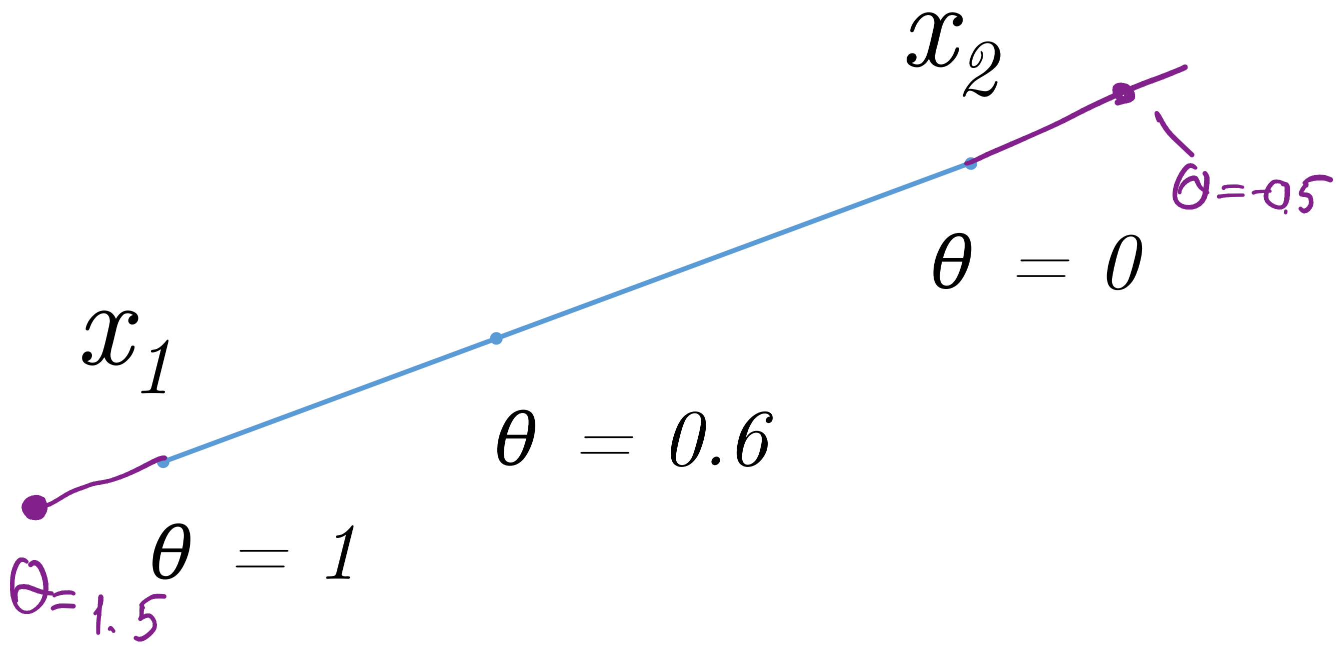
Convex set

Line segment Отрезок

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

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



Convex set

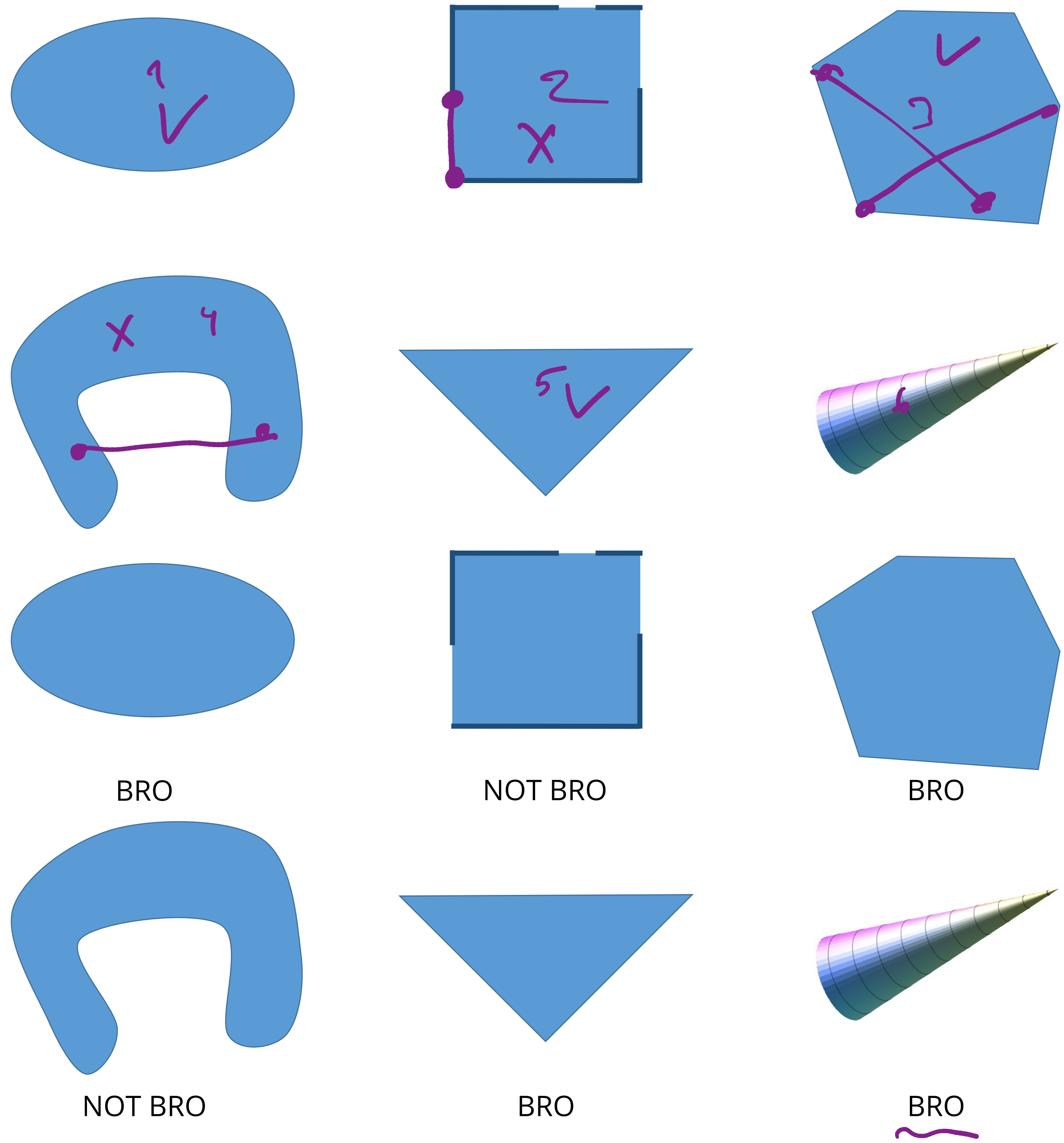
The set S is called **convex** if for any x_1, x_2 from S the line segment between them also lies in S, i.e.

S - вып. мн-во

$$\boxed{\begin{gathered}\forall \theta \in [0, 1], \; \forall x_1, x_2 \in S: \\ \theta x_1 + (1 - \theta)x_2 \in S\end{gathered}}$$

Examples:

- Any affine set
- Ray
- Line segment



Related definitions

ТОЧКА

Convex combination

lето k точк $x_1 \ldots x_k \in \mathbb{R}^n$

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

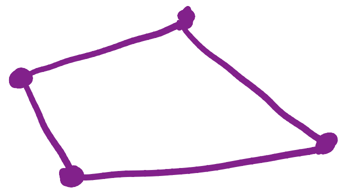
МНОЖЕСТВО

Convex hull

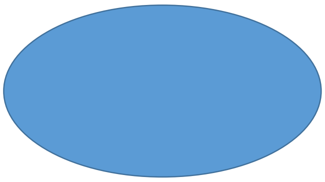
The set of all convex combinations of points from S is called the convex hull of the set S.

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

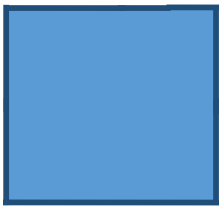
- The set $\mathbf{conv}(S)$ is the smallest convex set containing S. ← !
- The set S is convex if and only if $S = \mathbf{conv}(S)$.



