Convex set

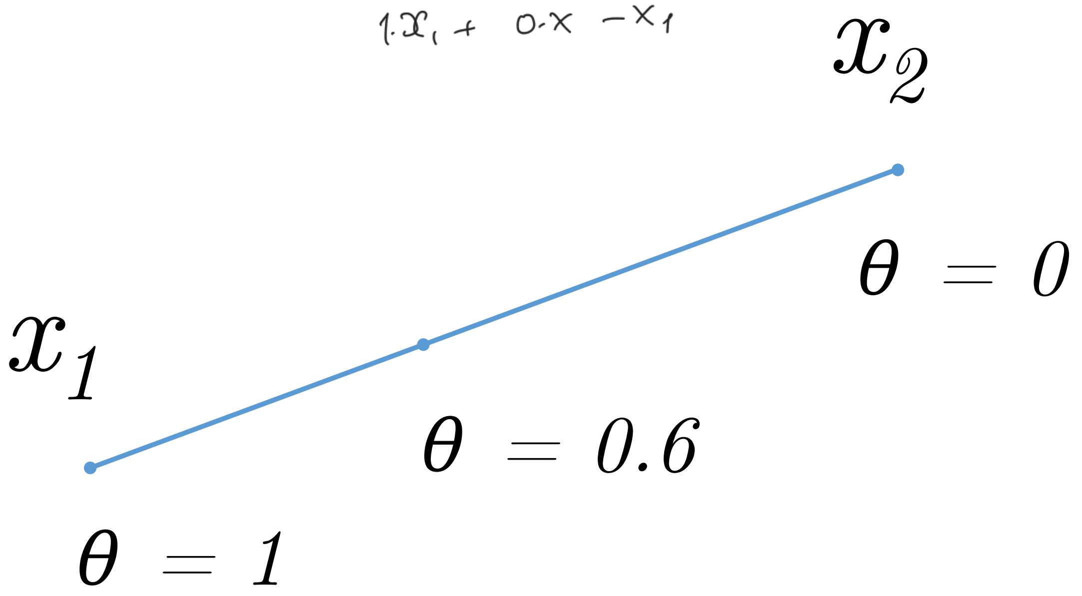
Line segment

Suppose x_1, x_2 are two points in $\mathbb{R}^n$. Then the line segment between them is defined as follows:

$$x = \theta x_1 + (1-\theta)x_2, \ \theta \in [0,1]$$

$$0 + x_2 = x_2$$

$$1 \cdot x_1 + 0 \cdot x - x_1$$



Convex set

The set S is called **convex** if for any x_1, x_2 from S the line segment between them also lies in S, i.e.

$$\forall \theta \in [0,1], \ \forall x_1, x_2 \in S: \\ \theta x_1 + (1-\theta)x_2 \in S \ \mapsto \ S\text{ - вып. мн.}$$

Examples:

