1 - Preliminary notions of convex analysis

G. Mastroeni

Department of Computer Science, University of Pisa

Optimization Methods and Game Theory

Master of Science in Artificial Intelligence and Data Engineering
University of Pisa – A.Y. 2023/24

Contents of the lessons

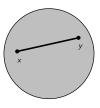
- Convex sets
- Cones
- Convex functions
- Strictly and strongly convex functions

Convex sets

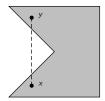
Definition (Convex set)

A set $C \subseteq \mathbb{R}^n$ is convex if, for every $x, y \in C$ and for every $\alpha \in [0, 1]$,

$$\alpha x + (1 - \alpha)y \in C$$
.



convex set



non-convex set

Examples of convex sets: affine sets

Definition (Affine set)

A set $C \subseteq \mathbb{R}^n$ is affine if, for every $x, y \in C$ and every $\alpha \in \mathbb{R}$,

$$\alpha x + (1 - \alpha)y \in C$$
.

Examples of affine sets:

- any single point {x}
- any line
- the solution set of a system of linear equations

$$C = \{x \in \mathbb{R}^n : Ax = b\},\$$

where A is a $m \times n$ matrix and $b \in \mathbb{R}^m$

any subspace

Examples of convex sets: subspaces

Note that a subspace is a particular affine set.

In fact, a set $S \subseteq \mathbb{R}^n$ is a subspace if, for every $x, y \in S$ and every $\alpha, \beta \in \mathbb{R}$,

$$\alpha x + \beta y \in S$$

Examples of subspaces:

- {0}
- any line which passes through zero
- the solution set of a homogeneous system of linear equations

$$S = \{x \in \mathbb{R}^n : Ax = 0\},\$$

where A is a $m \times n$ matrix.

Definition

A convex combination of the points $x^1, x^2,, x^k$ is a point

$$y = \sum_{i=1}^k \alpha_i x^i$$
 where $\alpha_1, \dots, \alpha_k \in [0, 1]$ and $\sum_{i=1}^k \alpha_i = 1$.

Remark

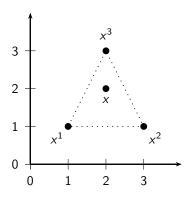
By definition, a set $C \subseteq \mathbb{R}^n$ is convex if it contains all the convex combinations of any two points in C.

Example. Consider the following 3 points in the plane:

$$x^1 = (1, 1),$$
 $x^2 = (3, 1),$ $x^3 = (2, 3).$

x = (2,2) is a convex combination of x^1 , x^2 e x^3 , in fact:

$$x = \frac{1}{4}x^1 + \frac{1}{4}x^2 + \frac{1}{2}x^3.$$



A convex set contains any convex combination of its points.

Lemma 1

If C is convex, then for any $x^1, \ldots, x^k \in C$ and $\alpha_1, \ldots, \alpha_k \in [0, 1]$ s.t. $\sum_{i=1}^k \alpha_i = 1$,

$$\sum_{i=1}^k \alpha_i x^i \in C.$$

Proof. By induction on k. For k=2, the thesis holds, by definition of convexity. Assume that the thesis holds for a given k and let us prove it holds for k+1.

Let $x^1, \ldots, x^{k+1} \in C$ and $\alpha_1, \ldots, \alpha_{k+1} \in [0,1]$ s.t. $\sum_{i=1}^{k+1} \alpha_i = 1$. With no loss of generality, we assume that $\alpha_1 \neq 0$.

$$\sum_{i=1}^{k+1} \alpha_i x^i = \alpha_1 x_1 + \sum_{i=2}^{k+1} \alpha_i x^i = \alpha_1 x_1 + \left(1 - \sum_{i=2}^{k+1} \alpha_i\right) \sum_{i=2}^{k+1} \frac{\alpha_i}{1 - \sum_{i=2}^{k+1} \alpha_i} x^i$$

Since $\sum_{i=2}^{k+1} \frac{\alpha_i}{1-\sum_{i=1}^{k+1} \alpha_i} = 1$, by inductive assumption we have:

$$\bar{x} := \sum_{i=2}^{k+1} \frac{\alpha_i}{1 - \sum_{i=2}^{k+1} \alpha_i} x^i \in C$$

and finally, since C is convex,

$$\alpha_1 x_1 + (1 - \sum_{i=2}^{k+1} \alpha_i) \bar{x} \in C.$$

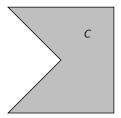
Proposition

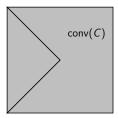
If $\{C_i\}_{i\in I}$ is any (possibly infinite) family of convex sets, then $\bigcap_{i\in I} C_i$ is convex.