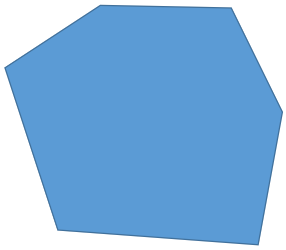
Examples:



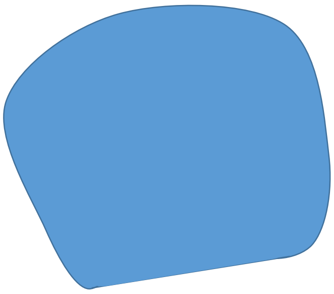
BRO



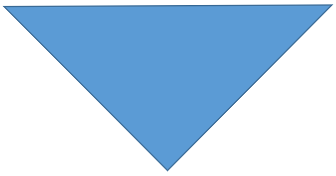
BRO



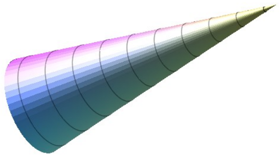
BRO



BRO



BRO



BRO

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.

1 By definition

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

$\varnothing$ – выпуклое по опр.

2 Preserving convexity

2.2 The linear combination of convex sets is convex

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 $$S$$

$S = c_1 \cdot S_x + c_2 \cdot S_y$
сумма Минковского

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

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

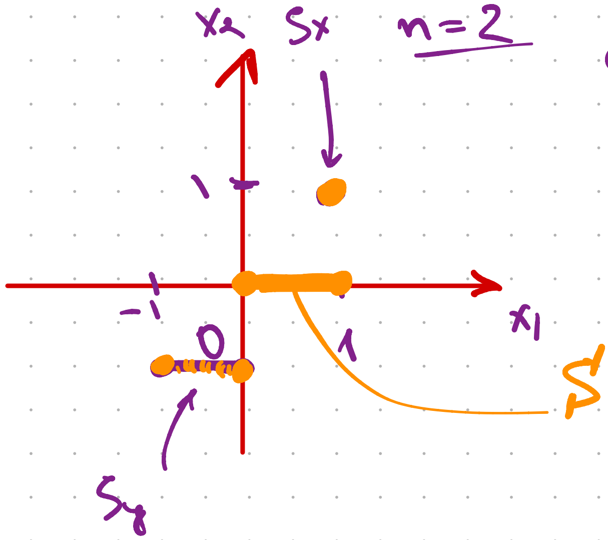
$$c_1(\theta x_1 + (1-\theta)x_2) + c_2(\theta y_1 + (1-\theta)y_2)$$

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

2.1 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.

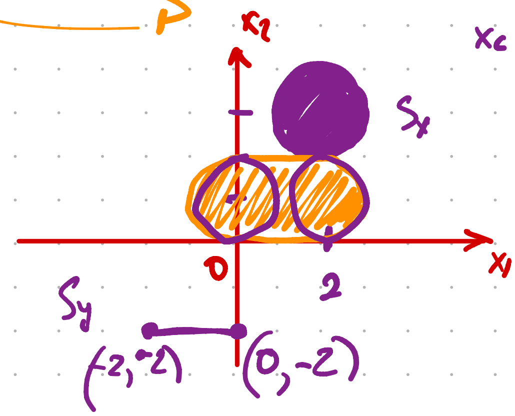
$n=2$



$$S = S_x + S_y$$

$$\Rightarrow S = \{ s \in \mathbb{R}^n : s = x + y \mid x \in S_x,\ y \in S_y \}$$

x_c $(2,3)$ $r = 1$



2.3

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

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

Решение:

1) Беру $x_1 \in S$; $x_2 \in S$ $\quad \|x_1 - x_c\| \le r$; $\|x_2 - x_c\| \le r$

