk} \cdot x_k$), $m - 1$ сложений (суммирование m произведений)
- Для всего вектора $\mathbf{y}$ (n элементов): $n \cdot (2m - 1)$ операций всего.

Временная сложность:

- Для квадратной матрицы $n \times n$: $\mathcal{O}(2n^2 - n) = \mathcal{O}(n^2)$
- Для прямоугольной матрицы $n \times m$: $\mathcal{O}(nm)$

> Естественный параллелизм
>
> Вычисление каждого элемента y_j **независимо** от других элементов!

- **Построчная параллелизация:** отдельный процессор может вычислять свой y_j

Умножение матрицы на вектор (matvec)

Просветленный подход :)

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

Умножение матрицы на вектор (matvec)

Просветленный подход :)

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

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

Умножение матрицы на вектор (matvec)

Просветленный подход :)

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

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

Умножение матрицы на вектор (matvec)

Просветленный подход :)

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

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

Умножение матрицы на вектор (matvec)

Просветленный подход :)

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

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

Умножение матрицы на вектор

Визуальное сравнение

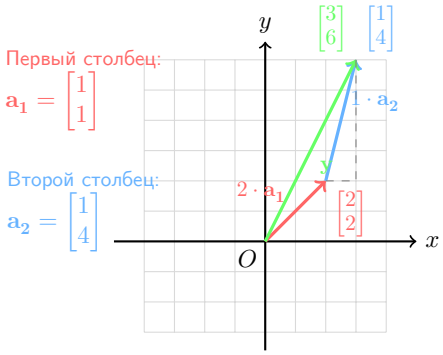
$$\underset{A}{\begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}} \quad \underset{\mathbf{x}}{\begin{bmatrix} 2 \\ 1 \end{bmatrix}} = \underset{\mathbf{y}}{\begin{bmatrix} 3 \\ 6 \end{bmatrix}}$$

Строка на столбец

Вычисления:

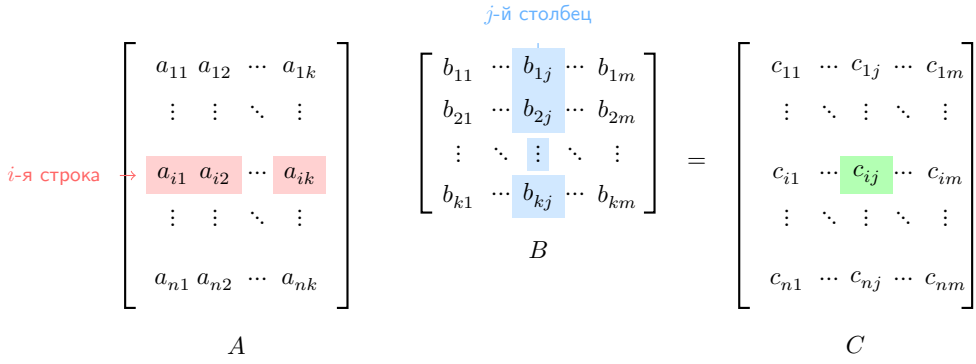
- $y_1 = 1 \cdot 2 + 1 \cdot 1 = 3$
- $y_2 = 1 \cdot 2 + 4 \cdot 1 = 6$

Комбинация столбцов



Умножение матриц (matmul, GEneral Matrix Multiplication - GEMM)

Обычный подход



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

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

Умножение матриц (matmul, GEneral Matrix Multiplication - GEMM)

Столбцовый подход

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} \quad \begin{bmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1m} \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & b_{k3} & \cdots & b_{km} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{i1} & c_{i2} & c_{i3} & \cdots & c_{im} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \cdots & c_{nm} \end{bmatrix}$$

A B C

Каждый столбец $C = A\times$ соответствующий столбец B

$$\mathbf{C_j} = A\mathbf{B_j}, \quad j = 1, 2, \ldots, m$$

Умножение матриц (matmul, GEneral Matrix Multiplication - GEMM)

Свойства матричного умножения

1. **Ассоциативность:**

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

2. **Дистрибутивность:**

$$A(B + C) = AB + AC$$
$$(A + B)C = AC + BC$$

3. **Умножение на скаляр:**

$$\alpha(AB) = (\alpha A)B = A(\alpha B)$$

Умножение матрицы на матрицу (matmul, GEMM)

Важные ограничения

4. **Некоммутативность:**

$$AB \neq BA \quad \text{(в общем случае)}$$

Пример: Для матриц 2×2:

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

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

5. **Транспонирование произведения:**

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

> **Note**
>
> Обратите внимание на **обратный порядок** матриц при транспонировании!

Умножение матриц (matmul, GEneral Matrix Multiplication - GEMM)

Вычислительная сложность и параллелизация

Вспомним общую формулу:

$$c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m$$

Анализ операций для матриц $A_{n \times k}$ и $B_{k \times m}$:

- Для вычисления одного элемента c_{ij}: k умножений, $k - 1$ сложений
- Общее количество операций: $n \times m \times (2k - 1)$

Временная сложность:

- Для квадратных матриц $n \times n$: $\mathcal{O}(2n^3 - n^2) = \mathcal{O}(n^3)$
- Для прямоугольных матриц: $\mathcal{O}(nmk)$

Естественный параллелизм

- **По элементам:** каждый c_{ij} вычисляется независимо
- **По строкам:** каждый процессор обрабатывает свои строки матрицы A
- **По столбцам:** каждый процессор обрабатывает свои столбцы матрицы B
- **Блочный подход:** разделение матриц на блоки для эффективного использования кэша

Пример. Простая, но важная идея о матричных вычислениях.

Предположим, у вас есть следующее выражение

$$b = A_1 A_2 A_3 x,$$

где $A_1, A_2, A_3 \in \mathbb{R}^{3\times 3}$ - случайные квадратные плотные (полностью заполненные числами) матрицы, и $x \in \mathbb{R}^3$ - вектор. Вам нужно вычислить b.

Какой способ лучше всего использовать?

1. $A_1 A_2 A_3 x$ (слева направо)
2. $(A_1 (A_2 (A_3 x)))$ (справа налево)
3. Не имеет значения
4. Результаты первых двух вариантов не будут одинаковыми.