$f \to \min\limits_{x,y,z}$

Basic linear algebra background

Формальности

Группа

$\boxed{\text{mipt23.fmin.xyz}}$

Vectors and matrices

We will treat all vectors as column vectors by default. The space of real vectors of length n is denoted by $\mathbb{R}^n$, while the space of real-valued $m \times n$ matrices is denoted by $\mathbb{R}^{m \times n}$. That's it: [1]

Курс годовой
- ТЕСТЫ 0.3
- домашки 0.3 после каждой темы
- проекты 0.2
- экзамен 0.2 ← теоретический минимум

fmin.xyz — commit

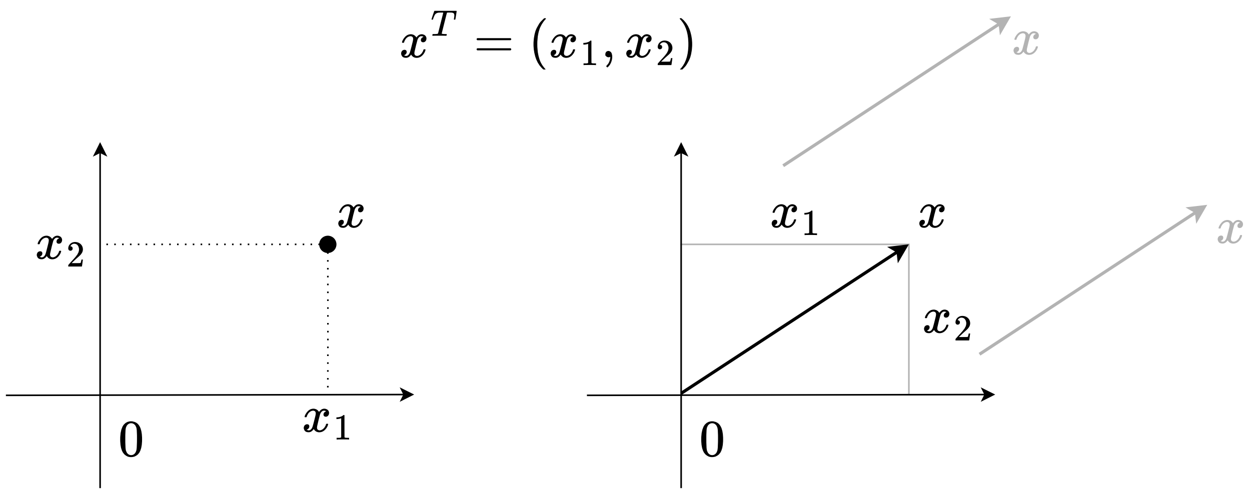
$$
x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad x \in \mathbb{R}^n, x_i \in \mathbb{R}
$$

Твёрдо и чётко:
markdown + latex
или
latex
$\boxed{\text{PDF}}$

Similarly, if $A \in \mathbb{R}^{m \times n}$ we denote transposition as $A^T \in \mathbb{R}^{n \times m}$:

$$
A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \quad A \in \mathbb{R}^{m \times n}, a_i
$$

We will write $x \geq 0$ and $x \neq 0$ to indicate componentwise relationships



A matrix is symmetric if $A = A^T$. It is denoted as $A \in \mathbb{S}^n$ (set of square symmetric matrices of dimension n). Note, that only square matrix could be symmetric by definition.

A matrix $A \in \mathbb{S}^n$ is called **positive (negative) definite** if for all $x \neq 0 : x^T Ax > (<) 0$

Basic linear algebra background

Vectors and matrices

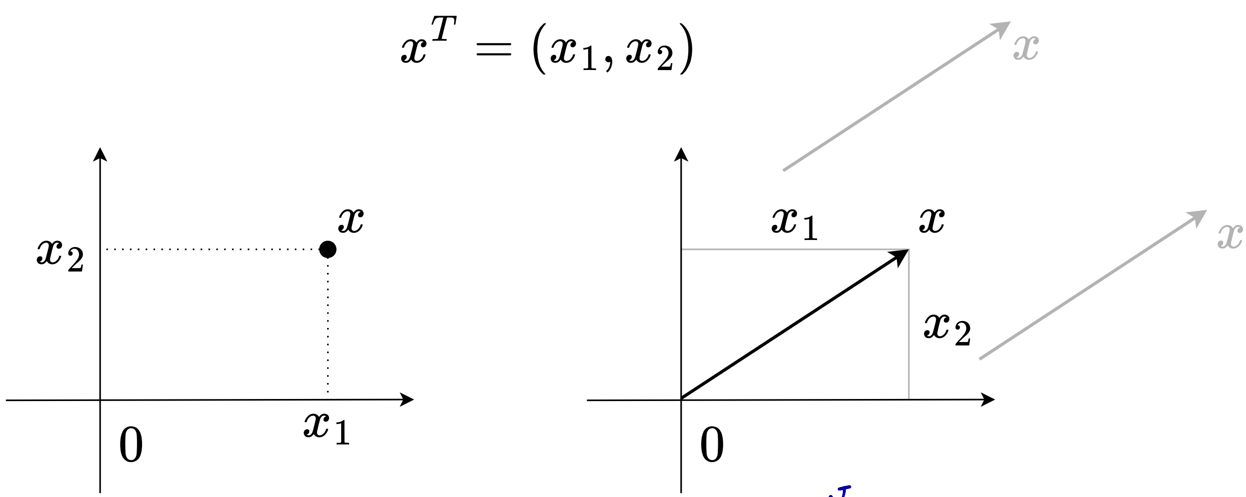
We will treat all vectors as column vectors by default. The space of real vectors of length n is denoted by $\mathbb{R}^n$, while the space of real-valued $m \times n$ matrices is denoted by $\mathbb{R}^{m \times n}$. That's it: [1]