$0 \le \theta \le 1$

2) Строю $x = \theta \cdot x_1 + (1-\theta) x_2$

3) $\underbrace{x \in S: \ \|x - x_c\| \le r}_{\text{доказать}}$

$\|\theta x_1 + (1-\theta) x_2 - x_c\|$

$\|\theta x_1 + (1-\theta) x_2 - \theta x_c - (1-\theta) x_c\|$

$-\theta + \theta = 0$

$\|\theta (x_1 - x_c) + (1-\theta)(x_2 - x_c)\| \le$

$\le \theta \cdot \underbrace{\|x_1 - x_c\|}_{\le r} + (1-\theta) \underbrace{\|x_2 - x_c\|}_{\le r} \le \theta r + (1-\theta) r \le r.$

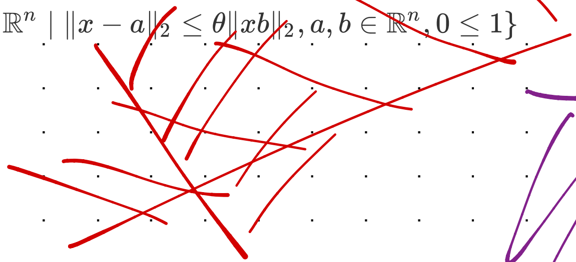
ч.т.д.

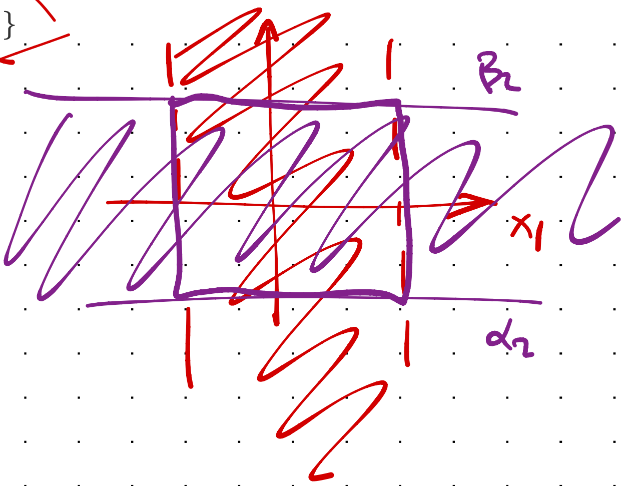
$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$ i-ая позиция

$\tilde{a}_i^T x \le \beta_i$

Example 2

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





Полуплоскость: $S = \{x \in \mathbb{R}^n \mid x^T a \leq b\}$

$a^T x \leq b$

$x_1: \quad a^T x_1 \leq b$

$a^T(\theta x_1 + (1-\theta) x_2)$

$x_2: \quad a^T x_2 \leq b$

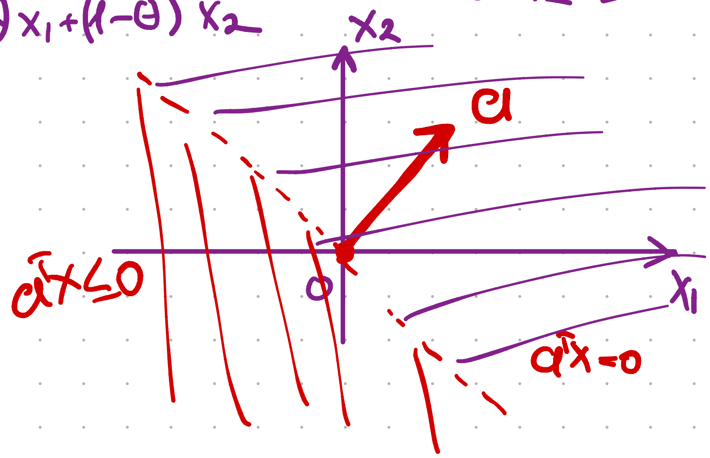
$a^T x \leq b$

$\theta \underset{\wedge\, b}{a^T x_1} + (1-\theta) \underset{\wedge\, b}{a^T x_2}$

