Convexity

V. Leclère (ENPC)

February 20th and 27th, 2026

Why should I bother to learn this stuff?

- Convex vocabulary and results are needed throughout the course, especially to obtain optimality conditions and duality relations.
- Convex analysis tools like Fenchel transform appears in modern machine learning theory
- $\Longrightarrow$ fundamental for M2 in continuous optimization
- $\Longrightarrow$ useful for M2 in operation research, machine learning (and some part of probability or mechanics)

Contents

Contents

Affine sets

Let X be a normed vector space (usually $X = \mathbb{R}^n$), and $C \subset X$

- C is affine if it contains the entire line through any two distinct points of C, *i.e.*,

$$\forall x, y \in C, \quad \forall \theta \in \mathbb{R}, \qquad \theta x + (1-\theta) y \in C.$$

- The affine hull of C is the set of affine combination of elements of C,

$$\text{aff}(C) := \Big\{ \sum_{i=1}^{K} \theta_i x_i \;\Big|\; \forall x_i \in C,\ \forall \theta_i \in \mathbb{R},\ \sum_{i=1}^{K} \theta_i = 1,\ \forall i \in [K], \forall K \in \mathbb{N} \Big\}$$

- $\text{aff}(C)$ is the smallest affine space containing C.
- The affine dimension of C is the dimension of $\text{aff}(C)$ (*i.e.*,the dimension of the vector space $\text{aff}(C) - x_0$ for $x_0 \in C$).
- The relative interior of C is defined as

$$\text{ri}(C) := \Big\{ x \in C \;\Big|\; \exists r > 0, \quad B(x,r) \cap \text{aff}(C) \subset C \Big\}$$

Affine sets

♡

Let X be a normed vector space (usually $X = \mathbb{R}^n$), and $C \subset X$

- C is affine if it contains the entire line through any two distinct points of C, *i.e.*,

$$\forall x, y \in C, \quad \forall \theta \in \mathbb{R}, \qquad \theta x + (1-\theta) y \in C.$$

- The affine hull of C is the set of affine combination of elements of C,

$$\mathrm{aff}(C) := \Big\{ \sum_{i=1}^{K} \theta_i x_i \;\Big|\; \forall x_i \in C, \ \forall \theta_i \in \mathbb{R}, \ \sum_{i=1}^{K} \theta_i = 1, \ \forall i \in [K], \forall K \in \mathbb{N} \Big\}$$

- $\mathrm{aff}(C)$ is the smallest affine space containing C.
- The affine dimension of C is the dimension of $\mathrm{aff}(C)$ (*i.e.*,the dimension of the vector space $\mathrm{aff}(C) - x_0$ for $x_0 \in C$).
- The relative interior of C is defined as

$$\mathrm{ri}(C) := \Big\{ x \in C \;\Big|\; \exists r > 0, \quad B(x,r) \cap \mathrm{aff}(C) \subset C \Big\}$$

Convex sets

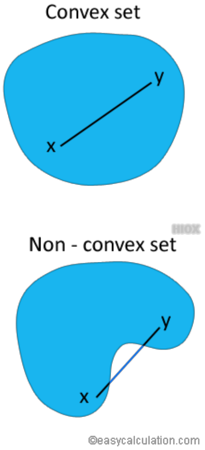
- C is convex if for any two points x and y in C the segment $[x, y] \subset C$, *i.e.*,

$$\forall x, y \in C,\ \forall \theta \in [0, 1],\ \theta x + (1 - \theta) y \in C.$$

- The convex hull of C is the set of convex combination of elements of C, *i.e.*,

$$\mathrm{conv}(C) := \Big\{ \sum_{i=1}^{K} \theta_i x_i \mid \ \forall x_i \in C, \\ \forall \theta_i \in [0, 1],\ \sum_{i=1}^{K} \theta_i = 1,\ \forall i \in [K],\ \forall K \in \mathbb{N} \Big\}$$

- $\mathrm{conv}(C)$ is the smallest convex set containing C.