* Any affine set
* Ray
* Line segment

Аффинное мн-во: $\forall x_1, x_2 \in S: \ \theta x_1 + (1-\theta)x_2 \in S$

мн-во решений: $Ax = b$ $\qquad \mathbb{R}^2, \mathbb{R}^3, \mathbb{R}^4$

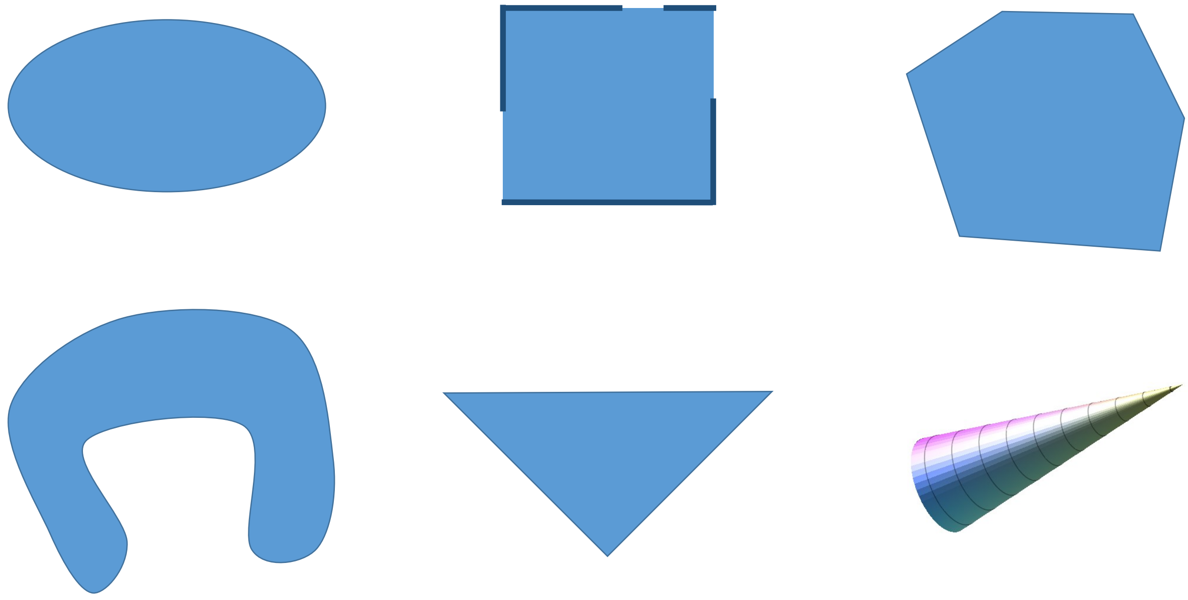
$$A \in \mathbb{R}^{m \times n}$$

$$x \in \mathbb{R}^n \quad b \in \mathbb{R}^m \quad m < n$$

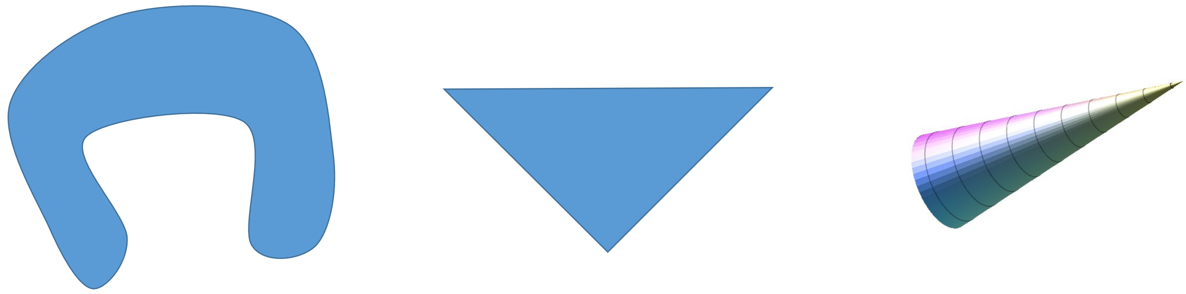
$$Ax_1 = b \qquad Ax_2 = b$$

$$A(\theta x_1 + (1-\theta)x_2) = \theta \underbrace{Ax_1}_{b} + (1-\theta)\underbrace{Ax_2}_{b} =$$

$$= \theta b + (1-\theta)b = b$$

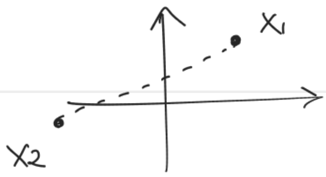


BRO NOT BRO BRO



NOT BRO BRO BRO

Related definitions



conv(x1, x2)

Convex combination

Let $x_1, x_2, \ldots, x_k \in S$, then the point $\theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_k x_k$ is called the convex combination of points $x_1, x_2, \ldots, x_k$ if $\boxed{\sum_{i=1}^{k} \theta_i = 1, \; \theta_i \geq 0}$

выпуклая комбин.

$\boxed{\sum_{i=1}^{k} \theta_i = 1 \text{ - афинная комбин.}}$

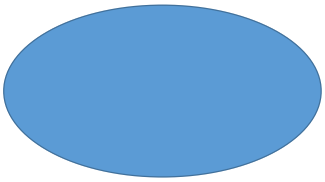
Convex hull

The set of all convex combinations of points from S is called the convex hull of the set S.

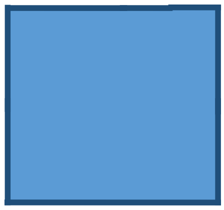
$$
\mathbf{conv}(S) = \left\{ \sum_{i=1}^{k} \theta_i x_i \mid x_i \in S, \sum_{i=1}^{k} \theta_i = 1, \; \theta_i \geq 0 \right\}
$$

- The set $\mathbf{conv}(S)$ is the smallest convex set containing S.
- The set S is convex if and only if $S = \mathbf{conv}(S)$.