$x: \quad \theta x_1 + (1-\theta) x_2$

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

$a^T x \geq 0$



пример:

$$S = \{X \in \mathbb{R}^{n \times n} \mid tr(X) \geq 228\}$$

$$tr(I^T X) \geq 228$$

$$\langle I, X \rangle \geq 228$$

x_i	1	2	...	21	22	23
p_i	$\frac{1}{100}$	$\frac{2}{100}$		$\frac{3}{100}$	$\frac{3}{100}$	$\frac{2}{100}$

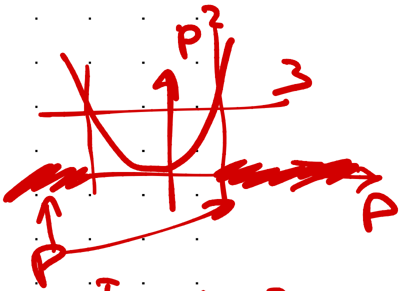
Example 3

Let $x \in \mathbb{R}$ is a random variable with a given probability distribution of $\boxed{\mathbb{P}(x = a_i) = p_i}$, 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. $\mathbb{V}x \leq \alpha$

Пример:

Вероятностный симплекс:

$$\begin{cases} \mathbf{1}^T \cdot p = 1 \\ p \succeq 0 \end{cases} \qquad \begin{cases} \sum_i p_i = 1 \\ p_i \geq 0 \end{cases}$$



$p \in \mathbb{R}^n$

?

