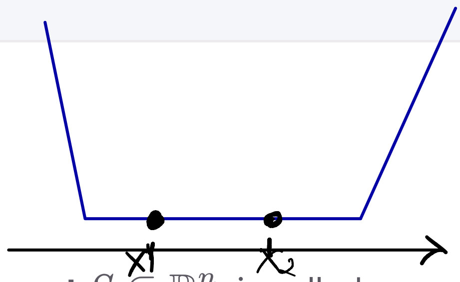
Theory / Convex function

Convex function

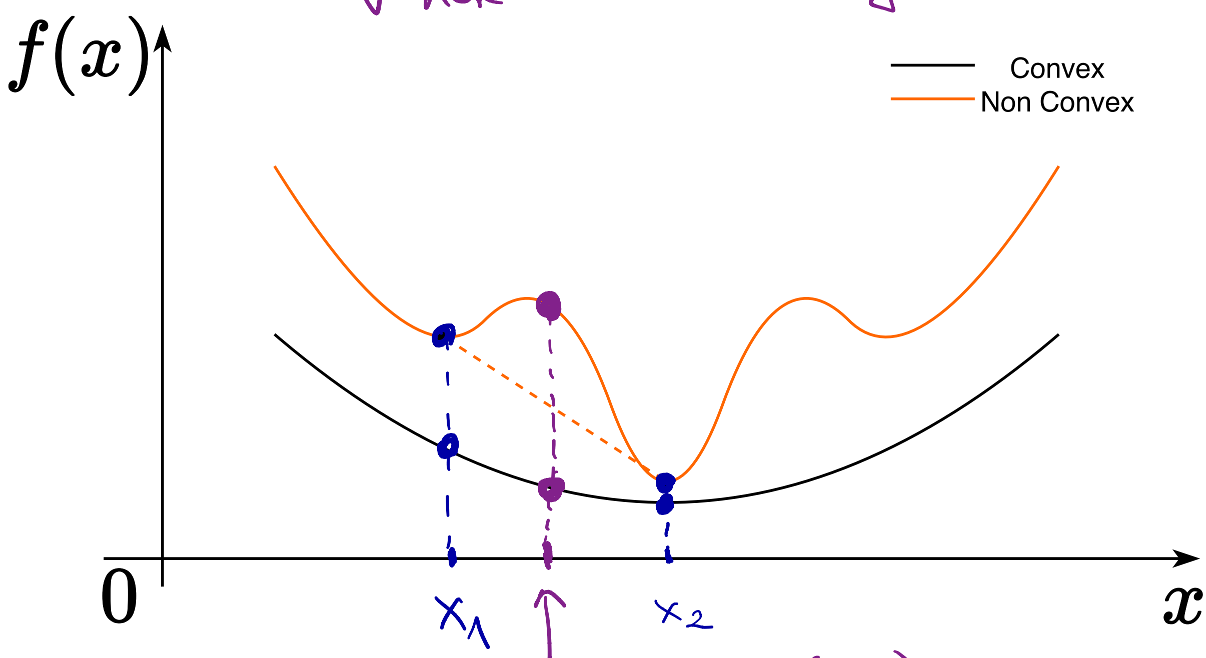


The function $f(x)$, **which is defined on the convex set** $S \subseteq \mathbb{R}^n$, is called **convex** on S, if:

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

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 выпуклой функции
$\forall$ локальный минимум = глобальный