Definition (Convex hull)

The convex hull conv(C) of a set C is the intersection of all the convex sets containing C.

In other words, it is the smallest convex set containing C.

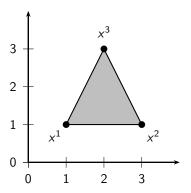




The convex hull of the points

$$x^1 = (1, 1),$$
 $x^2 = (3, 1),$ $x^3 = (2, 3).$

is the grey triangle with vertexes the three points:



Proposition

$$conv(C) = \{all convex combinations of points in C\}$$

Proof. It can be proved that the set of convex combinations of points in C is a convex set containing C, so that

$$conv(C) \subseteq \{all\ convex\ combinations\ of\ points\ in\ C\}.$$

Since $C \subseteq \text{conv}(C)$ and conv(C) is convex, by Lemma 1 it contains any convex combination of its points, and therefore

 $conv(C) \supseteq \{all convex combinations of points in C\}.$

Remark

Observe that C is convex if and only if C = conv(C).

Examples of convex sets: Polyhedra

Definition (Polyhedron)

A polyhedron P is the intersection of a finite number of closed halspaces in \mathbb{R}^n .

A closed halfspace is the set of solutions of a linear inequality:

$$a^{\mathsf{T}}x \leq \beta$$
, where $a \in \mathbb{R}^n$ e $\beta \in \mathbb{R}$.

Consequently, a polyhedron

$$P = \{x \in \mathbb{R}^n : Ax \le b\}$$

is the solution set of a system of linear inequalities where A is a $m \times n$ matrix and $b \in \mathbb{R}^m$.

A polyhedron P is convex since any closed halfspace is a convex set and the intersection of convex sets is convex.

Examples of convex sets: Balls

• A ball is defined by $B(\bar{x},r) := \{z \in \mathbb{R}^n : ||z - \bar{x}|| \le r\}$, where $||\cdot||$ is any norm, e.g.

$$\begin{split} \|x\|_2 &= \sqrt{\sum_{i=1}^n x_i^2} \text{ (Euclidean norm)} \\ \|x\|_1 &= \sum_{i=1}^n |x_i| \text{ (Manhattan distance)} \\ \|x\|_\infty &= \max_{i=1,\dots,n} |x_i| \text{ (Chebyshev norm)} \\ \|x\|_p &= \sqrt[p]{\sum_{i=1}^n |x_i|^p}, \text{ with } 1 \leq p < +\infty \\ \|x\|_A &= \sqrt{x^T A x}, \text{ where } A \text{ is a symmetric and positive definite matrix, i.e.,} \\ x^T A x > 0 \qquad \forall \ x \neq 0. \end{split}$$

Norms

Recall that a norm on a real vector space X is a function $p: X \to \mathbb{R}$ such that:

By the previous conditions it follows that $p(x) \ge 0$, $\forall x \in X$.

Exercise 1.1 Find the unit ball B(0,1) w.r.t. $\|\cdot\|_1$, $\|\cdot\|_\infty$ and $\|\cdot\|_A$, where

$$A = \left(\begin{array}{cc} 4 & 0 \\ 0 & 1 \end{array}\right).$$

Operations that preserve convexity

Algebraic operations

Sum and product by a constant

If C_1 and C_2 are convex, then $C_1 + C_2 := \{x + y : x \in C_1, y \in C_2\}$ is convex.

If C is convex and $\alpha \in \mathbb{R}$, then $\alpha C := \{\alpha x : x \in C, \}$ is convex.

Consequently, if C_1 and C_2 are convex, then $C_1 - C_2 := \{x - y : x \in C_1, y \in C_2\}$ is convex.

Topological operations

Closure and interior

If C is convex, then cl(C) is convex.