$\mathbb{E}x \geq 10$

Выпуклое

$\sum p_i \cdot x_i \geq 10$

$x^T p \geq 10$

$p^T p \geq 3$

$p^2 \geq 3$

$\boxed{\mathbb{E}f(x) \geq 100}$

$\sum\limits_{i=1}^{n} p_i \cdot f(x_i) \geq 100$

$\vec{f}(x)^T \cdot p \geq 100$

Convex function

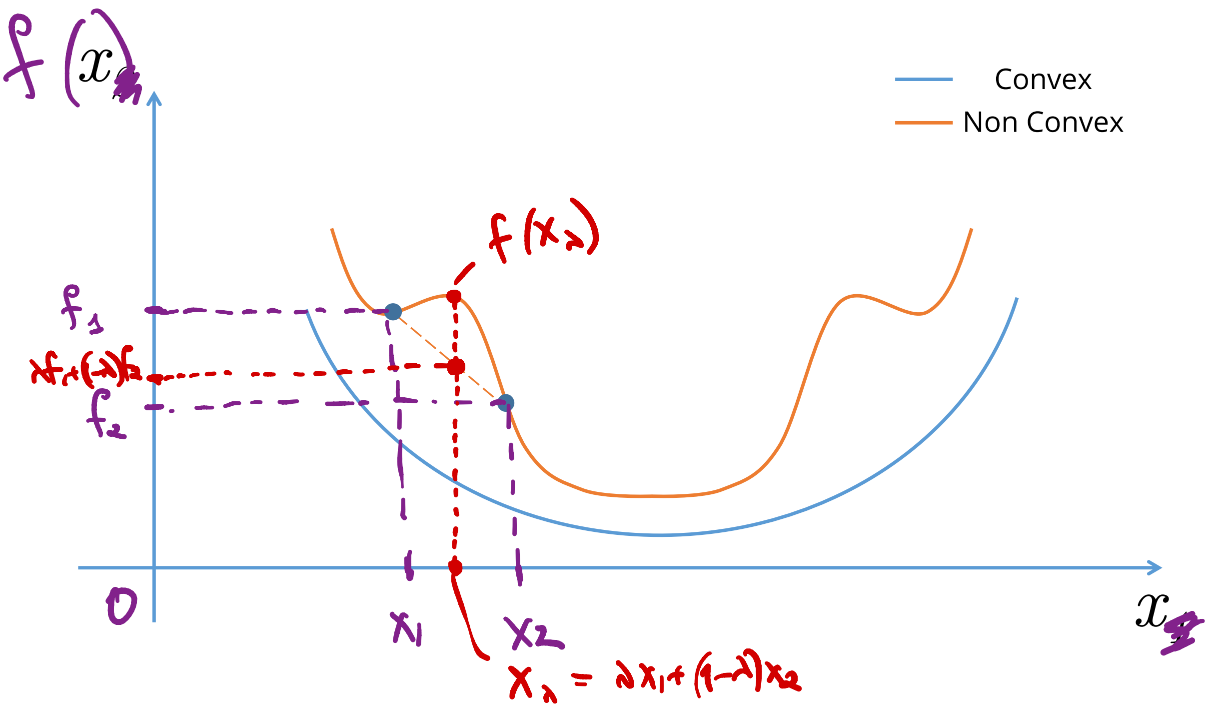
Convex function

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

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

неравенство Йенсена

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

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