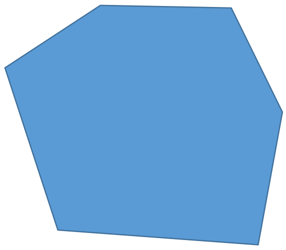
Examples:



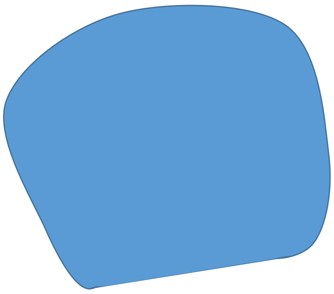
BRO



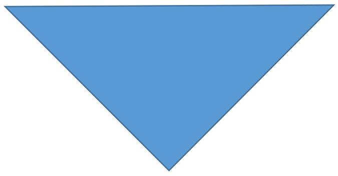
BRO



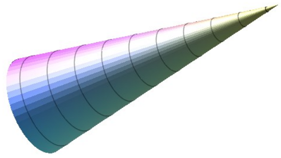
BRO



BRO



BRO



BRO

$\min_{x \in S} f(x)$

Finding convexity

In practice it is very important to understand whether a specific set is convex or not. Two approaches are used for this depending on the context.

- By definition.
- Show that S is derived from simple convex sets using operations that preserve convexity.

By definition

$$
x_1, x_2 \in S,\ 0 \leq \theta \leq 1 \quad \rightarrow \quad \theta x_1 + (1-\theta)x_2 \in S
$$

Preserving convexity

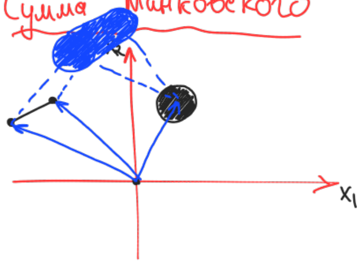
The linear combination of convex sets is convex

$\alpha S_x + \beta S_y$ - выпукло

Let there be 2 convex sets S_x, S_y, let the set $S = \{s \mid s = c_1 x + c_2 y,\ x \in S_x,\ y \in S_y,\ c_1, c_2 \in \mathbb{R}\}$

Take two points from S: $s_1 = c_1 x_1 + c_2 y_1, s_2 = c_1 x_2 + c_2 y_2$ and prove that the segment between them $${theta s_1 + (1 - \theta)s_2, \theta \in [0,1]$$ also belongs to $$S$$

Сумма Минковского



$$
\theta s_1 + (1-\theta)s_2
$$

$$
\theta(c_1 x_1 + c_2 y_1) + (1-\theta)(c_1 x_2 + c_2 y_2)
$$

$$
c_1 \underbrace{(\theta x_1 + (1-\theta)x_2)}_{x} + c_2 \underbrace{(\theta y_1 + (1-\theta)y_2)}_{y}
$$

$$
c_1 x + c_2 y \in S
$$

The intersection of any (!) number of convex sets is convex

If the desired intersection is empty or contains one point, the property is proved by definition. Otherwise, take 2 points and a segment between them. These points must lie in all intersecting sets, and since they are all convex, the segment between them lies in all sets and, therefore, in their intersection.

The image of the convex set under affine mapping is convex

$$S \subseteq \mathbb{R}^n \text{ convex} \rightarrow f(S) = \{f(x) \mid x \in S\} \text{ convex} \quad (f(x) = \mathbf{A}x + \mathbf{b})$$

Examples of affine functions: extension, projection, transposition, set of solutions of linear matrix inequality $\{x \mid x_1 A_1 + \ldots + x_m A_m \preceq B\}$ Here $A_i, B \in \mathbf{S}^p$ are symmetric matrices $p \times p$.

Note also that the prototype of the convex set under affine mapping is also convex.

