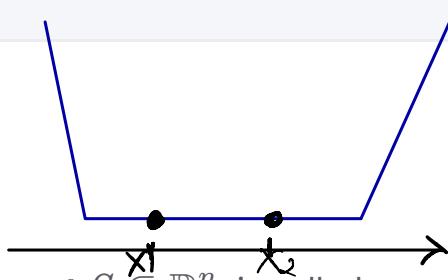




$f \rightarrow \min_{x,y,z}$

Theory / Convex function

# Convex function



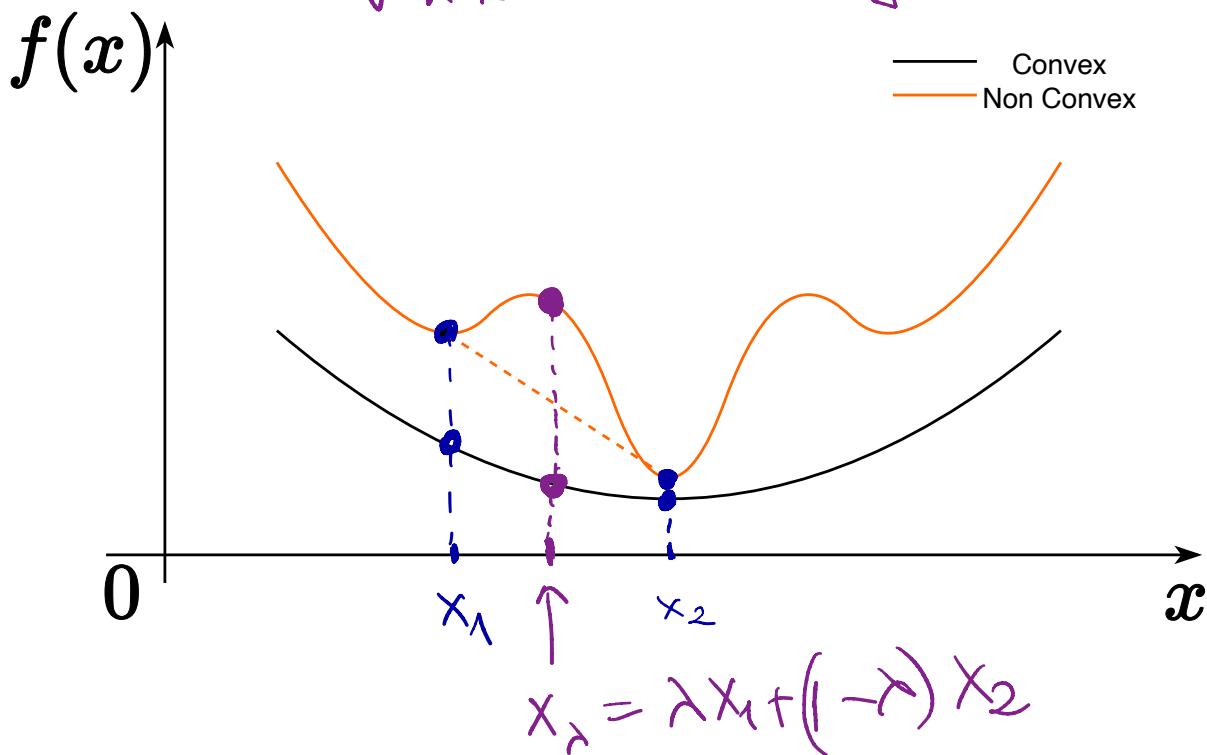
The function  $f(x)$ , which is defined on the convex set  $S \subseteq \mathbb{R}^n$ , is called **convex** on  $S$ , if:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

for any  $x_1, x_2 \in S$  and  $0 \leq \lambda \leq 1$ .

If above inequality holds as strict inequality  $x_1 \neq x_2$  and  $0 < \lambda < 1$ , then function is called strictly convex on  $S$ .

Для вогнуткої функції  
А локальної мінімуму = ГЛОБАЛЬНИЙ



## EXAMPLE

- $f(x) = x^p, p > 1, x \in \mathbb{R}_+$
- $f(x) = \|x\|^p, p > 1, x \in \mathbb{R}^n$
- $f(x) = e^{cx}, c \in \mathbb{R}, x \in \mathbb{R}$
- $f(x) = -\ln x, x \in \mathbb{R}_{++}$
- $f(x) = x \ln x, x \in \mathbb{R}_{++}$
- The sum of the largest  $k$  coordinates  $f(x) = x_{(1)} + \dots + x_{(k)}, x \in \mathbb{R}^n$

$$\begin{aligned} f(x) &= \|x\| \\ x_\lambda &= \lambda x_1 + (1 - \lambda) x_2 \\ f(x_\lambda) &\leq \lambda f(x_1) + (1 - \lambda) f(x_2) \end{aligned}$$

