

Теоретический минимум для ML

Матрицы. Основные понятия и операции.

Глеб Карпов

OTUS ML Basic

Основные разделы математики для ML и DA

- Линейная алгебра: язык оперирования с данными, предоставляет нам инструменты для преобразования данных с целью упрощения их последующего анализа, инструменты для сжатия данных и более компактного хранения.

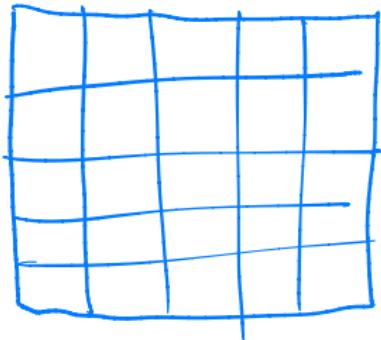
Основные разделы математики для ML и DA

- Линейная алгебра: язык оперирования с данными, предоставляет нам инструменты для преобразования данных с целью упрощения их последующего анализа, инструменты для сжатия данных и более компактного хранения.
- Теория вероятностей и статистика: выявление закономерностей в данных, выявление общих характеристик из ограниченного набора данных, важные инструменты для аналитики данных.

Основные разделы математики для ML и DA

- Линейная алгебра: язык оперирования с данными, предоставляет нам инструменты для преобразования данных с целью упрощения их последующего анализа, инструменты для сжатия данных и более компактного хранения.
- Теория вероятностей и статистика: выявление закономерностей в данных, выявление общих характеристик из ограниченного набора данных, важные инструменты для аналитики данных.
- Математический анализ (оптимизация): главные концепции для обучения моделей, знание которых позволяет более гибко настраивать процессы обучения. Обычно спрятан ``под капотом'' у известных ML-пакетов, и на первых этапах может не очень активно требоваться.

Матрица



$n \times m$ чисел

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nm} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} A \in \mathbb{N}^{2 \times 2}$$

$$\begin{pmatrix} \pi & \sin(60) \\ 0 & -\ln_{10} 30 \end{pmatrix} B \in \mathbb{R}^{2 \times 2}$$

$$A_{n \times m} \quad n \times m$$
$$A \in \mathbb{R}^{n \times m}$$

$$D = \begin{pmatrix} 0 & 2+i \\ 2-i & 0+4i \end{pmatrix};$$
$$D \in \mathbb{C}^{2 \times 2}$$

Матрица

$A \in \mathbb{R}^{n \times n}$; квадратная



$A \in \mathbb{R}^{n \times m}$, $n > m$; прямоуг.



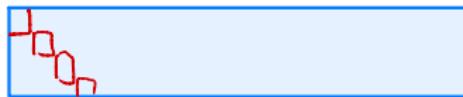
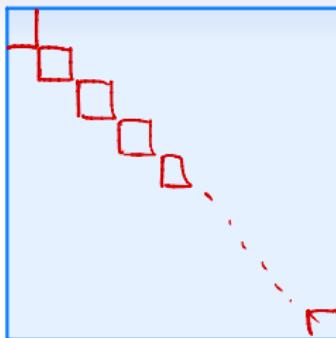
$A \in \mathbb{R}^{n \times m}$, $m > n$, Прямоуг



Матрица

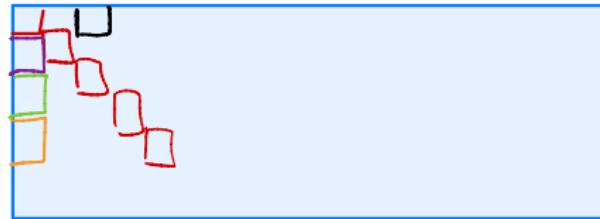
$$\{a_{ij} : i=j\}$$

Гл. диагональ



Матрица

$$B = A^T$$



$$B = A^T$$

$$a_{12} \rightarrow b_{21}$$

$$a_{ii} \rightarrow b_{ii}$$

$$a_{13} \rightarrow b_{31}$$

$$a_{14} \rightarrow b_{41}$$

$$a_{31} \rightarrow b_{13}$$

$$b_{ij} = a_{ji}$$

$$(A^T)^T = A$$

A

Операции:

Матрица

1. + $A \in \mathbb{R}^{n \times m}$ $\exists C = A + B$
 $B \in \mathbb{R}^{n \times m}$ $\forall i, j \quad c_{ij} = a_{ij} + b_{ij}$

2. умножение на скаляр

$$A \in \mathbb{R}^{n \times m} \quad \exists D = d \cdot A$$

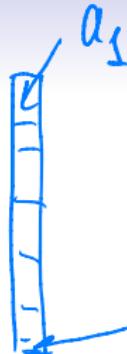
$$d \in \mathbb{R} \quad \forall i, j \quad d_{ij} = d \cdot a_{ij}$$