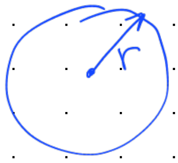
$$S \subseteq \mathbb{R}^m \text{ convex} \rightarrow f^{-1}(S) = \{x \in \mathbb{R}^n \mid f(x) \in S\} \text{ convex} \quad (f(x) = \mathbf{A}x + \mathbf{b})$$

Example 1

$S_{\shortparallel}$

Дано:

Prove, that ball in $\mathbb{R}^n$ (i.e. the following set $\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_c\| \leq r\}$) - is convex.



$x, y \in S$: $\|x - x_c\| \le r$ $\|y - x_c\| \le r$

доказать

$\|\theta x + (1-\theta)y - x_c\| \le r$

$\|\theta x + (1-\theta)y - (\theta + 1 - \theta)x_c\| = \|\theta(x - x_c) + (1-\theta)(y - x_c)\| \le$

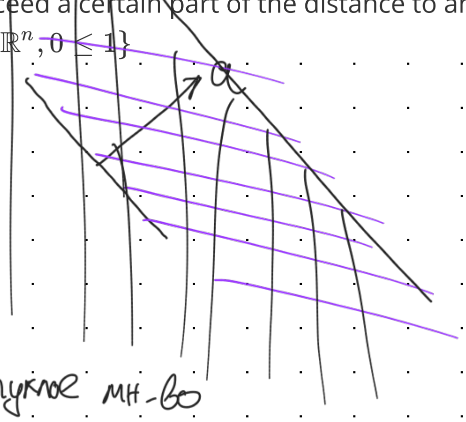
$\le \theta \underbrace{\|x - x_c\|}_{\le r} + (1-\theta)\underbrace{\|y - x_c\|}_{\le r} \le \theta r + (1-\theta) r = r$

Example 2

Which of the sets are convex: 1. Stripe, $\{x \in \mathbb{R}^n \mid \alpha \leq a^\top x \leq \beta\}$ 1. Rectangle, $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i = \overline{1,n}\}$ 1. Kleen, $\{x \in \mathbb{R}^n \mid a_1^\top x \leq b_1, a_2^\top x \leq b_2\}$ 1. A set of points closer to a given point than a given set that does not contain a point, $\{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \leq \|x - y\|_2, \forall y \in S \subseteq \mathbb{R}^n\}$ 1. A set of points, which are closer to one set than another, $\{x \in \mathbb{R}^n \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T), S, T \subseteq \mathbb{R}^n\}$ 1. A set of points, $\{x \in \mathbb{R}^n \mid x + X \subseteq S\}$, where $S \subseteq \mathbb{R}^n$ is convex and $X \subseteq \mathbb{R}^n$ is arbitrary. 1. A set of points whose distance to a given point does not exceed a certain part of the distance to another given point is $\{x \in \mathbb{R}^n \mid \|x - a\|_2 \leq \theta\|xb\|_2, a, b \in \mathbb{R}^n, 0 \leq 1\}$

$a^T x \le \beta$

$a^T x \ge \alpha$



$a^T x \ge \alpha$ – выпуклое мн-во

Example 3

Let $x \in \mathbb{R}$ is a random variable with a given probability distribution of $\mathbb{P}(x = a_i) = p_i$, where $i = 1, \ldots, n$, and $a_1 < \ldots < a_n$. It is said that the probability vector of outcomes of $p \in \mathbb{R}^n$ belongs to the probabilistic simplex, i.e.
$P = \{p \mid \mathbf{1}^T p = 1, p \succeq 0\} = \{p \mid p_1 + \ldots + p_n = 1, p_i \geq 0\}$. Determine if the following sets of p are convex: 1. $\alpha < \mathbb{E} f(x) < \beta$, where $\mathbb{E} f(x)$ stands for expected value of $f(x) : \mathbb{R} \to \mathbb{R}$, i.e. $\mathbb{E} f(x) = \sum\limits_{i=1}^{n} p_i f(a_i)$ 1. $\mathbb{E} x^2 \leq \alpha$ 1. $\mathbb{V} x \leq \alpha$