If C is convex, then int(C) is convex, provided that $int(C) \neq \emptyset$.

Relative interior

Given a set $C \subseteq \mathbb{R}^n$ we denote by aff(C) the smallest affine set containing C.

Definition (relative interior)

Let $C \subseteq \mathbb{R}^n$ be a convex set.

The relative interior of C is defined by

$$ri(C) = \{x \in C : \exists \epsilon > 0 \text{ s.t. aff}(C) \cap B(x, \epsilon) \subseteq C\}$$

Examples

• Let $C := \{(x_1, x_2) \in \mathbb{R}^2 : 1 \le x_1 \le 3, x_2 = 0\}$. Then

$$ri(C) := \{(x_1, x_2) \in \mathbb{R}^2 : 1 < x_1 < 3, x_2 = 0\}.$$

• Let $C = \{\bar{x}\}$, then ri(C) = C.

Theorem

Let C be a nonempty convex set in \mathbb{R}^n . Then the relative interior of C is a nonempty convex set.

Separation of convex sets

The sets A and B in \mathbb{R}^n are said to be linearly separable if there exists $a \in \mathbb{R}^n$, $a \neq 0$, $\beta \in \mathbb{R}$, such that

$$a^T x \ge \beta \quad \forall x \in A, \qquad a^T x \le \beta \quad \forall x \in B,$$

The separation is said to be proper if strict inequality holds for at least one $x \in A \cup B$.

Theorem

Let A, B be nonempty convex sets in \mathbb{R}^n . Then A and B are properly linearly separable if and only if

$$ri(A) \cap ri(B) = \emptyset$$
.

In particular two disjoint convex sets are always properly linearly separable.

Example Let
$$A := \{(x_1, x_2) \in \mathbb{R}^2 : 1 \le x_1 \le 2, x_2 = 0\},$$

 $B:=\{(x_1,x_2)\in\mathbb{R}^2:2\leq x_1\leq 4,x_2=0\}.$

Then $ri(A) \cap ri(B) = \emptyset$ and the sets are properly separable by the hyperplane of equation $x_1 = 2$.

Operations that preserve convexity

Affine functions

Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be affine, i.e. f(x) = Ax + b, with $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.

- If $C \subseteq \mathbb{R}^n$ is convex, then $f(C) = \{f(x) : x \in C\}$ is convex
- If $C \subseteq \mathbb{R}^m$ is convex, then $f^{-1}(C) = \{x \in \mathbb{R}^n : f(x) \in C\}$ is convex

Examples:

- $f(x) = \alpha x$, with $\alpha \in \mathbb{R}$
- f(x) = x + b, with $b \in \mathbb{R}^n$
- $\bullet \ f(x) = \left(\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right) x \text{, with } \theta \in (0,2\pi) \text{ (rotation)}$

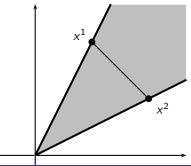
Definition (Cone)

A set $C \subseteq \mathbb{R}^n$ is a cone if, for every $x \in C$ and for every $\lambda \geq 0$, it results

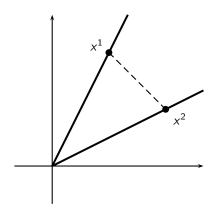
$$\lambda x \in C$$
.

In other words, if C contains a point x different from 0, then it contains the whole halfline starting from 0 and passing through x.

Example. A cone may be convex



or non convex:



Examples of cones

- \mathbb{R}^n_+ is a convex cone.
- $\{x \in \mathbb{R}^2 : x_1 x_2 = 0\}$ is a non-convex cone.
- Given a polyhedron $P = \{x : Ax \le b\}$, the recession cone of P is defined as

$$rec(P) := \{d: x + \alpha d \in P \text{ for any } x \in P, \alpha \ge 0\}.$$

It can be proved that $rec(P) = \{d : Ad \le 0\}$, thus it is a polyhedral cone.