Cones

- C is a cone if for all $x \in C$ the ray $\mathbb{R}_+ x \subset C$, *i.e.*,

$$\forall x \in C, \quad \forall \theta \in \mathbb{R}_+, \qquad \theta x \in C.$$

- The (convex) conic hull of C is the set of all conic combination of elements of C *i.e.*,

$$\mathrm{cone}(C) := \Big\{ \sum_{i=1}^{K} \theta_i x_i \mid \forall x_i \in C, \forall \theta_i \in \mathbb{R}_+, \forall i \in [K], \forall K \in \mathbb{N} \Big\}$$

- $\mathrm{cone}(C)$ is the smallest convex cone containing C.
- A cone C is pointed if it does not contain any full line $\mathbb{R}x$ for $x \neq 0$.
- For C convex, $\mathrm{cone}(C) = \bigcup_{t \geq 0} tC$

Examples

Let $X = \mathbb{R}^n$.

- Any affine space is convex.
- Any hyperplane of X can be defined as $H := \{x \in X \mid \ a^\top x = b\}$ for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ and is an affine space of dimension $n-1$.
- H divide X into two half-spaces $\{x \in \mathbb{R}^n \mid a^\top x \leq b\}$ and $\{x \in \mathbb{R}^n \mid a^\top x \geq b\}$ which are (closed) convex sets.
- For any norm $\|\cdot\|$ the ball $B_{\|\cdot\|}(x_0, r) := \{x \in X \mid \|x - x_0\| \leq r\}$ is a (closed) convex set. ♣ Exercise: Prove it.
- The set $C = \{(x,t) \in X \times \mathbb{R} \mid \|x\| \leq t\ \}$ is a cone.
- The set $C = \{x \in X \mid Ax \leq b\}$ where A and b are given is a (closed) convex set called polyhedron and polytope if it is bounded.

Examples

Let $X = \mathbb{R}^n$.

- Any affine space is convex.
- Any hyperplane of X can be defined as $H := \{x \in X \mid \; a^\top x = b\}$ for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ and is an affine space of dimension $n-1$.
- H divide X into two half-spaces $\{x \in \mathbb{R}^n \mid a^\top x \leq b\}$ and $\{x \in \mathbb{R}^n \mid a^\top x \geq b\}$ which are (closed) convex sets.
- For any norm $\|\cdot\|$ the ball $B_{\|\cdot\|}(x_0, r) := \{x \in X \mid \|x - x_0\| \leq r\}$ is a (closed) convex set.
  ♣ Exercise: Prove it.
- The set $C = \{(x,t) \in X \times \mathbb{R} \mid \|x\| \leq t \}$ is a cone.
- The set $C = \{x \in X \mid Ax \leq b\}$ where A and b are given is a (closed) convex set called polyhedron and polytope if it is bounded.

Operations preserving convexity ♡

Assume that all sets denoted by C (indexed or not) are convex.

- $C_1 + C_2$ and $C_1 \times C_2$ are convex sets.
- For any arbitrary index set $\mathcal{I}$ the intersection $\bigcap_{i \in \mathcal{I}} C_i$ is convex.
- Let f be an affine function. Then $f(C)$ and $f^{-1}(C)$ are convex.
- In particular, $C + x_0$, and tC are convex. The projection of C on any affine space is convex.
- The closure $\mathrm{cl}(C)$ and relative interior $\mathrm{ri}(C)$ are convex.

♣ Exercise: Prove these results.

Perspective and linear-fractional function

Let $P : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be the perspective function defined as $P(x,t) = x/t$, with $\mathrm{dom}(P) = \mathbb{R}^n \times \mathbb{R}^*_+$.

Theorem

If $C \subset \mathrm{dom}(P)$ is convex, then $P(C)$ is convex.
If $C \subset \mathbb{R}^n$ is convex, then $P^{-1}(C)$ is convex.

♠ Exercise: Prove this result.