$$\mathbb{E} f(x) < \beta = \boxed{\sum_{i=1}^{n} p_i f(x_i) < \beta}$$

$p^T \cdot f(x) < \beta$

c

$c^T p < \beta$

полуплоскосів

$$\mathbb{V} x = \mathbb{E}(x^2) - (\mathbb{E} x)^2 =$$

$= \ldots$

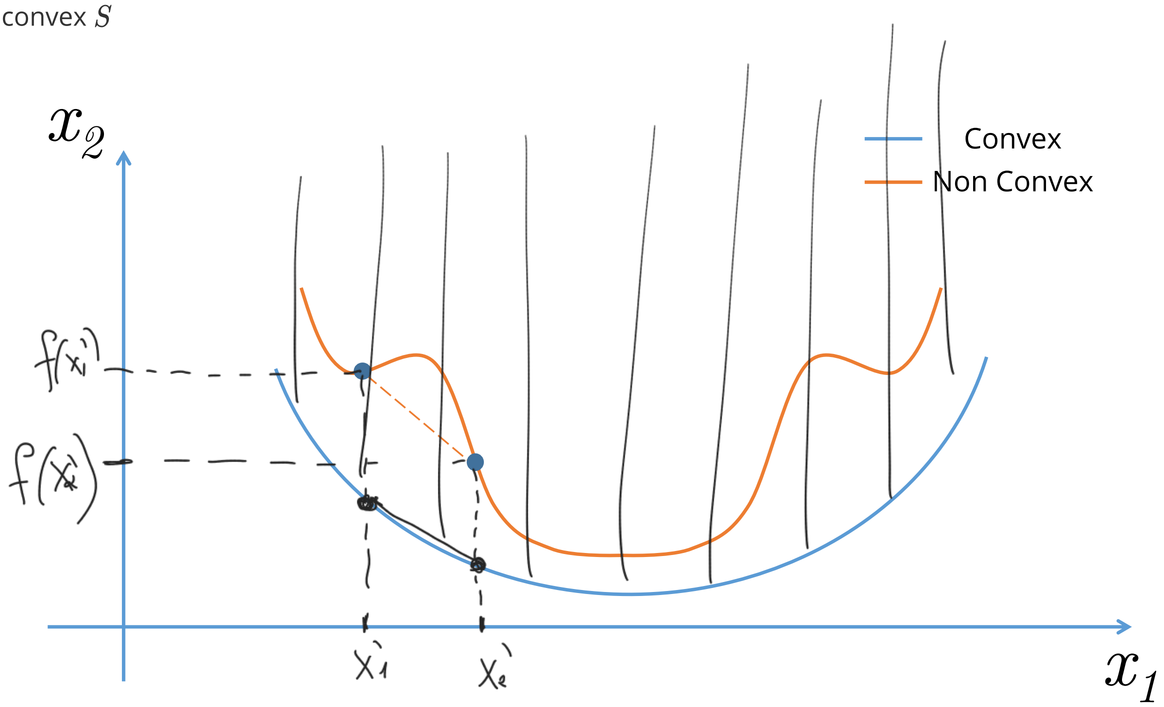
Convex function

Convex function

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



Examples

$x \succeq 0$

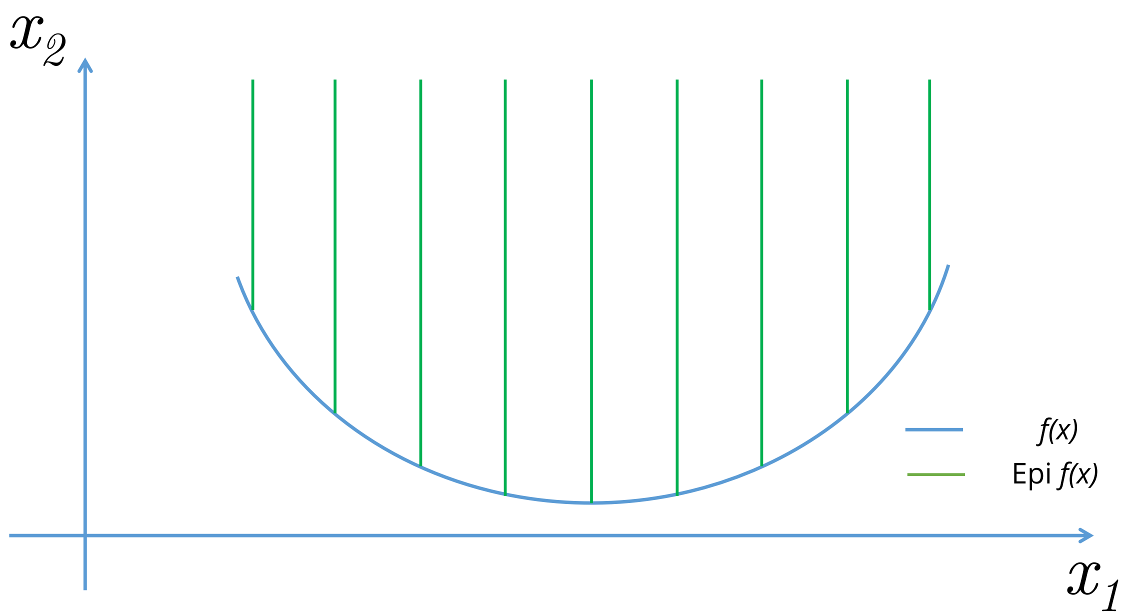
- $f(x) = x^p, p > 1, \quad S = \mathbb{R}_+$
- $f(x) = \|x\|^p, \quad p > 1, S = \mathbb{R}$
- $f(x) = e^{cx}, \quad c \in \mathbb{R}, S = \mathbb{R}$
- $f(x) = -\ln x, \quad S = \mathbb{R}_{++}$ — $x > 0$
- $f(x) = x \ln x, \quad S = \mathbb{R}_{++}$
- The sum of the largest k coordinates $f(x) = x_{(1)} + \ldots + x_{(k)}, \quad S = \mathbb{R}^n$
- $f(X) = \lambda_{max}(X), \quad X = X^T$
- $f(X) = -\log \det X, \quad S = S^n_{++}$