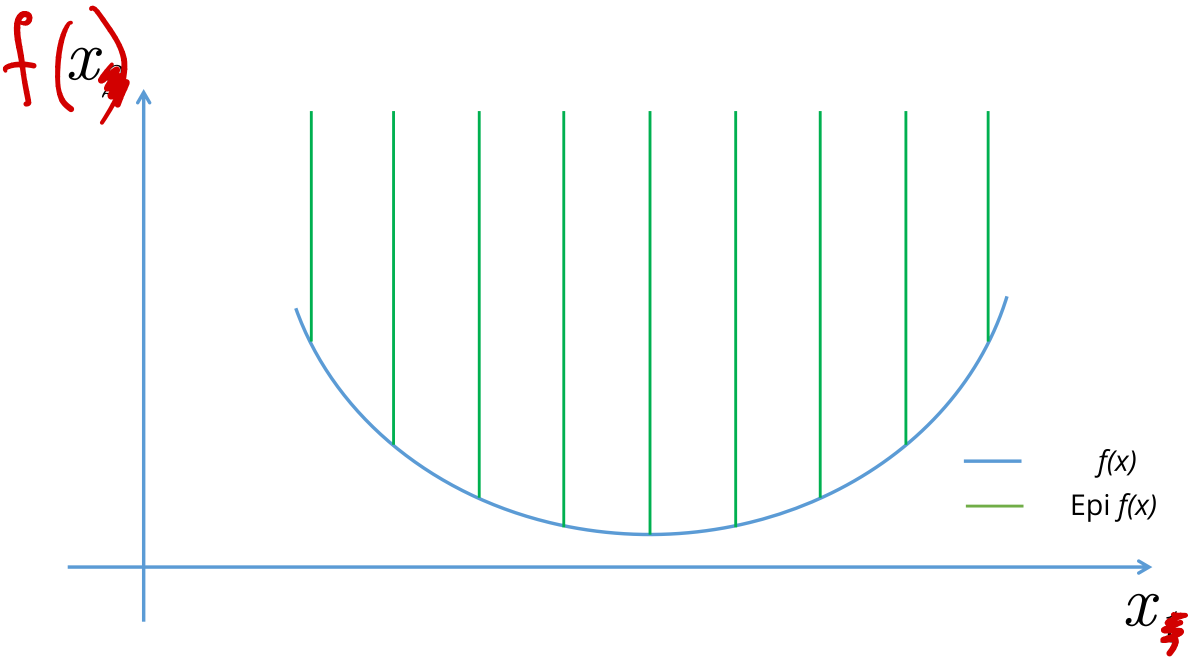
Epigraph

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

Надграфик

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

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

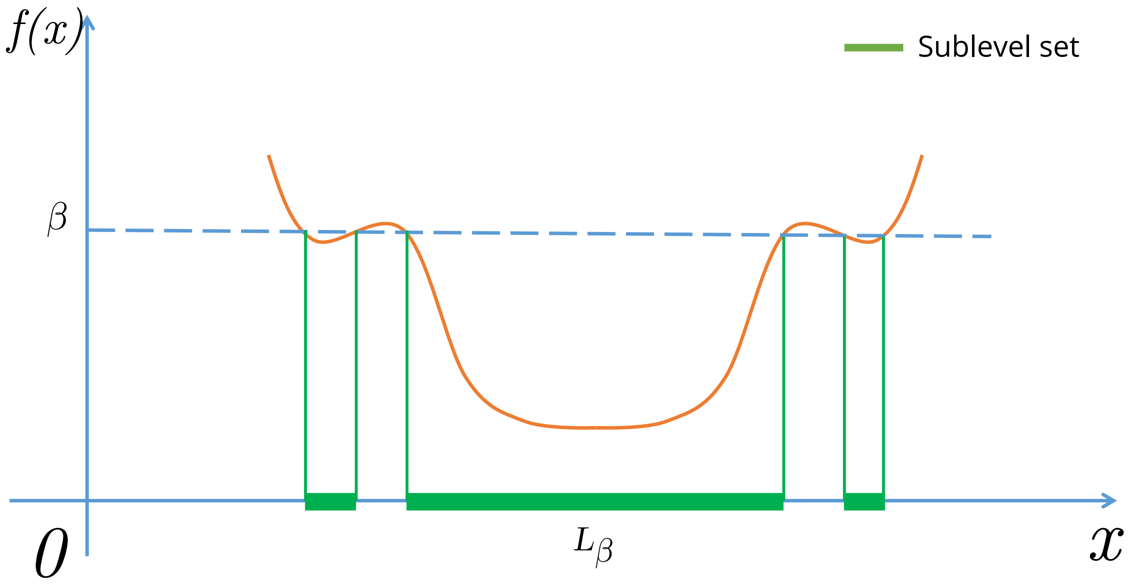


Sublevel set

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

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

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



Criteria of convexity

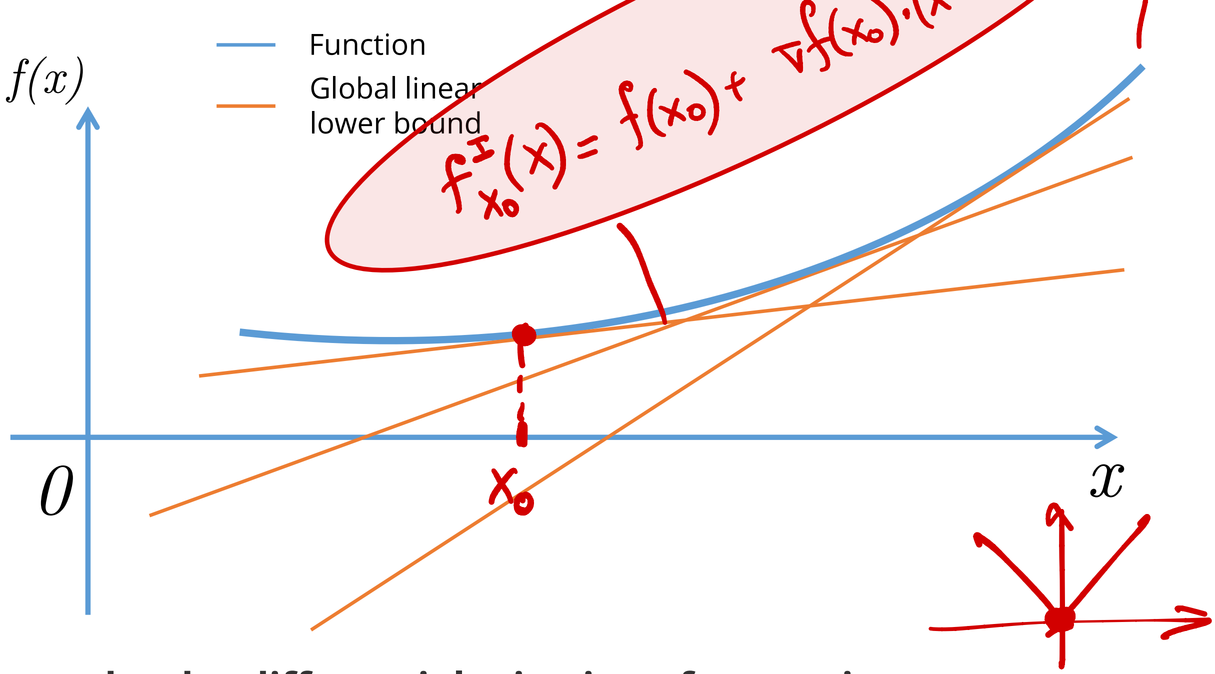
First order differential criterion of convexity