Ex.

$$2 \cdot \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix} - 4 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 8 & 2 \end{pmatrix} + \begin{pmatrix} -4 & -8 \\ -12 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -8 \\ -4 & 2 \end{pmatrix}$$

$a \in \mathbb{R}^{n \times 1}$; $a \in \mathbb{R}^n$

Вектор



1. + $a \in \mathbb{R}^n$
 $b \in \mathbb{R}^n$ 2. $c = a + b$; $c \in \mathbb{R}^n$

2. $\times d$ $d \in \mathbb{R}$
 $a \in \mathbb{R}^n$ $d \cdot a = d \cdot a$; $d \in \mathbb{R}^n$

$$a^T = \underbrace{\begin{array}{c} a_1 \\ \vdots \\ a_n \end{array}}_{\text{Horizontal vector}}$$

Умножение матрицы на вектор

"мат век"

$$A \in \mathbb{R}^{n \times m}$$

$$x \in \mathbb{R}^m$$

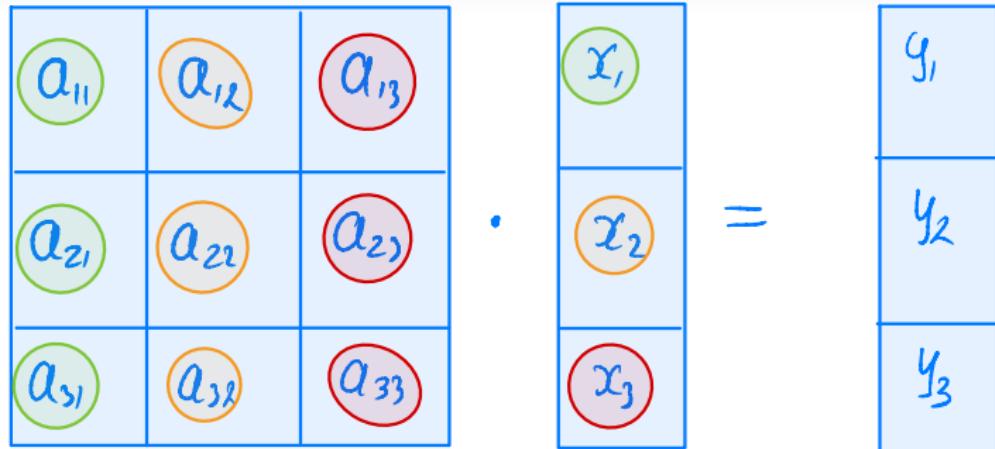
$$\exists y \in \mathbb{R}^n : y = A \cdot x$$

$$y_i = a_{i1} \cdot x_1 + a_{i2} \cdot x_2 + \dots = \sum_{j=1}^m a_{ij} \cdot x_j$$

$$N_{\text{op}} = (m_{\text{mult}} + (m-1)_{\text{add}}) \cdot n$$

$$= (mn)_{\text{mult}} + n \cdot (m-1)_{\text{add}}$$

Умножение матрицы на вектор



$$y_1 = a_{11} \cdot x_1 + a_{12} \cdot x_2 + a_{13} \cdot x_3$$

$$y = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

Умножение матрицы на матрицу

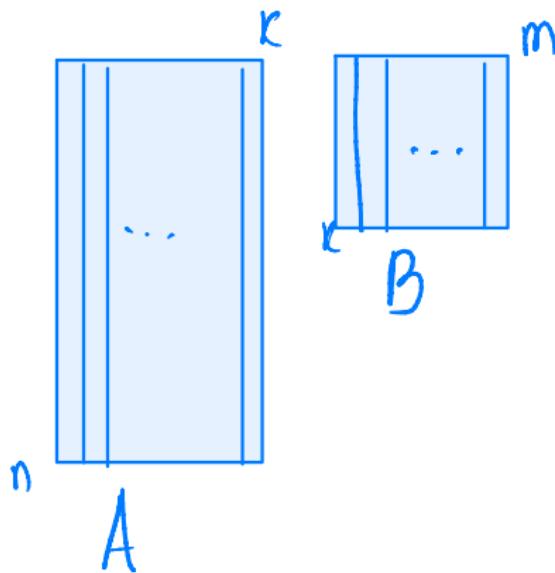
$$A \in \mathbb{R}^{n \times k}$$

$$B \in \mathbb{R}^{k \times m}$$

$$\Rightarrow \exists C \in \mathbb{R}^{n \times m}$$

$$; C = AB$$

$$c_i = A \cdot B_i$$



$$N_{op} = \\ = m(nk + n(k-l))$$

Умножение матрицы на матрицу

$$AB \neq BA$$

$$A \in \mathbb{R}^{n \times m}$$

$$A \cdot B \quad \checkmark \quad (n \times m) \quad (m \times n) = C_{n \times n}$$

$$B \in \mathbb{R}^{m \times n}$$

$$B \cdot A \quad \checkmark \quad (m \times n) \quad (n \times m) = D_{m \times m}$$