Epigraph

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 $f(x)$

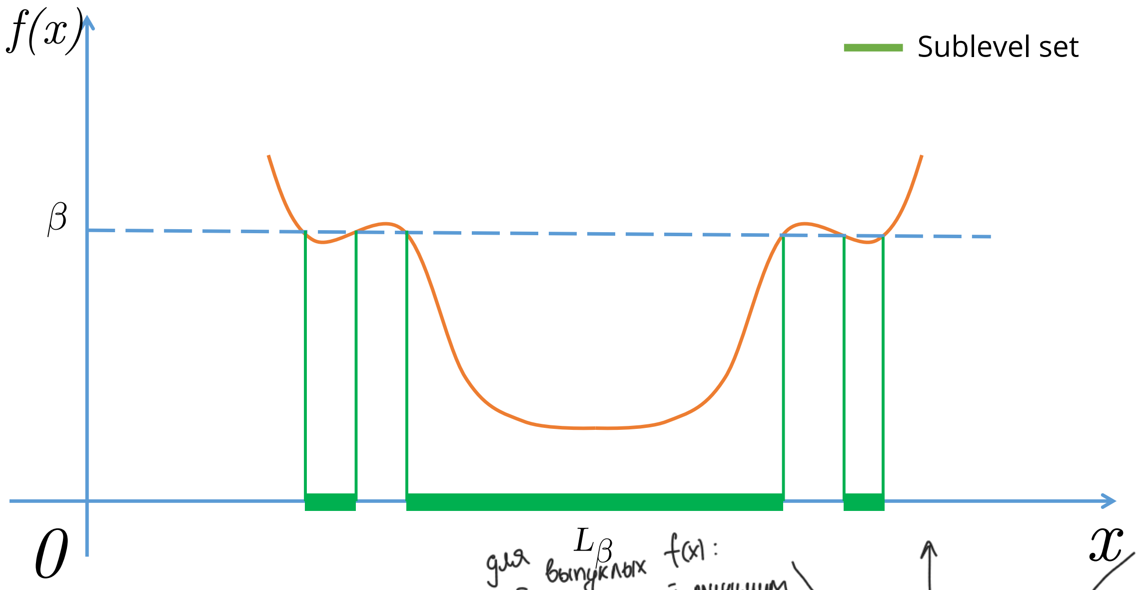


Sublevel set

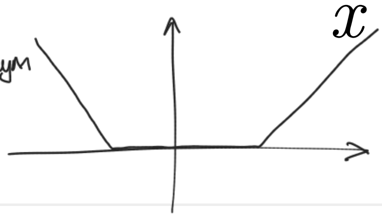
For the function $f(x)$, defined on $S \subseteq \mathbb{R}^n$, the following set:

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

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



для выпуклых f(x):
любой локальный минимум
является глобальным



Criteria of convexity

First order differential criterion of convexity

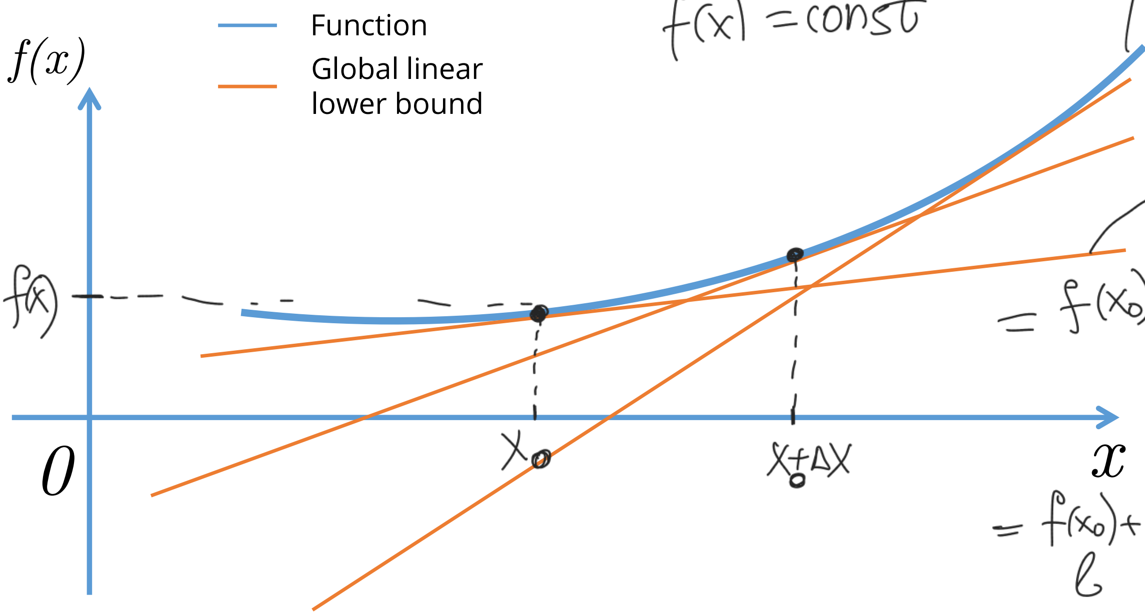
The differentiable function $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is convex if and only if $\forall x, y \in S$:

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

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

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

f(x) = const



$f^I_{x_0}(x) =$

$= f(x_0) + \nabla f^T(x_0)(x - x_0)$

$= f(x_0) + \nabla f^T(x_0) \cdot \Delta x$

b k 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 $\forall x \in \textbf{int}(S) \neq \emptyset$:

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

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 convex set.

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.

Reduction to a line

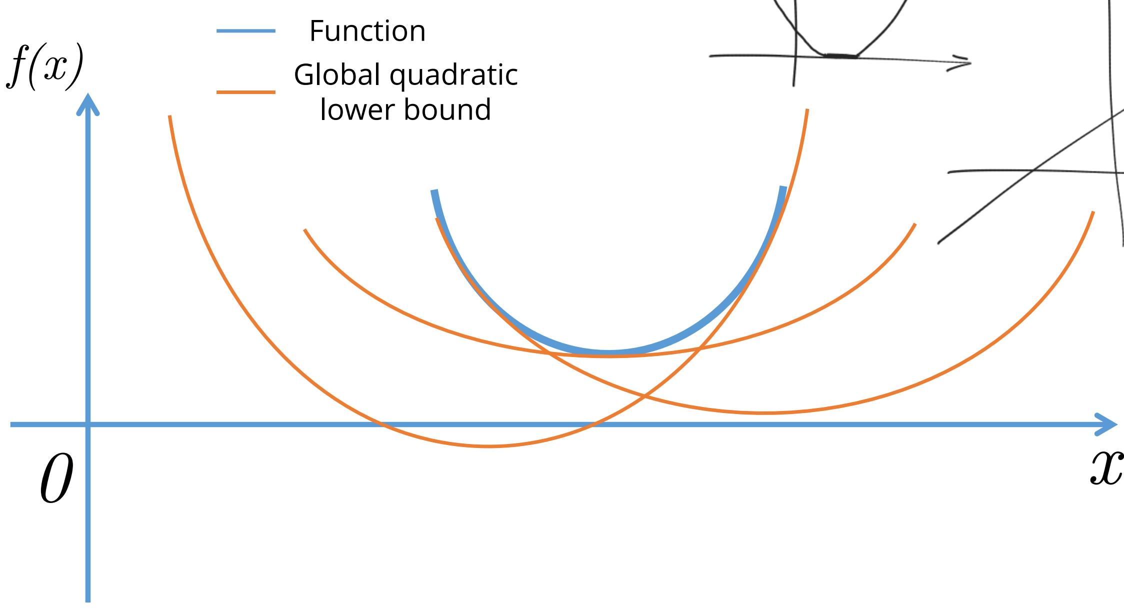
$f : S \to \mathbb{R}$ is convex if and only if S is convex set and the function $g(t) = f(x + tv)$ defined on $\{t \mid x + tv \in S\}$ is convex for any $x \in S, v \in \mathbb{R}^n$, which allows to check convexity of the scalar function in order to establish covexity of the vector function.