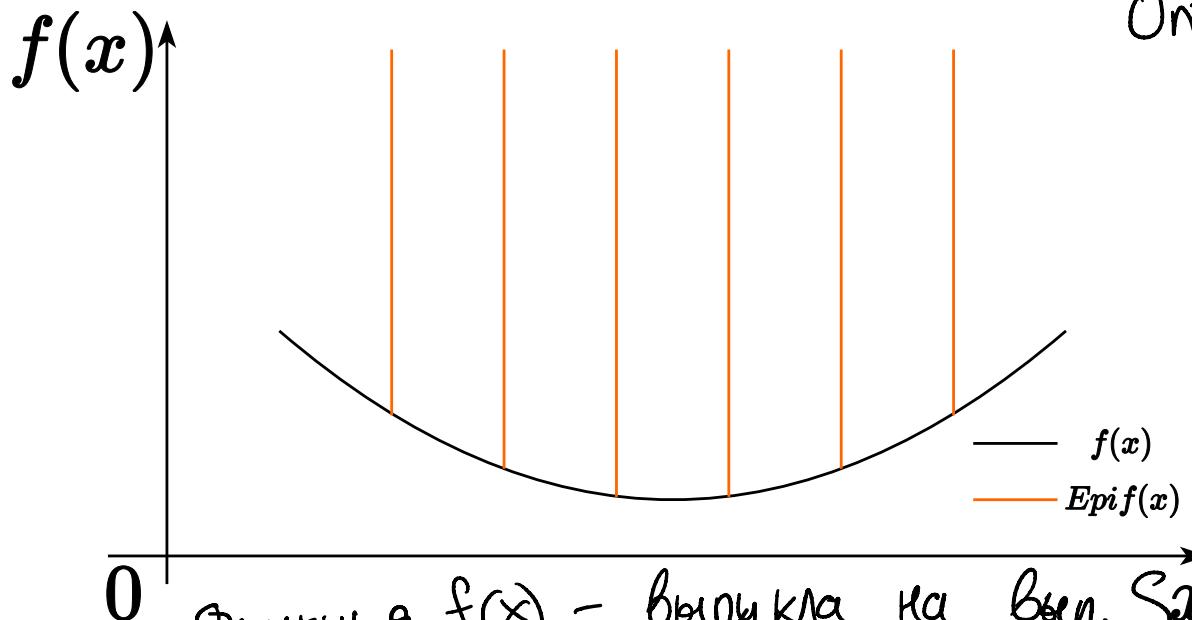
- $f(X) = \lambda_{\max}(X), \quad X = X^T$
- $f(X) = -\log \det X, \quad X \in S_{++}^n$

## Epigraph Награфик

For the function  $f(x)$ , defined on  $S \subseteq \mathbb{R}^n$ , the following set:

$$\text{epi } f = \{[x, \mu] \in S \times \mathbb{R} : f(x) \leq \mu\}$$

is called **epigraph** of the function  $f(x)$ .



## Sublevel set

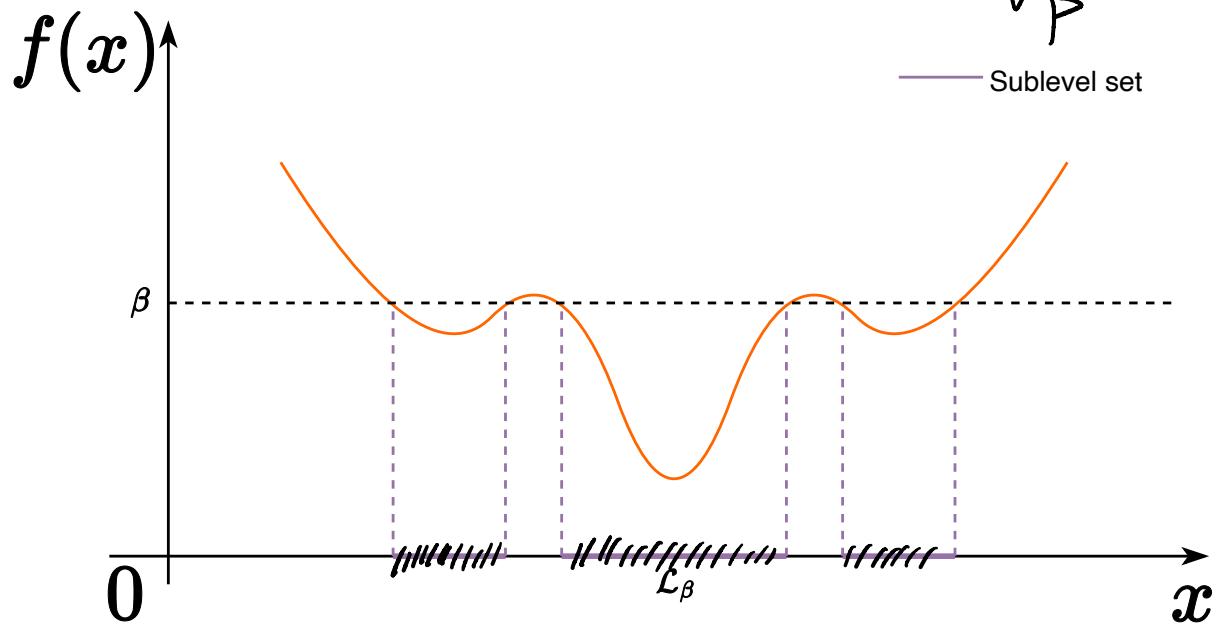
For the function  $f(x)$ , defined on  $S \subseteq \mathbb{R}^n$ , the following set:

$$\mathcal{L}_\beta = \{x \in S : f(x) \leq \beta\}$$

is called **sublevel set** or Lebesgue set of the function  $f(x)$ .