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

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

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 $\forall x \in \mathbf{int}(S) \neq \emptyset$:

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

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

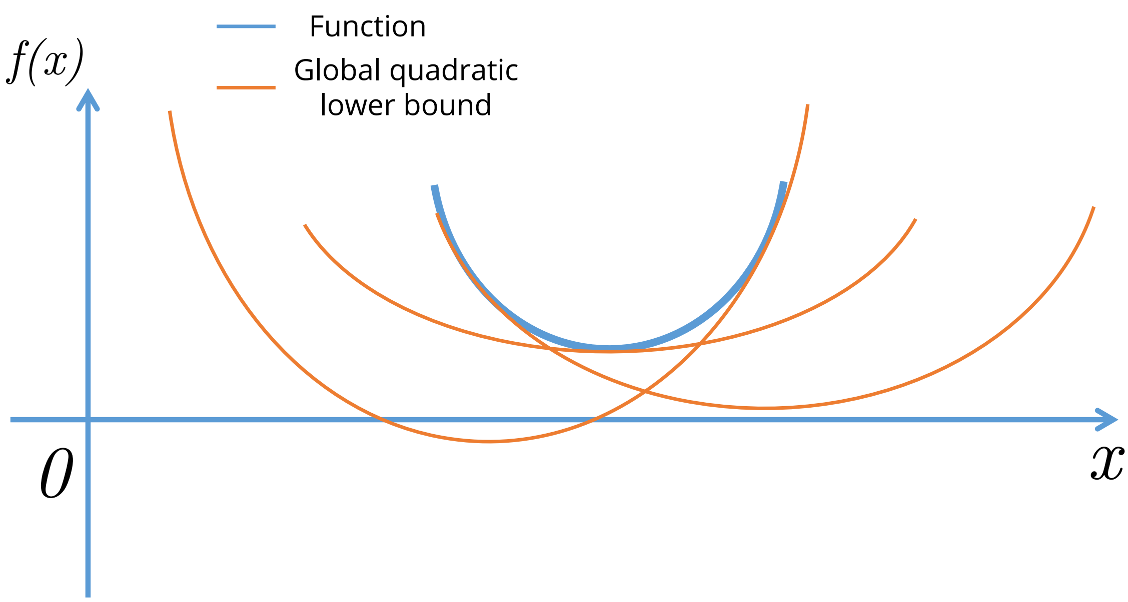
$f : S \to \mathbb{R}$ is convex if and only if S is convex set and the function $g(t) = f(x + tv)$ defined on $\{t \mid x + tv \in S\}$ is convex for any $x \in S, v \in \mathbb{R}^n$, which allows to check convexity of the scalar function in order to establish covexity of the vector function.

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 $\forall x \in \mathbf{int}(S) \neq \emptyset$:

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

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 $-f(x)$ - (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 $f_1(x), \dots, f_m(x)$ 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 $y \in Y$: $g(x) = \sup\limits_{y \in Y} f(x, y)$ 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 $\operatorname{range}(f_1) \subseteq S_2$. 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 $x_1, \dots, x_n$
- 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: $\langle \nabla f(y), x - y\rangle \geq 0 \longrightarrow f(x) \geq f(y)$

- Discrete convexity: $f : \mathbb{Z}^n \to \mathbb{Z}$; "convexity + matroid theory."

References

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

$-f(x)$ – выпуклая
$f(x)$ – вогнутая

Example 4

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

1) $\nabla f = c \qquad df = \langle c, dx \rangle$

2) $\nabla^2 f = 0^{n\times n} \succeq 0 \to f$ - выпуклая

$\nabla^2(-f) = 0^{n\times n} \succeq 0$ – $-f$ - выпуклая
f - впуклая

Example 5

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

$\nabla f = (A + A^T)x = 2Ax$

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

Example 6

$f(x) = \|Ax - b\|^2$

$\nabla^2 f = A^T A \succeq 0$? $\quad x \in \mathbb{R}^n$

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

$x^T (A^T A) x \succeq 0$

$(Ax)^T Ax \succeq 0 \qquad \|Ax\| \succeq 0$

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

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?