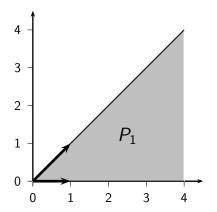
- $\left\{x \in \mathbb{R}^3 : x_3 \ge \sqrt{x_1^2 + x_2^2}\right\}$ is a non-polyhedral cone.
- Given $\bar{x} \in cl(C) \subseteq \mathbb{R}^n$, the set

$$T_C(\bar{x}) = \left\{ d \in \mathbb{R}^n : \exists \{z_k\} \subset C, \exists \{t_k\} > 0, \ z_k \to \bar{x}, \ t_k \to 0, \lim_{k \to \infty} \frac{z_k - \bar{x}}{t_k} = d \right\}$$

is called the *tangent cone* to C at \bar{x} .

Example

$$P_1 = \{x \in \mathbb{R}^2: \ x_2 \leq x_1, \quad x_2 \geq 0\}$$

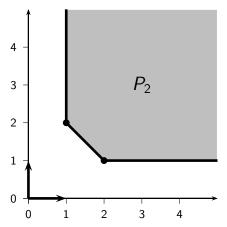


is a polyhedral cone.

$$rec(P_1) = P_1, \quad T_{P_1}((0,0)) = P_1.$$

Example

$$P_2 = \{x \in \mathbb{R}^2: \ x_1 \geq 1, \quad x_2 \geq 1, \quad x_1 + x_2 \geq 3\}$$



$$rec(P_2) = \{ d \in \mathbb{R}^2 : d_1 \ge 0, \quad d_2 \ge 0 \}$$

$$T_{P_2}((1,2)) = \{ d \in \mathbb{R}^2 : d_1 \ge 0, \quad d_1 + d_2 \ge 0 \}$$

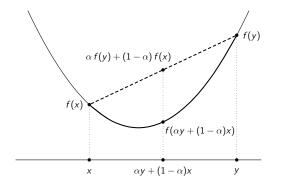
Exercises

- **1.2** Let $P = \{x : Ax \le b\}$ where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Prove that $rec(P) = \{d \in \mathbb{R}^n : Ad \le 0\}$.
- **1.3** If C_1 and C_2 are convex, then is $C_1 \cup C_2$ convex?
- **1.4** Prove that $B(\bar{x}, r) := \{z \in \mathbb{R}^n : ||z \bar{x}|| \le r\}$, is a convex set, whatever the norm $||\cdot||$ may be.
- **1.5** Write the vector (1,1) as a convex combination of the vectors (0,0),(3,0),(0,2),(3,2).

Definition (Convex function)

Let $C \subseteq \mathbb{R}^n$ be convex. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex on C if

$$f(\alpha y + (1 - \alpha)x) \le \alpha f(y) + (1 - \alpha)f(x)$$
 $\forall x, y \in C, \forall \alpha \in [0, 1]$



Remark

When $C = \mathbb{R}^n$ we will simply say that f is convex.

Theorem

A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex on \mathbb{R}^n if and only if the set

$$epi f_C := \{(x,y) \in C \times \mathbb{R} : y \ge f(x)\}$$

is convex.

Definition (Concave function)

Given a convex set $C \subseteq \mathbb{R}^n$, a function $f : \mathbb{R}^n \to \mathbb{R}$ is concave on C if -f is convex, i.e.,

$$f(\alpha y + (1 - \alpha)x) \ge \alpha f(y) + (1 - \alpha)f(x)$$
 $\forall x, y \in C, \forall \alpha \in [0, 1]$

Examples.

- A linear (affine) function $f(x) = c^{T}x + b$ is both convex and concave.
- Let $\|\cdot\|$ be any norm, then $f(x) = \|x\|$ is convex.

Theorem (continuity of convex functions)

Let $f: \mathbb{R}^n \to \mathbb{R}$ be convex on the convex set $C \subseteq \mathbb{R}^n$. Then f is continuous on ri(C).

Strictly convex and strongly convex functions

Definition (strictly convex function)

Given a convex set $C \subseteq \mathbb{R}^n$, a function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex on C if

$$f(\alpha y + (1 - \alpha)x) < \alpha f(y) + (1 - \alpha)f(x)$$
 $\forall x, y \in C, x \neq y, \forall \alpha \in (0, 1)$

Definition (strongly convex function)

Given a convex set $C \subseteq \mathbb{R}^n$, a function $f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex on C if there exists $\tau > 0$ s.t.

$$f(\alpha y + (1 - \alpha)x) \le \alpha f(y) + (1 - \alpha)f(x) - \frac{\tau}{2}\alpha(1 - \alpha)\|y - x\|_2^2$$
$$\forall x, y \in C, \forall \alpha \in [0, 1]$$

Remark

Similarly to convex functions, we say that f is strictly (strongly) concave on C if -f is strictly (strongly) convex on C.