$$f(x) - \text{бюг K нак} \iff L_\beta - \text{бюг K нак}$$

$\forall \beta$



## Criteria of convexity

### First order differential criterion of convexity

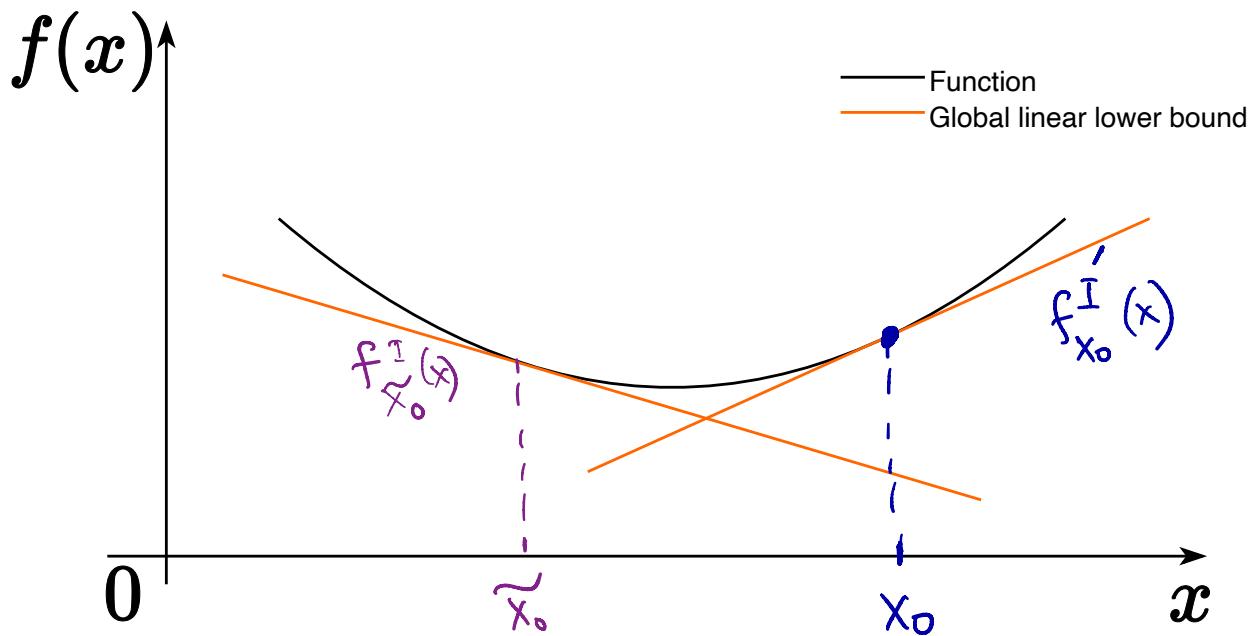
The differentiable function  $f(x)$  defined on the convex set  $S \subseteq \mathbb{R}^n$  is convex if and only if  $\forall x, y \in S$ :

$$f(y) \geq f(x) + \nabla f^T(x)(y - x)$$

Let  $y = x + \Delta x$ , then the criterion will become more tractable:

$$f_x^I(y) \geq f_x^I(y)$$

$$f(x + \Delta x) \geq f(x) + \nabla f^T(x)\Delta x$$



## Second order differential criterion of convexity

Twice differentiable function  $f(x)$  defined on the convex set  $S \subset \mathbb{R}^n$  is convex if and only if  $\forall x \in \text{int}(S) \neq \emptyset$ :

$$\nabla^2 f(x) \succeq 0$$

$$f(x, y) = x \cdot y$$

In other words,  $\forall y \in \mathbb{R}^n$ :

$$\langle y, \nabla^2 f(x)y \rangle \geq 0$$

## Connection with epigraph

The function is convex if and only if its epigraph is a convex set.

### EXAMPLE

Let a norm  $\|\cdot\|$  be defined in the space  $U$ . Consider the set:

$$K := \{(x, t) \in U \times \mathbb{R}^+ : \|x\| \leq t\}$$

гокажем, что  $K$ -внукло

настрадафик  
 $\|x\|$

which represents the epigraph of the function  $x \mapsto \|x\|$ . This set is called the cone norm. According to statement above, the set  $K$  is convex.

In the case where  $U = \mathbb{R}^n$  and  $\|x\| = \|x\|_2$  (Euclidean norm), the abstract set  $K$  transitions into the set:

$$\{(x, t) \in \mathbb{R}^n \times \mathbb{R}^+ : \|x\|_2 \leq t\}$$

## Connection with sublevel set

If  $f(x)$  - is a convex function defined on the convex set  $S \subseteq \mathbb{R}^n$ , then for any  $\beta$  sublevel set  $\mathcal{L}_\beta$  is convex.

The function  $f(x)$  defined on the convex set  $S \subseteq \mathbb{R}^n$  is closed if and only if for any  $\beta$  sublevel set  $\mathcal{L}_\beta$  is closed.  $\text{f}(x) = \det X \quad g(t) = \det(X + tY)$

## Reduction to a line

$f : S \rightarrow \mathbb{R}$  is convex if and only if  $S$  is a convex set and the function  $g(t) = f(x + tv)$  defined on  $\{t \mid x + tv \in S\}$  is convex for any  $x \in S, v \in \mathbb{R}^n$ , which allows to check convexity of the scalar function in order to establish convexity of the vector function.