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad x \in \mathbb{R}^n, x_i \in \mathbb{R}$$

Similarly, if $A \in \mathbb{R}^{m \times n}$ we denote transposition as $A^T \in \mathbb{R}^{n \times m}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \quad A \in \mathbb{R}^{m \times n}, a_{i}$$

We will write $x \geq 0$ and $x \neq 0$ to indicate componentwise relationships



$A = A^T$

A matrix is symmetric if $A = A^T$. It is denoted as $A \in \mathbb{S}^n$ (set of square symmetric matrices of dimension n). Note, that only square matrix could be symmetric by definition.

A matrix $A \in \mathbb{S}^n$ is called **positive (negative) definite** if for all $x \neq 0 : x^T Ax > (<)0$

1xn nxn nx1

1x1

. We denote this as $A \succ (\prec) 0$. The set of such matrices is denoted as $\mathbb{S}^n_{++} (\mathbb{S}^n_{--})$

A matrix $A \in \mathbb{S}^n$ is called **positive (negative) semidefinite** if for all $x : x^T A x \geq (\leq) 0$. We denote this as $A \succeq (\preceq) 0$. The set of such matrices is denoted as $\mathbb{S}^n_{+} (\mathbb{S}^n_{-})$

> 🤔 **QUESTION**
>
> Is it correct, that positive definite matrix has all positive entries?

Matrix and vector product

Let A be a matrix of size $m \times n$, and B be a matrix of size $n \times p$, and let the product AB be:

СТРОКА на столбец

$$
\underset{m\times p}{C} = \underset{m\times n}{A}\,\underset{n\times p}{B}
$$

then C is a $m \times p$ matrix, with element (i, j) given by:

$$
\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.}
$$

This operation in a naive form requires $\mathcal{O}(n^3)$ arithmetical operations, where n is usually assumed as the largest dimension of matrices.

$O(n^{\log_2 7})$

> 🤔 **QUESTION**
>
> Is it possible to multiply two matrices faster, then $\mathcal{O}(n^3)$? How about $\mathcal{O}(n^2)$, $\cancel{\mathcal{O}(n)}$?

3

2

Let A be a matrix of shape $m \times n$, and x be $n \times 1$ vector, then the i-th component of the product:

$$
\boxed{\underset{m\times 1}{z} = \underset{m\times n}{A}\,\underset{n\times 1}{x}}
$$

$O(n^2)$

is given by:

$$
z_i = \sum_{k=1}^{n} a_{ik} x_k
$$

Remember, that:

- $C = AB \quad C^T = B^T A^T$
- $AB \neq BA$

- $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$ - матричная экспонента
- $e^{A+B} \neq e^A e^B$ (but if A and B are commuting matrices, which means that $AB = BA$, $e^{A+B} = e^A e^B$)
- $\langle x, Ay \rangle = \langle A^T x, y \rangle$

Norms and scalar products

Norm is a **qualitative measure of smallness of a vector** and is typically denoted as $\|x\|$.

The norm should satisfy certain properties:

1. $\|\alpha x\| = |\alpha| \|x\|$, $\alpha \in \mathbb{R}$
2. $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)
3. If $\|x\| = 0$ then $x = 0$

The distance between two vectors is then defined as

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

The most well-known and widely used norm is **euclidean norm**:

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

which corresponds to the distance in our real life. If the vectors have complex elements, we use their modulus.

Euclidean norm, or 2-norm, is a subclass of an important class of p-norms:

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

There are two very important special cases:

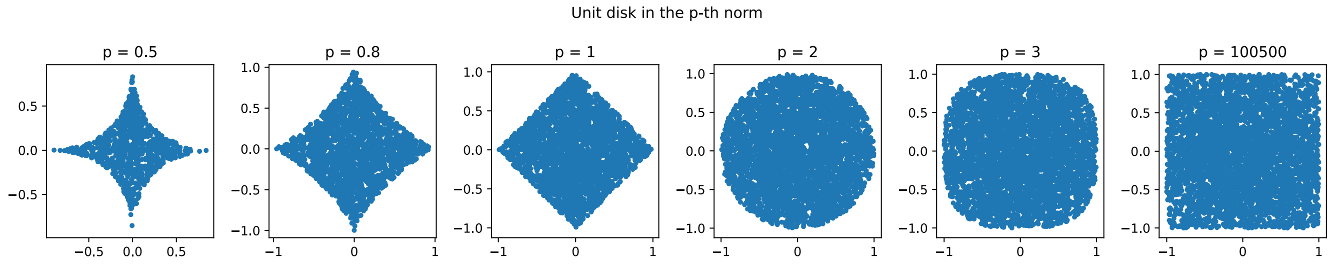
- Infinity norm, or Chebyshev norm is defined as the element of the maximal absolute value:

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

- L_1 norm (or **Manhattan distance**) which is defined as the sum of modules of the elements of x:

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

L_1 norm plays very important role: it all relates to the **compressed sensing** methods that emerged in the mid-00s as one of the most popular research topics. The code for picture below is available here: Open In Colab



In some sense there is no big difference between matrices and vectors (you can vectorize the matrix), and here comes the simplest matrix norm **Frobenius** norm:

$$\boxed{\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{1/2}}$$

Spectral norm, $\|A\|_2$ is one of the most used matrix norms (along with the Frobenius norm).

спектральная норма

$$\boxed{\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2},}$$

It can not be computed directly from the entries using a simple formula, like the Frobenius norm, however, there are efficient algorithms to compute it. It is directly related to the **singular value decomposition** (SVD) of the matrix. It holds

$$\boxed{\|A\|_2 = \sigma_1(A) = \sqrt{\lambda_{\max}(A^TA)}}$$

where $\sigma_1(A)$ is the largest singular value of the matrix A.

🤔 **QUESTION**

Is it true, that all matrix norms satisfy submultiplicativity property: $\|AB\| \leq \|A\|\|B\|$? Hint: consider Chebyshev matrix norm $\|A\|_C = \max\limits_{i,j} |a_{ij}|$.

The standard **scalar (inner) product** between vectors x and y from $\mathbb{R}^n$ is given by

$\langle X, X\rangle = \|X\|^2$

$$\boxed{\langle x, y\rangle = \underset{1\times n}{x^T}\underset{n\times 1}{y} = \sum_{i=1}^{n} x_i y_i = y^T x = \langle y, x\rangle}$$

Here x_i and y_i are the scalar i-th components of corresponding vectors.

> 🤔 **QUESTION**
>
> Is there any connection between the norm $\|\cdot\|$ and scalar product $\langle\cdot,\cdot\rangle$?

> 🧑‍🏫 **EXAMPLE**
>
> $= x^T A y$
>
> Prove, that you can switch the position of a matrix inside scalar product with transposition: $\boxed{\langle x, Ay\rangle = \langle A^T x, y\rangle}$ and $\langle x, yB\rangle = \langle xB^T, y\rangle$

The standard **scalar (inner) product** between matrices X and Y from $\mathbb{R}^{m\times n}$ is given by

$$\boxed{\langle X, Y\rangle = \operatorname{tr}(\underbrace{\underset{n\times m}{X^T}\underset{m\times n}{Y}}_{n\times n}) = \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}Y_{ij} = \operatorname{tr}(Y^T X) = \langle Y, X\rangle}$$

> 🤔 **QUESTION**
>
> $\|X\|_F^2 = \langle X, X\rangle$
>
> Is there any connection between the Frobenious norm $\|\cdot\|_F$ and scalar product between matrices $\langle\cdot,\cdot\rangle$?

> 🧑‍🏫 **EXAMPLE**
>
> Simplify the following expression:
>
> $$\sum_{i=1}^{n}\langle S^{-1}a_i, a_i\rangle = \varkappa, \text{ where } \boxed{S = \sum_{i=1}^{n} a_i a_i^T}, a_i \in \mathbb{R}^n, \det(S) \neq 0$$
>
> ▼ Solution
>
> 1) $\sum_{i=1}^{n}\langle S^{-1}a_i, a_i\rangle = \sum_{i=1}^{n} m_{ii} = \operatorname{tr}(M)$
>
> 2) $S = A\cdot A^T$ SKELETON, где $A = \begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix}_{n\times n}$, $A^T = \begin{pmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_n^T & - \end{pmatrix}$; $S^{-1} = (AA^T)^{-1} = A^{-T}\cdot A^{-1} = S^{-T}$
>
> 3) $m_{ii} = a_i^T S^{-T} a_i$ $\quad$ $a_i^T \cdot S^{-1}\cdot a_i$
>
> Если $M = \underbrace{A^T}_{i\text{-ая строка}} \cdot \underbrace{S^{-1}A}_{i\text{-ый столбец}}$ $\quad$ $a_i^T \cdot S^{-1}a_i$ $\quad$ $\boxed{S^{-1}}\begin{pmatrix} a_1 & \cdots & a_n\end{pmatrix}$
>
> 4) $\varkappa = \operatorname{tr}(A^T S^{-1} A) = \operatorname{tr}(AA^T\cdot S^{-1}) = \operatorname{tr}(SS^{-1}) = \operatorname{tr} I = n.$