$x_\lambda = \lambda x_1 + (1-\lambda) x_2$

🧑‍🏫 **EXAMPLE**

$f(x) = \|x\|$

$x_\lambda = \lambda x_1 + (1-\lambda) x_2$

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

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

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

Epigraph

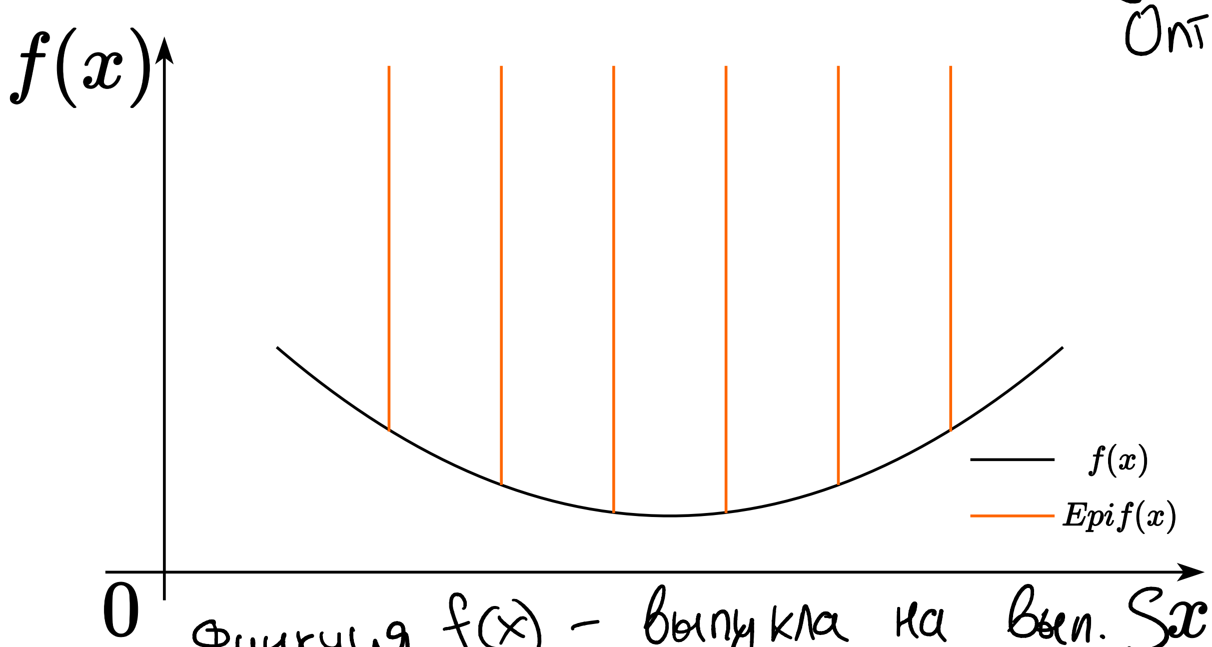
Надграфик

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

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

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

Хадан
Методы
Оптимиза



Функция $f(x)$ - выпукла на вып. S_x

$\Longleftrightarrow$ epi f - выпуклое мн-во

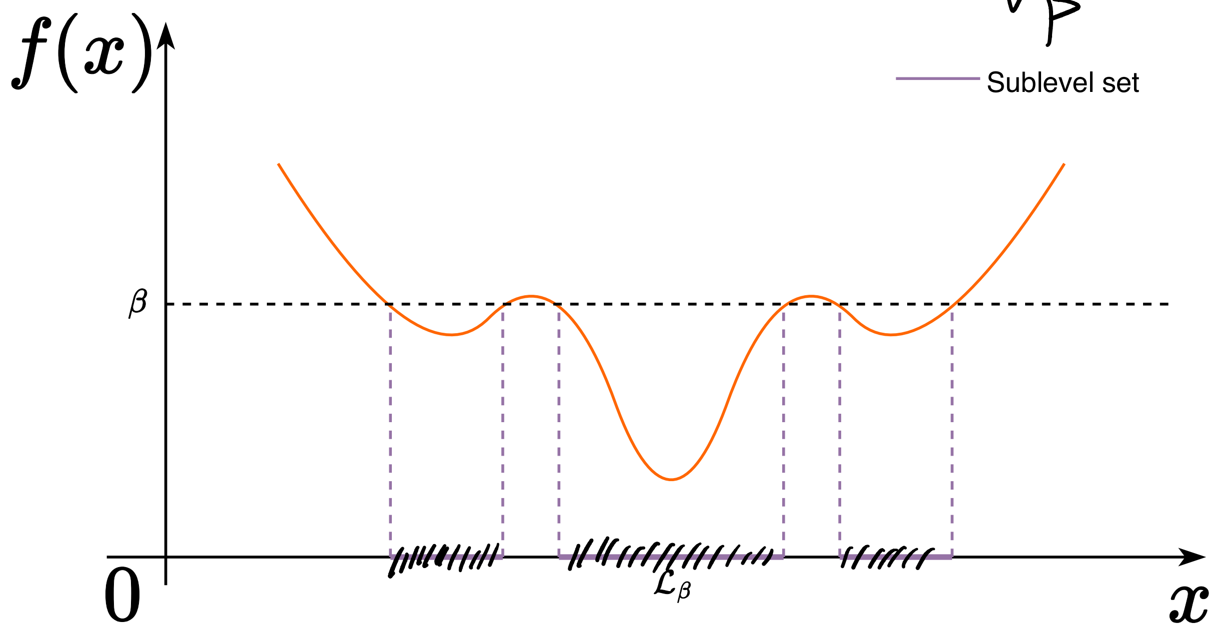
Sublevel set

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

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

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

$f(x)$ – выпуклая $\iff$ L_β – выпуклое $\forall \beta$



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 $\forall x, y \in S$:

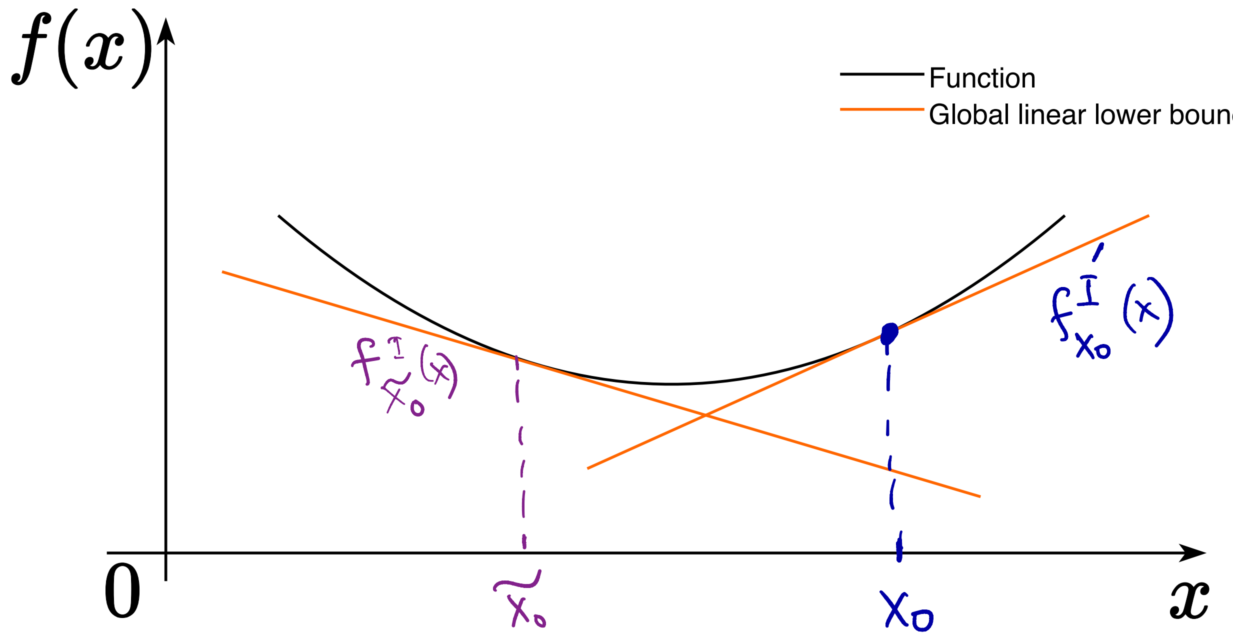
$$
\boxed{f(y) \geq f(x) + \nabla f^T(x)(y - x)}
$$

$f_x^I(y)$

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

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

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



Second order differential criterion of convexity

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

$$\boxed{\nabla^2 f(x) \succeq 0}$$

$\boxed{f(x, y) = x \cdot y}$

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**

докажем, что K- выпукло

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

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

надграфик $\|x\|$

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 β 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 β sublevel set $\mathcal{L}_\beta$ is closed.

$f(X) = \det X \quad g(t) = \det(X + t \cdot Y)$

Reduction to a line

$f : S \to \mathbb{R}$ 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.