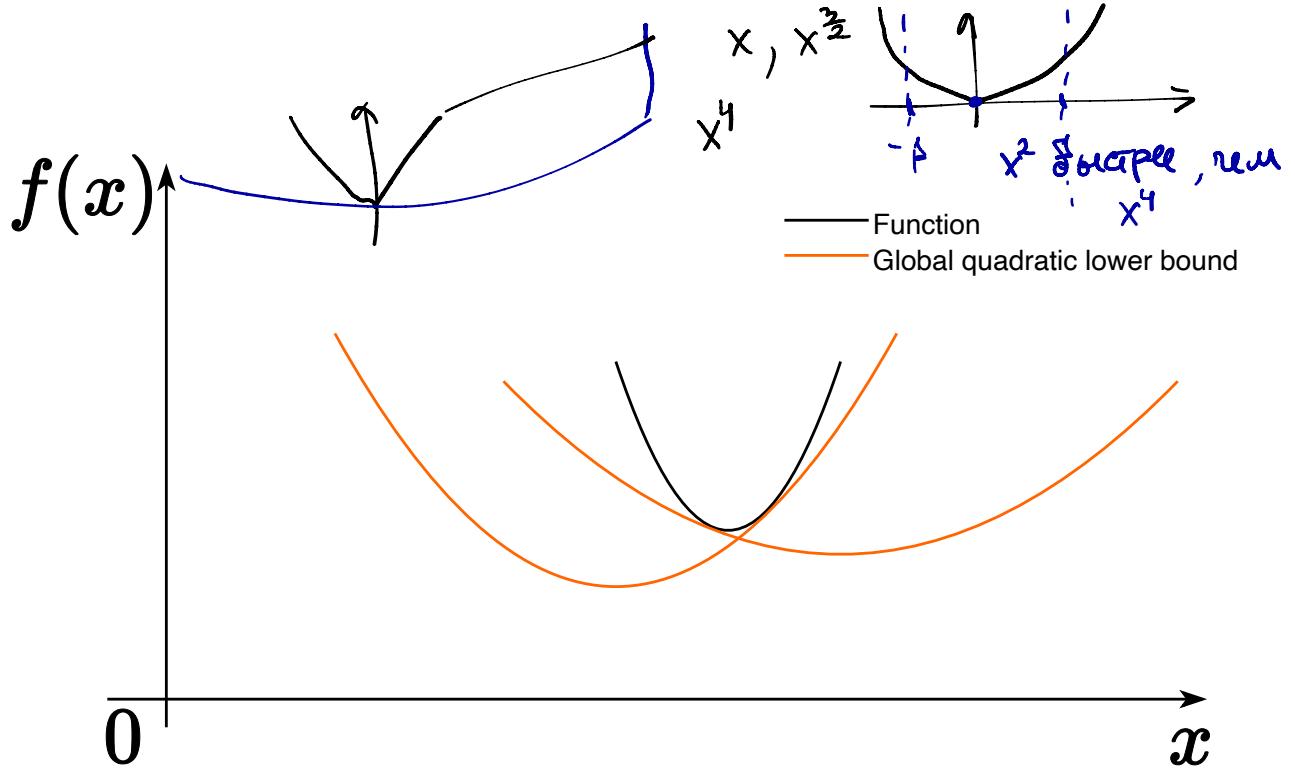
## Strong convexity

$f(x)$ , defined on the convex set  $S \subseteq \mathbb{R}^n$ , is called  $\mu$ -strongly convex (strongly convex) on  $S$ , if:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) - \mu\lambda(1 - \lambda)\|x_1 - x_2\|^2$$

for any  $x_1, x_2 \in S$  and  $0 \leq \lambda \leq 1$  for some  $\mu > 0$ .

$\mu = 0$    $f(x) - \text{bennykncd}$



## Criteria of strong convexity

### First order differential criterion of strong convexity

Differentiable  $f(x)$  defined on the convex set  $S \subseteq \mathbb{R}^n$  is  $\mu$ -strongly convex if and only if  $\forall x, y \in S$ :

$$f(y) \geq f(x) + \nabla f^T(x)(y - x) + \frac{\mu}{2}\|y - x\|^2$$

Let  $y = x + \Delta x$ , then the criterion will become more tractable:

$$f(x + \Delta x) \geq f(x) + \nabla f^T(x)\Delta x + \frac{\mu}{2}\|\Delta x\|^2$$

### Second order differential criterion of strong convexity

Twice differentiable function  $f(x)$  defined on the convex set  $S \subseteq \mathbb{R}^n$  is called  $\mu$ -strongly convex if and only if  $\forall x \in \text{int}(S) \neq \emptyset$ :

enough

$$\nabla^2 f(x) \succeq \mu I$$

$$\min(\nabla^2 f(x)) = \mu > 0$$

In other words:  $\lambda_{\min} > \mu > 0$

$$\langle y, \nabla^2 f(x)y \rangle \geq \mu\|y\|^2$$

# Facts

- $f(x)$  is called (strictly) concave, if the function  $-f(x)$  - is (strictly) convex.
- Jensen's inequality for the convex functions:

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i)$$

for  $\alpha_i \geq 0$ ;  $\sum_{i=1}^n \alpha_i = 1$  (probability simplex)

For the infinite dimension case:

$$f\left(\int_S x p(x) dx\right) \leq \int_S f(x) p(x) dx$$

If the integrals exist and  $p(x) \geq 0$ ,  $\int_S p(x) dx = 1$

- If the function  $f(x)$  and the set  $S$  are convex, then any local minimum  $x^* = \arg \min_{x \in S} f(x)$  will be the global one. Strong convexity guarantees the uniqueness of the solution.
- Let  $f(x)$  - be a convex function on a convex set  $S \subseteq \mathbb{R}^n$ . Then  $f(x)$  is continuous  $\forall x \in \text{ri}(S)$ .

## Operations that preserve convexity

- Non-negative sum of the convex functions:  $\alpha f(x) + \beta g(x)$ , ( $\alpha \geq 0, \beta \geq 0$ ).
- Composition with affine function  $f(Ax + b)$  is convex, if  $f(x)$  is convex.
- Pointwise maximum (supremum): If  $f_1(x), \dots, f_m(x)$  are convex, then  $f(x) = \max\{f_1(x), \dots, f_m(x)\}$  is convex.
- If  $f(x, y)$  is convex on  $x$  for any  $y \in Y$ :  $g(x) = \sup_{y \in Y} f(x, y)$  is convex.
- If  $f(x)$  is convex on  $S$ , then  $g(x, t) = tf(x/t)$  - is convex with  $x/t \in S, t > 0$ .
- Let  $f_1 : S_1 \rightarrow \mathbb{R}$  and  $f_2 : S_2 \rightarrow \mathbb{R}$ , where  $\text{range}(f_1) \subseteq S_2$ . If  $f_1$  and  $f_2$  are convex, and  $f_2$  is increasing, then  $f_2 \circ f_1$  is convex on  $S_1$ .

affine vs linear  
 $Ax + b$        $f(b) = 0$

# Other forms of convexity

- Log-convex:  $\log f$  is convex; Log convexity implies convexity.
- Log-concavity:  $\log f$  concave; **not** closed under addition!
- Exponentially convex:  $[f(x_i + x_j)] \succeq 0$ , for  $x_1, \dots, x_n$
- Operator convex:  $f(\lambda X + (1 - \lambda)Y) \preceq \lambda f(X) + (1 - \lambda)f(Y)$
- Quasiconvex:  $f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$
- Pseudoconvex:  $\langle \nabla f(y), x - y \rangle \geq 0 \rightarrow f(x) \geq f(y)$
- Discrete convexity:  $f : \mathbb{Z}^n \rightarrow \mathbb{Z}$ ; "convexity + matroid theory."

## EXAMPLE

Show, that  $f(x) = c^\top x + b$  is convex and concave.

### Solution

$$\begin{aligned} &\Rightarrow \text{Kumeguri I nösgök} \\ &\nabla f = c \quad f(y) \geq f(x) + \nabla f(x)^\top (y-x) \\ &2) \nabla^2 f = 0^n \quad c^\top y + b \geq c^\top x + b + c^\top (y-x) \\ &g(x) = f(x) \quad \nabla g = -c \quad b \geq b \quad \text{r.T.g.} \\ &\nabla^2 g = 0 \end{aligned}$$

o.