Theorem

f is strongly convex if and only if $\exists \ \tau > 0$ such that $f(x) - \frac{\tau}{2} \|x\|_2^2$ is convex

Remark

By the previous theorem it follows that f is strongly convex if and only if there exists a convex function ψ and $\tau > 0$ such that $f(x) = \psi(x) + \frac{\tau}{2} \|x\|_2^2$.

Exercise 1.6

- Prove that: strongly convex ⇒ strictly convex ⇒ convex
- convex ⇒ strictly convex ?
- strictly convex ⇒ strongly convex ?

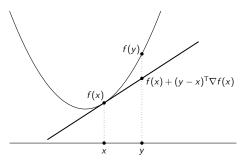
First order conditions

Assume that $C \subseteq \mathbb{R}^n$ is open and convex, $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable on C.

Theorem

f is convex on C if and only if

$$f(y) \ge f(x) + (y - x)^{\mathsf{T}} \nabla f(x) \qquad \forall \ x, y \in C.$$



First-order approximation of f is a global understimator

First order conditions

Theorem

• f is strictly convex on C if and only if

$$f(y) > f(x) + (y - x)^{\mathsf{T}} \nabla f(x)$$
 $\forall x, y \in C$, with $x \neq y$.

• f is strongly convex on C if and only if there exists $\tau > 0$ such that

$$f(y) \ge f(x) + (y - x)^{\mathsf{T}} \nabla f(x) + \frac{\tau}{2} ||y - x||_2^2 \qquad \forall \ x, y \in C.$$

Second order conditions

Assume that $C \subseteq \mathbb{R}^n$ is open and convex, $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable on C.

Theorem

• f is convex on C if and only if for all $x \in C$ the Hessian matrix $\nabla^2 f(x)$ is positive semidefinite, i.e.

$$v^{\mathsf{T}} \nabla^2 f(x) v \ge 0 \qquad \forall \ v \in \mathbb{R}^n, \ \forall x \in C,$$

or, equivalently, the eigenvalues of $\nabla^2 f(x)$ are ≥ 0 , $\forall x \in C$.

- If $\nabla^2 f(x)$ is positive definite for all $x \in C$, then f is strictly convex on C.
- f is strongly convex on C if and only if there exists $\tau > 0$ such that $\nabla^2 f(x) \tau I$ is positive semidefinite for all $x \in C$, i.e.

$$v^{\mathsf{T}} \nabla^2 f(x) v \ge \tau ||v||_2^2 \qquad \forall \ v \in \mathbb{R}^n, \quad \forall \ x \in C,$$

or, equivalently, the eigenvalues of $\nabla^2 f(x)$ are $\geq \tau$, $\forall x \in C$.

Convexity of quadratic functions

Let $f: \mathbb{R}^n \to \mathbb{R}$ be defined by

$$f(x) = \frac{1}{2}x^{\mathsf{T}}Qx + c^{\mathsf{T}}x$$

where Q is a $n \times n$ symmetric matrix, $c \in \mathbb{R}^n$. It is easy to check that

- $\nabla f = \frac{1}{2}(Qx + (x^{T}Q)^{T}) + c = Qx + c$
- Q is the Hessian of f.

Then f is:

- o convex iff Q is positive semidefinite
- strongly convex iff Q is positive definite
- concave iff Q is negative semidefinite
- strongly concave iff Q is negative definite

Examples

Let $f: \mathbb{R} \to \mathbb{R}$ and $C:=\mathbb{R}_+ \setminus \{0\}$.

- $f(x) = e^{px}$ for any $p \in \mathbb{R} \setminus \{0\}$ is strictly convex (on \mathbb{R}), but not strongly convex
- $f(x) = x^p$ is strictly convex on C if p > 1 or p < 0. Is it strongly convex?
- $f(x) = x^p$ is strictly concave on C if 0
- $f(x) = \log(x)$ is strictly concave, but not strongly concave on C

Let $f: \mathbb{R}^n \to \mathbb{R}$.