Произведения векторов

1 Скалярное (inner product)

$$a \in \mathbb{R}^n \quad \langle a, b \rangle = a_1 \cdot b_1 + a_2 \cdot b_2 + \dots + a_n \cdot b_n = \\ b \in \mathbb{R}^n \quad = \sum a_i \cdot b_i;$$

$$\boxed{a_1 \rightarrow a_n} \quad \begin{array}{c} b_1 \\ \downarrow \\ b_n \end{array} = a_1 \cdot b_1 + \dots + a_n \cdot b_n = \sum a_i \cdot b_i; \\ = a^T b$$

$$\langle a, b \rangle = \langle b, a \rangle$$
$$a^T b \underset{\text{||}}{=} b^T a$$

$$\langle a, b \rangle = \|a\| \cdot \|b\| \cdot \cos(a, b)$$

Произведения векторов

2. outer product

$$a \in \mathbb{R}^n$$
$$b \in \mathbb{R}^K$$

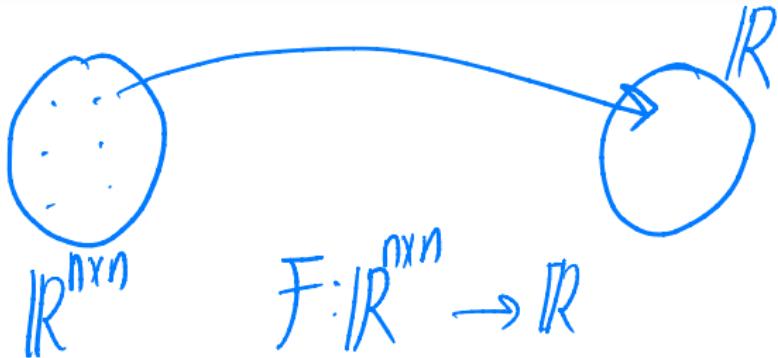
$$ab^T = \begin{bmatrix} n \times 1 \end{bmatrix} \begin{matrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{matrix} \begin{matrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{matrix} \begin{matrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{matrix} = \begin{bmatrix} a_1 b_1 & a_2 b_2 & a_3 b_3 & \dots \\ a_1 b_1 & a_2 b_2 & a_3 b_3 & \dots \\ a_1 b_1 & a_2 b_2 & a_3 b_3 & \dots \\ a_1 b_1 & a_2 b_2 & a_3 b_3 & \dots \\ a_1 b_1 & a_2 b_2 & a_3 b_3 & \dots \end{bmatrix}^K$$

Diagram illustrating the outer product ab^T . A column vector a of size $n \times 1$ is multiplied by a row vector b^T of size $1 \times K$. The result is a matrix of size $n \times K$, where each element is the product of the corresponding elements from a and b^T . The diagram shows a red scribble over the multiplication symbol.

$$ba^T = \begin{bmatrix} K \times 1 \end{bmatrix} \begin{matrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{matrix} \begin{matrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{matrix} \begin{matrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{matrix} = \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix}^{K \times n}$$

Diagram illustrating the outer product ba^T . A row vector b of size $K \times 1$ is multiplied by a column vector a^T of size $1 \times n$. The result is a matrix of size $K \times n$, where each row is a copy of the vector b .

Определитель



1. $A \in \mathbb{R}^{1 \times 1}$ $\det(A) = a_{11}$

2.
A diagram of a 2×2 matrix A with elements $a_{11}, a_{12}, a_{21}, a_{22}$ labeled. The matrix is enclosed in a red oval. To its right, the formula for the determinant is given as $\det(A) = a_{11} \cdot a_{22} - a_{21} \cdot a_{12}$, where the terms $a_{11} \cdot a_{22}$ and $-a_{21} \cdot a_{12}$ are each enclosed in a purple oval.

$A \in \mathbb{R}^{2 \times 2}$ $\begin{bmatrix} (a) & (b) \end{bmatrix}$
A diagram showing a parallelogram formed by two vectors originating from the same point. One vector is labeled a and the other b . Below the parallelogram, the formula $|\det(A)| = S_{ab}$ is written, where S_{ab} represents the area of the parallelogram.

Определитель

3. $A \in \mathbb{R}^{3 \times 3}$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\det A = a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} -$$

$$- a_{12} \cdot \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} +$$

$$+ a_{13} \cdot \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$