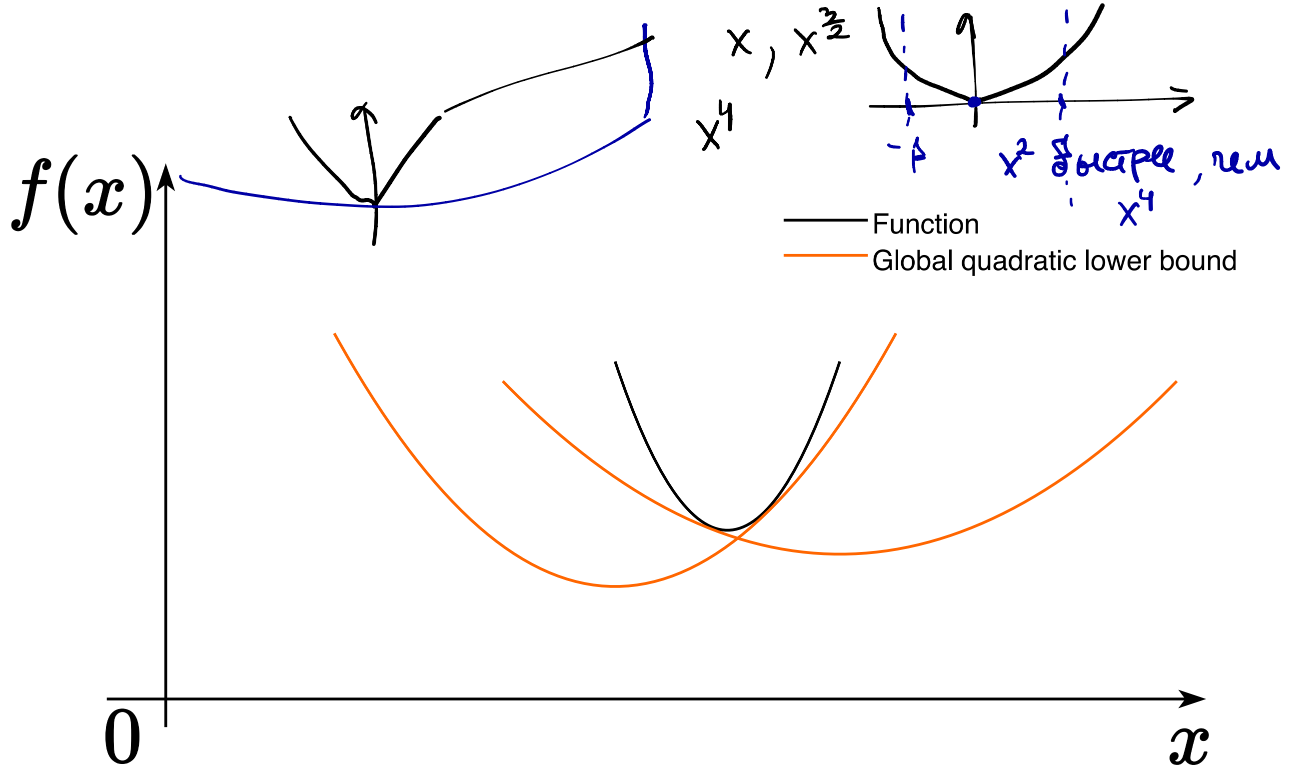
Strong convexity

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

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

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

$\mu = 0 \rightarrow f(x)$ – выпуклая



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 μ-strongly convex if and only if $\forall x, y \in S$:

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

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 μ-strongly convex if and only if $\forall x \in \textbf{int}(S) \neq \emptyset$:

$\mu > 0$

спектр

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

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

0 λ_{min} λ_{max}

In other words:

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

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 $\alpha_i \geq 0; \quad \sum_{i=1}^n \alpha_i = 1$ (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 $p(x) \geq 0, \quad \int_S p(x) dx = 1$

- 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 $\forall x \in \mathbf{ri}(S)$.

Operations that preserve convexity

- Non-negative sum of the convex functions: $\boxed{\alpha f(x) + \beta g(x)}$, $(\alpha \geq 0, \beta \geq 0)$.
- Composition with affine function $\boxed{f(Ax + b) \text{ is convex, if } f(x) \text{ is convex.}}$
- Pointwise maximum (supremum): If $f_1(x), \ldots, f_m(x)$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ 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 $\boxed{g(x, t) = t f(x/t)}$ - is convex with $x/t \in S, t > 0$.
- Let $f_1 : S_1 \to \mathbb{R}$ and $f_2 : S_2 \to \mathbb{R}$, 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

$Ax+b$ $\quad f(0)=0$

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 $x_1, \ldots, x_n$
- 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: $\langle \nabla f(y), x - y\rangle \geq 0 \longrightarrow f(x) \geq f(y)$
- Discrete convexity: $f : \mathbb{Z}^n \to \mathbb{Z}$; "convexity + matroid theory."

🧑‍🏫 **EXAMPLE**

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

0. Вып вп.

▼ Solution

1) Критерий I порядок

$\nabla f = c$

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

$c^T y + b \geq c^T x + b + c^T(y - x)$

$b \geq b$ ч.т.д.

2) $\nabla^2 f = 0^n$

$g(x) = -f(x)$ $\nabla g = -c$

$\nabla^2 g = 0$

🧑‍🏫 **EXAMPLE**

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

▼ Solution

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

Найдём μ.

$\mu = 0$

$\nabla^2 f = \frac{A + A^T}{2} \cdot 2 \succeq 0$

если $A \in S^n_{++}$

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

🧑‍🏫 **EXAMPLE**

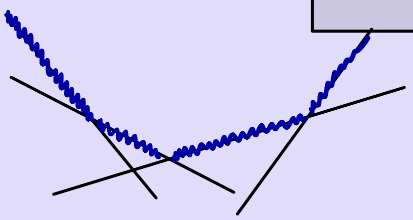
Show, that $f(A) = \lambda_{max}(A)$ - is convex, if $A \in S^n_+$.

▼ Solution По определению: $Ax = \lambda x$

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

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

$f(A)$ - лин. по A



счётность vs несчётность 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. $f(x, y) = \dfrac{(y - \sin x)^2}{2}$

References

- Steven Boyd lectures
- Suvrit Sra lectures
- Martin Jaggi lectures
- Example pf Pl non-convex function Open in Colab