(B610)

(BN)

## EXAMPLE

Show, that  $f(x) = x^\top Ax$ , where  $A \succeq 0$  - is convex on  $\mathbb{R}^n$ .

### Solution

Hügeli M.

$$\mu = 0$$

$A \in S^n_+$

$$\Rightarrow \mu = 2 \lambda_{\min}(A)$$

$$\nabla^2 f = 2A \succeq 0$$

- Bolinyknaš

$$\frac{A + A^\top}{2} \cdot 2 \succeq 0$$

## EXAMPLE

Show, that  $f(A) = \lambda_{\max}(A)$  - is convex, if  $A \in S^n_+$ .

▼ Solution

$$\text{No определено: } Ax = \lambda x$$

$$x^T A x = \lambda x^T x$$

$$\Rightarrow \lambda_{\max} = \sup_{x \neq 0} \frac{x^T A x}{x^T x}$$

$f(A)$  - нум.  
но  $A$

сигнатура vs характеристика

функций под max

это хороший вопрос.



### EXAMPLE

PL inequality holds if the following condition is satisfied for some  $\mu > 0$ ,

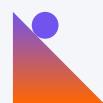
$$\|\nabla f(x)\|^2 \geq \mu(f(x) - f^*) \forall x$$

The example of function, that satisfy PL-condition, but is not convex.  $f(x, y) =$

$$\frac{(y - \sin x)^2}{2}$$

## References

- [Steven Boyd lectures](#)
- [Suvrit Sra lectures](#)
- [Martin Jaggi lectures](#)
- Example pf PI non-convex function [Open in Colab](#)



# Conjugate (dual) set

Пусть  $S \subseteq \mathbb{R}^n$  - произвольное непустое множество. Тогда сопряженное к нему множество определяется, как:

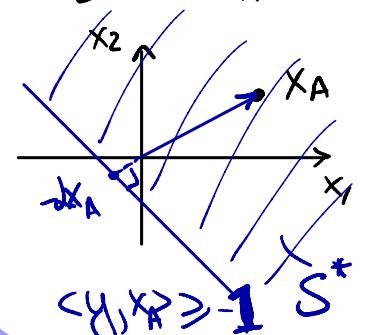
ТУПА:

- 1)  $S^*$  - выпукло
- 2)  $0 \in S^*$

$$S^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in S\}$$

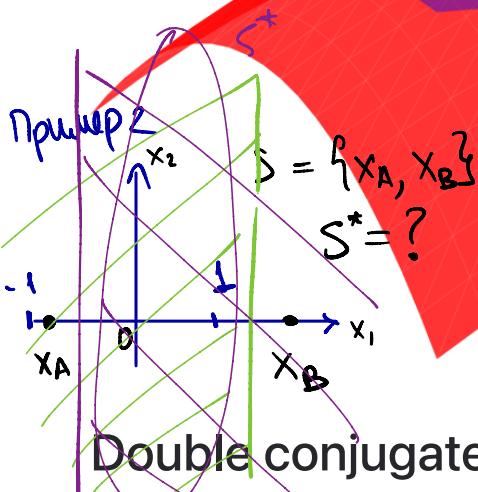
пример 1

$$S = x_A \in \mathbb{R}^2$$



$$\langle y, x \rangle \geq 0$$

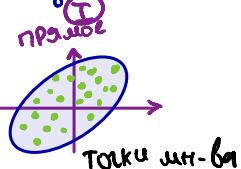
$$\begin{aligned} & \langle y, x \rangle \geq -1 \\ & \langle -2x_A, x_A \rangle \geq -1 \\ & 2\|x_A\|^2 \leq 1 \\ & d \leq \frac{1}{\|x_A\|^2} \end{aligned}$$



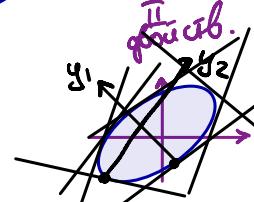
Double conjugate set

Множество  $S^{**}$  называется вторым сопряженным к множеству  $S$ , если:

Выпуклые множества, содержащие 0 допускают двойственное описание



точки множ.



$y_1, y_2, \dots$   
набор  
шерпанс,  
подп.

$$S^{**} = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in S^*\}$$

Inter-conjugate and self-conjugate sets

- Множества  $S_1$  и  $S_2$  называются **взаимосопряженными**, если  $S_1^* = S_2$ ,  $S_2^* = S_1$ .
- Множество  $S$  называется **самосопряженным**, если  $S^* = S$ .

Properties

- Сопряженное множество всегда замкнуто, выпукло и содержит нуль.
- Для произвольного множества  $S \subseteq \mathbb{R}^n$ :

$$S^{**} = \overline{\text{conv}(S \cup \{0\})}$$

- Если  $S_1 \subseteq S_2$ , то  $S_2^* \subseteq S_1^*$ .

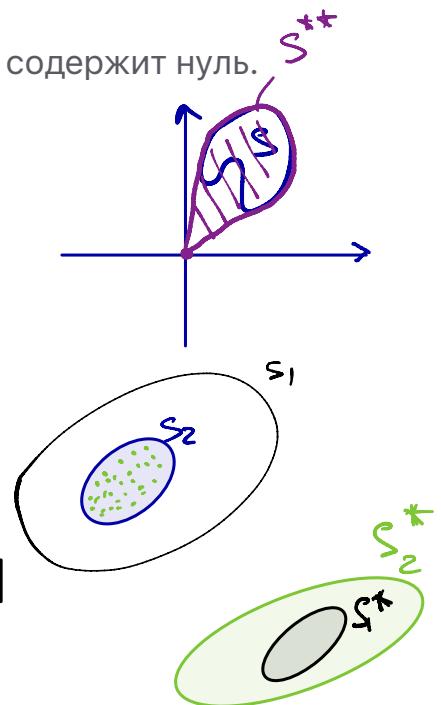
*m-около*

*Когда*

- Если  $S$  - замкнуто, выпукло, включает 0, то  $S^{**} = S$ .