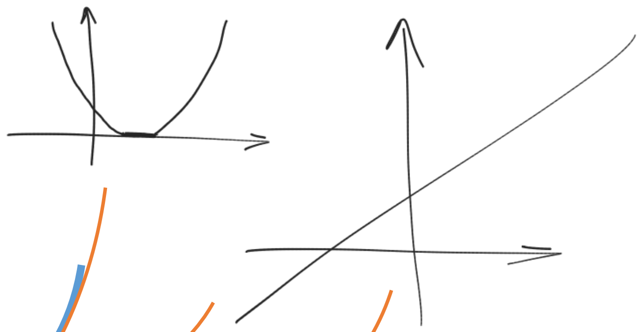
Strong convexity

$f(x)$, **defined on the convex set** $S \subseteq \mathbb{R}^n$, is called μ-strongly convex (strogly 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 $\mu > 0$.





Criteria of strong convexity

First order differential criterion of strong convexity

Differentiable $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ μ-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 μ-strongly convex if and only if $\forall x \in \mathbf{int}(S) \neq \emptyset$:

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

$\nabla^2 f - \mu I \succeq 0$

In other words:

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

$\exists \mu > 0$

Facts

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

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

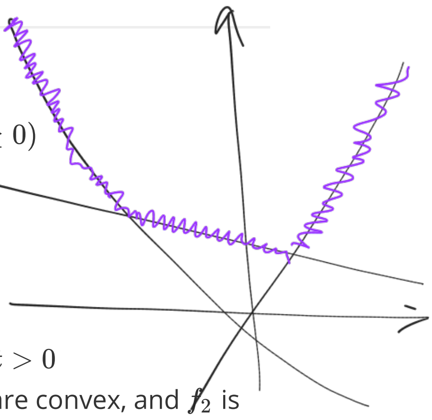
  for $\alpha_i \geq 0; \quad \sum\limits_{i=1}^n \alpha_i = 1$ (probability simplex)

  For the infinite dimension case:

  $$f\left(\int\limits_S x p(x)dx\right) \leq \int\limits_S f(x)p(x)dx$$

  If the integrals exist and $p(x) \geq 0, \quad \int\limits_S p(x)dx = 1$
- If the function $f(x)$ and the set S are convex, then any local minimum $x^* = \arg\min\limits_{x \in S} f(x)$ will be the global one. Strong convexity guarantees the uniqueness of the solution.

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 $f_1(x), \ldots, f_m(x)$ 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 $y \in Y$: $g(x) = \sup\limits_{y \in Y} f(x,y)$ 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 $f_1 : S_1 \to \mathbb{R}$ 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

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."

References

- Steven Boyd lectures
- Suvrit Sra lectures
- Martin Jaggi lectures

Example 4

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

Example 5

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

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

Example 6

Show, that $f(x)$ is convex, using first and second order criteria, if $f(x) = \sum\limits_{i=1}^{n} x_i^4$.

$$\nabla^2 f = \operatorname{diag}\left(12x^2\right)$$

Example 7

Find the set of $x \in \mathbb{R}^n$, where the function $f(x) = \dfrac{-1}{2(1 + x^\top x)}$ is convex, strictly convex, strongly convex?