- f(x) = ||x|| is convex, but not strictly convex
- $f(x) = \max\{x_1, \dots, x_n\}$ is convex, but not strictly convex

Exercises

- **1.7** Prove that f(x) = ||x|| is convex, whatever the norm $||\cdot||$ may be.
- **1.8** Prove that if f is convex, then for any $x^1, \ldots, x^k \in C$ and $\alpha_1, \ldots, \alpha_k \in (0,1)$

s.t.
$$\sum\limits_{i=1}^k \alpha_i = 1$$
, one has $f\left(\sum\limits_{i=1}^k \alpha_i x^i\right) \leq \sum\limits_{i=1}^k \alpha_i f(x^i)$.

Hint. Follow the proof given in Lemma 1.

- **1.9** Prove that $f(x_1, x_2) = \frac{1}{x_1 x_2}$ is convex on the set $\{x \in \mathbb{R}^2 : x_1, x_2 > 0\}$.
- 1.10 Analyse the convexity properties of the function

$$f(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + x_3^2 + 3x_1x_2 + x_2x_3 - 6x_1 - 4x_2 - 3x_3$$

1.11 Let f_1 and f_2 be convex, then is the product f_1 f_2 convex?

Operations that preserve convexity

Theorem

- If f is convex and $\alpha > 0$, then αf is convex
- If f_1 and f_2 are convex, then $f_1 + f_2$ are convex
- If f is convex, then f(Ax + b) is convex

Examples

Log barrier for linear inequalities:

$$f(x) = -\sum_{i=1}^{m} log(b_i - a_i^{\mathsf{T}} x)$$
 $C = \{x \in \mathbb{R}^n : b_i - a_i^{\mathsf{T}} x > 0 \quad \forall i = 1, ..., m\}$

• Norm of affine function: f(x) = ||Ax + b||

Pointwise maximum

Theorem

- If f_1, \ldots, f_m are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex.
- If $\{f_i\}_{i\in I}$ is a family of convex functions, then $f(x) = \sup_{i\in I} f_i(x)$ is convex.

Example. If $\psi(x,\lambda): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is convex in x and concave in λ , then

$$\begin{split} p(x) &= \sup_{\lambda} \psi(x,\lambda) & \text{ is convex} \\ d(\lambda) &= \inf_{x} \psi(x,\lambda) & \text{ is concave} \end{split}$$

Composition

 $f: \mathbb{R}^n \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}$.

Theorem

- If f is convex and g is convex and nondecreasing, then $g \circ f$ is convex.
- If f is concave and g is convex and nonincreasing, then $g \circ f$ is convex.
- If f is concave and g is concave and nondecreasing, then $g \circ f$ is concave.
- If f is convex and g is concave and nonincreasing, then $g \circ f$ is concave.

Examples Let $f: \mathbb{R} \to \mathbb{R}$.

- If f is convex, then $e^{f(x)}$ is convex
- If f is concave and positive, then $\log f(x)$ is concave
- If f is convex, then $-\log(-f(x))$ is convex on $\{x: f(x) < 0\}$
- If f is concave and positive, then $\frac{1}{f(x)}$ is convex
- If f is convex and nonnegative, then $f(x)^p$ is convex for all $p \ge 1$

Sublevel sets

Given $f: \mathbb{R}^n \to \mathbb{R}$ and $k \in \mathbb{R}$, the set

$$S_k(f) = \{x \in \mathbb{R}^n : f(x) \le k\}$$

is said the k-sublevel set of f.

Exercise 1.12 Prove that if f is convex, then $S_k(f)$ is a convex set for any $k \in \mathbb{R}$.

Is the converse true?

Quasiconvex functions

Definition (Quasiconvex convex function)

Given a convex set $C \subseteq \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ is said quasiconvex on C if the sets

$$S_k(f) \cap C = \{x \in C : f(x) \le k\}$$

are convex for all $k \in \mathbb{R}$.

f is said quasiconcave on C if -f is quasiconvex on C.

Examples

- $f(x) = \sqrt{|x|}$ is quasiconvex on \mathbb{R}
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$
- $f(x) = \log x$ is quasiconvex and quasiconcave