- 

$$S^* = (\overline{S})^*$$



## Examples

1

Доказать, что  $S^* = (\overline{S})^*$ .

Решение:

- $S \subset \overline{S} \rightarrow (\overline{S})^* \subset S^*$
- Пусть  $p \in S^*$  и  $x_0 \in \overline{S}$ ,  $x_0 = \lim_{k \rightarrow \infty} x_k$ . Тогда в силу непрерывности функции  $f(x) = p^T x$ , имеем:  $p^T x_k \geq -1 \rightarrow p^T x_0 \geq -1$ . Значит,  $p \in (\overline{S})^*$ , отсюда  $S^* \subset (\overline{S})^*$

2

Доказать, что  $(\text{conv}(S))^* = S^*$ .

Решение:

- $S \subset \text{conv}(S) \rightarrow (\text{conv}(S))^* \subset S^*$
- Пусть  $p \in S^*$ ,  $x_0 \in \text{conv}(S)$ , т.е.  $x_0 = \sum_{i=1}^k \theta_i x_i \mid x_i \in S, \sum_{i=1}^k \theta_i = 1, \theta_i \geq 0$ .

Значит,  $p^T x_0 = \sum_{i=1}^k \theta_i p^T x_i \geq \sum_{i=1}^k \theta_i (-1) = 1 \cdot (-1) = -1$ . Значит,  $p \in (\text{conv}(S))^*$ , отсюда  $S^* \subset (\text{conv}(S))^*$

3

Доказать, что если  $B(0, r)$  - шар радиуса  $r$  по некоторой норме с центром в нуле, то  $(B(0, r))^* = B(0, 1/r)$ .

Решение:

$$X^* \subset Y \quad \text{I}$$

$$\text{II} \quad Y \subset X^*$$

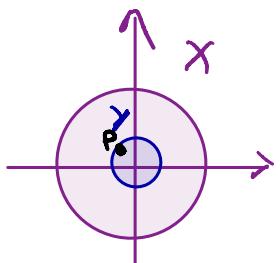
- Пусть  $B(0, r) = X, B(0, 1/r) = Y$ . Возьмем вектор нормали  $p \in X^*$ , тогда для любого  $x \in X : p^T x \geq -1$ .

- Из всех точек шара  $X$   $\langle p, x \rangle$  возьмем такую  $x \in X$ , что скалярное произведение её на  $p$ :  $p^T x$  было бы минимально, тогда это точка  $x = -\frac{p}{\|p\|}r$

$$\text{I}$$

$$p^T x = p^T \left( -\frac{p}{\|p\|}r \right) = -\|p\|r \geq -1$$

$$\|p\| \leq \frac{1}{r} \in Y$$



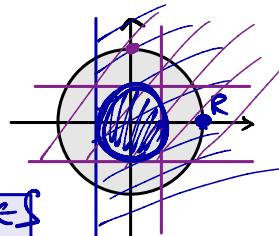
Значит,  $X^* \subset Y$ .

- Теперь пусть  $p \in Y$ . Нам надо показать, что  $p \in X^*$ , т.е.  $\langle p, x \rangle \geq -1$ .

Достаточно применить неравенство Коши - Буняковского:

для  $p$  - маселькин

$$\|\langle p, x \rangle\| \leq \|p\| \|x\| \leq \frac{1}{r} \cdot r = 1$$



Последнее исходит из того, что  $p \in B(0, 1/r)$ , а  $x \in B(0, r)$ .

Значит,  $Y \subset X^*$ .

$$S^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in S\}$$

$$K^* = \{y \in \mathbb{R}^n \mid \lambda x \in K, \lambda \geq 0 \quad \langle y, \lambda x \rangle \geq -1\}$$

## Dual cones

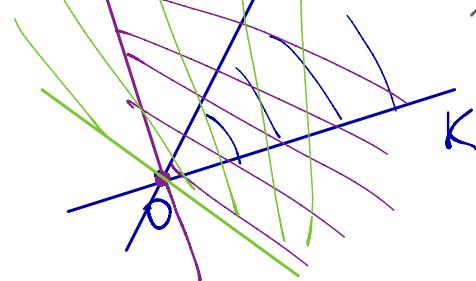
Сопряженным конусом к конусу  $K$  называется такое множество  $K^*$ , что:

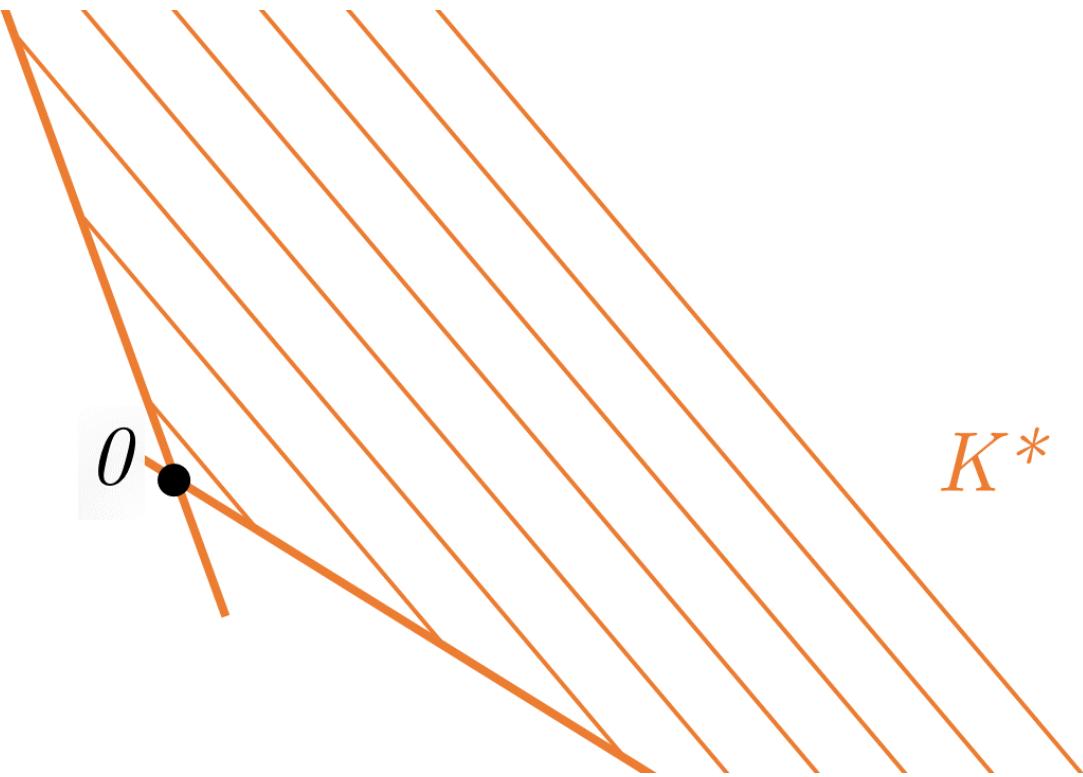
$$K^* = \{y \mid \langle x, y \rangle \geq 0 \quad \forall x \in K\}$$

$$\langle y, x \rangle \geq -\frac{1}{\lambda}$$

Чтобы показать, что это определение непосредственно следует из теории выше, вспомним, что такое сопряженное множество и что такое конус  $\forall \lambda > 0$

$$\{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \quad \forall x \in S\} \rightarrow \{\lambda y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -\frac{1}{\lambda} \quad \forall x \in S\}$$





## Dual cones properties

- Пусть  $K$  - замкнутый выпуклый конус. Тогда  $K^{**} = K$ .
- Для произвольного множества  $S \subseteq \mathbb{R}^n$  и конуса  $K \subseteq \mathbb{R}^n$ :

$$(S + K)^* = S^* \cap K^*$$

- Пусть  $K_1, \dots, K_m$  - конусы в  $\mathbb{R}^n$ , тогда:

$$\left( \sum_{i=1}^m K_i \right)^* = \bigcap_{i=1}^m K_i^*$$

- Пусть  $K_1, \dots, K_m$  - конусы в  $\mathbb{R}^n$ . Пусть также их пересечение имеет внутреннюю точку, тогда:

$$\left( \bigcap_{i=1}^m K_i \right)^* = \sum_{i=1}^m K_i^*$$

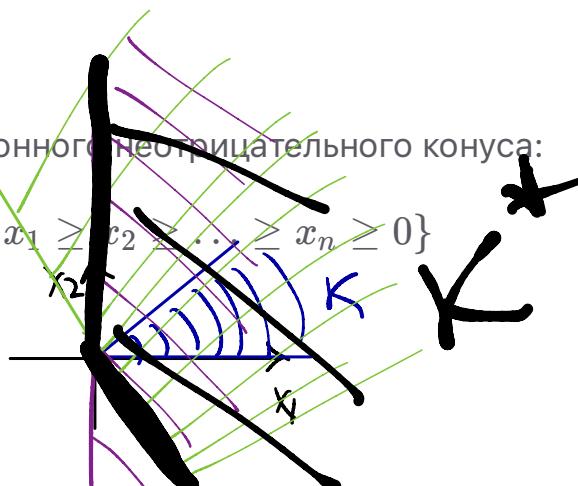
## Examples

Найти сопряженный конус для монотонного неотрицательного конуса:

$$K = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \dots \geq x_n \geq 0\}$$

Решение:

$$K^* = ?$$



Заметим, что:

По опр.  $K^*$ :  $\{y \in \mathbb{R}^n : \langle y, x \rangle \geq 0 \quad \forall x \in K\}$

$$\sum_{i=1}^n x_i \cdot y_i \geq 0$$

$K^* \in ?$

$$\sum_{i=1}^n x_i y_i = y_1(x_1 - x_2) + (y_1 + y_2)(x_2 - x_3) + \dots + (y_1 + y_2 + \dots + y_{n-1})(x_{n-1} - x_n)$$

~~$x_1 y_1 - x_2 y_1 + x_2 y_1 + x_2 y_2 - x_3 y_1 - x_3 y_2 + (y_1 + y_2 + y_3)(x_3 - x_4)$~~

Так как в представленной сумме в каждом слагаемом второй множитель неотрицательный, то:

$$y_1 \geq 0, \quad y_1 + y_2 \geq 0, \quad \dots, \quad y_1 + \dots + y_n \geq 0$$

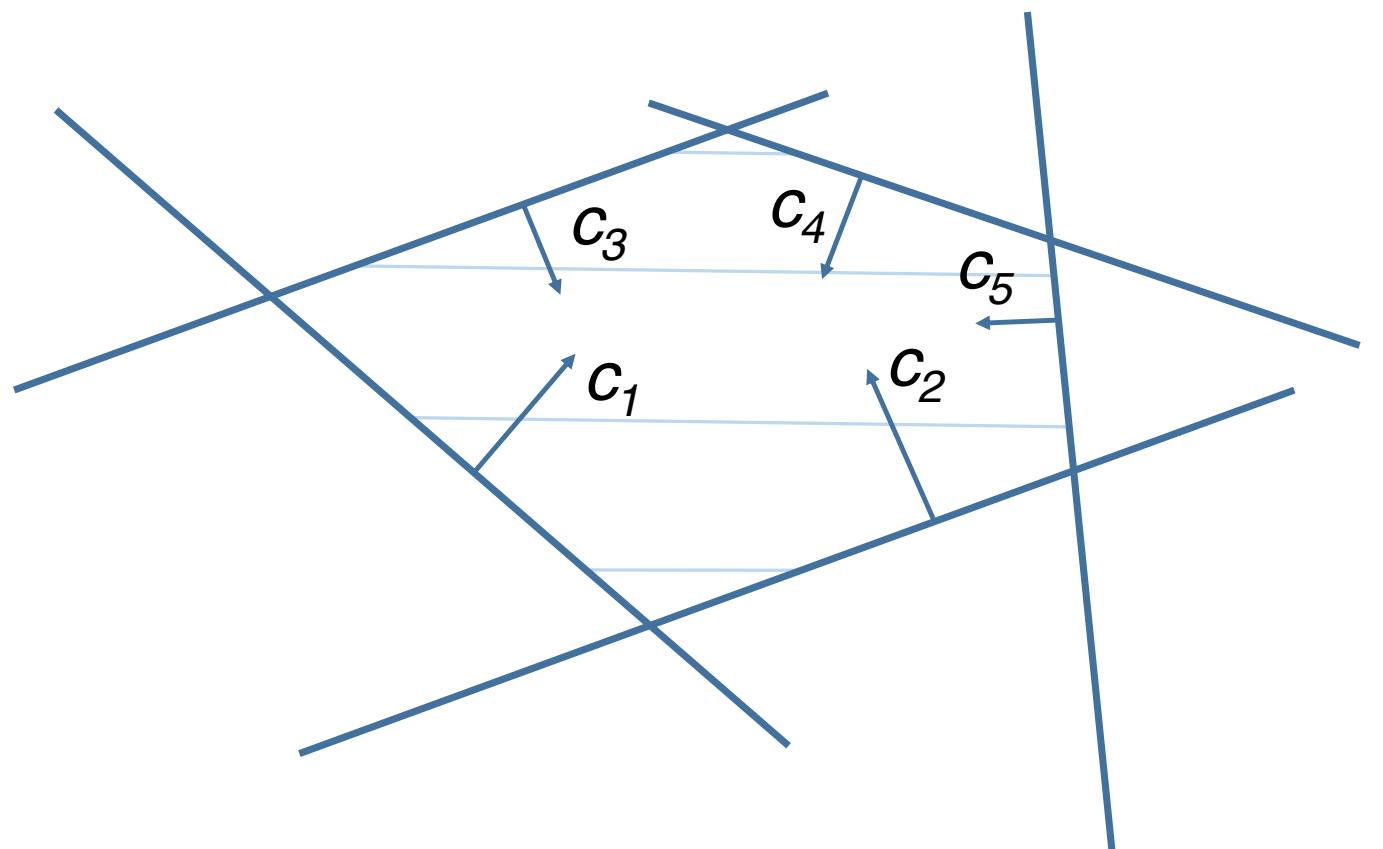
Значит,  $K^* = \left\{ y \mid \sum_{i=1}^k y_i \geq 0, k = \overline{1, n} \right\}$

## Polyhedra

Множество решений системы линейных неравенств и равенств представляет собой многогранник:

$$Ax \leq b, \quad Cx = d$$

Здесь  $A \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{p \times n}$ , а неравенство – поэлементное.



ТЕОРЕМА:

Пусть  $x_1, \dots, x_m \in \mathbb{R}^n$ . Сопряженным к многогранному множеству:

$$S = \mathbf{conv}(x_1, \dots, x_k) + \mathbf{cone}(x_{k+1}, \dots, x_m)$$

является полиэдр (многогранник):

$$S^* = \{p \in \mathbb{R}^n \mid \langle p, x_i \rangle \geq -1, i = \overline{1, k}; \langle p, x_i \rangle \geq 0, i = \overline{k+1, m}\}$$

ДОКАЗАТЕЛЬСТВО:

- Пусть  $S = X, S^* = Y$ . Возьмем некоторый  $p \in X^*$ , тогда  $\langle p, x_i \rangle \geq -1, i = \overline{1, k}$ . В то же время для любых  $\theta > 0, i = \overline{k+1, m}$ :

$$\langle p, x_i \rangle \geq -1 \rightarrow \langle p, \theta x_i \rangle \geq -1$$

$$\langle p, x_i \rangle \geq -\frac{1}{\theta} \rightarrow \langle p, x_i \rangle \geq 0$$

Значит,  $p \in Y \rightarrow X^* \subset Y$ .

- Пусть, напротив,  $p \in Y$ . Для любой точки  $x \in X$ :

$$x = \sum_{i=1}^m \theta_i x_i \quad \sum_{i=1}^k \theta_i = 1, \theta_i \geq 0$$

Значит:

$$\langle p, x \rangle = \sum_{i=1}^m \theta_i \langle p, x_i \rangle = \sum_{i=1}^k \theta_i \langle p, x_i \rangle + \sum_{i=k+1}^m \theta_i \langle p, x_i \rangle \geq \sum_{i=1}^k \theta_i (-1) + \sum_{i=1}^k \theta_i \cdot$$

Значит,  $p \in X^* \rightarrow Y \subset X^*$ .

5

Найти и изобразить на плоскости множество, сопряженное к многогранному конусу:

$$S = \mathbf{cone} \{(-3, 1), (2, 3), (4, 5)\}$$

Решение:



Используя теорему выше:

$$S^* = \{-3p_1 + p_2 \geq 0, 2p_1 + 3p_2 \geq 0, 4p_1 + 5p_2 \geq 0\}$$

**Лемма (теорема) Фаркаша (Фаркаша - Минковского)**

Пусть  $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ . Тогда имеет решение одна и только одна из

следующих двух систем:

$$1) Ax = b, x \geq 0$$

$$2) p^\top A \geq 0, \langle p, b \rangle < 0$$

$Ax = b$  при  $x \geq 0$  означает, что  $b$  лежит в конусе, натянутым на столбцы матрицы  $A$ .

$pA \geq 0, \langle p, b \rangle < 0$  означает, что существует разделяющая гиперплоскость между вектором  $b$  и конусом из столбцов матрицы  $A$ .

СЛЕДСТВИЕ:

Пусть  $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ . Тогда имеет решение одна и только одна из следующих двух систем:

$$1) Ax \leq b$$

$$2) p^\top A = 0, \langle p, b \rangle < 0, p \geq 0$$

Если в задаче линейного программирования на минимум допустимое множество непусто и целевая функция ограничена на нём снизу, то задача имеет решение.

---