Perspective and linear-fractional function

Let $P : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ be the perspective function defined as $P(x, t) = x/t$, with $\mathrm{dom}(P) = \mathbb{R}^n \times \mathbb{R}_+^*$.

Theorem

If $C \subset \mathrm{dom}(P)$ is convex, then $P(C)$ is convex.
If $C \subset \mathbb{R}^n$ is convex, then $P^{-1}(C)$ is convex.

♠ Exercise: Prove this result.

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear-fractional function of the form $f(x) := (Ax + b)/(c^\top x + d)$, with $\mathrm{dom}(f) = \{x \,|c^\top x + d > 0\}$.

Theorem

If $C \subset \mathrm{dom}(f)$ is convex, then $f(C)$ and $f^{-1}(C)$ are convex.

♣ Exercise: prove this result.

Cone ordering

Let $K \subset \mathbb{R}^n$ be a closed, convex, pointed cone with non-empty interior. We define the cone ordering according to K by

$$x \preceq_K y \quad \Longleftrightarrow \quad y - x \in K.$$

Examples:

- For $K = \mathbb{R}^n_+$, $\preceq_K$ is the componentwise ordering, *i.e.*,

$$x \preceq_{\mathbb{R}^n_+} y \quad \Longleftrightarrow \quad \forall i \in [n],\ x_i \leq y_i.$$

- For $K = S_n^+(\mathbb{R})$, $\preceq_K$ is the Löwner ordering, *i.e.*,

$$A \preceq_{S_n^+} B \quad \Longleftrightarrow \quad B - A \in S_n^+(\mathbb{R}).$$

♣ Exercise: Prove that $\preceq_K$ is a partial order (*i.e.*,reflexive, antisymmetric, transitive) compatible with scalar product, addition and limits.

Contents

Separation

Let X be a Banach space, and X^* its topological dual (i.e. the set of all continuous linear forms on X).

Theorem (Simple separation)

Let A and B be convex non-empty, disjoint subsets of X. There exists a separating hyperplane $(x^*, \alpha) \in X^* \times \mathbb{R}$ such that

$$\langle x^*, a\rangle \leq \alpha \leq \langle x^*, b\rangle \qquad \forall a, b \in A \times B.$$

Theorem (Strong separation)

Let A and B be convex non-empty, disjoint subsets of X. Assume that, A is closed, and B is compact (e.g. a point), then there exists a strict separating hyperplane $(x^*, \alpha) \in X^* \times \mathbb{R}$ such that, there exists $\varepsilon > 0$,

$$\langle x^*, a\rangle + \varepsilon \leq \alpha \leq \langle x^*, b\rangle - \varepsilon \qquad \forall a, b \in A \times B.$$

Remark: In Banach spaces these theorems require the Zorn Lemma which is equivalent to the axiom of choice. We will stay in finite dimension where these theorems are easier to prove.

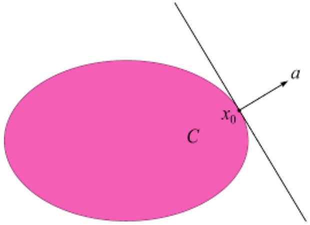
Supporting hyperplane

Theorem

Let $x_0 \notin \mathrm{ri}(C)$ and C convex. Then there exists $a \neq 0$ such that

$$a^\top x \geq a^\top x_0, \qquad \forall x \in C$$

If $x_0 \in C$, say that $H = \{x \mid a^\top x = a^\top x_0\}$ is a supporting hyperplane of C at x_0.



♣ Exercise: prove this theorem

Remark: there can be more than one supporting hyperplane at a given point.

Convex set as intersection of half-spaces

- The closed convex hull of $C \subset X$, denoted $\overline{\mathrm{conv}}(C)$ is the smallest closed convex set containing C.
- $\overline{\mathrm{conv}}(C)$ is the intersection of all the half-spaces containing C.
- A polyhedron is a finite intersection of half-spaces while a convex set is a possibly non-finite intersection of half-spaces.

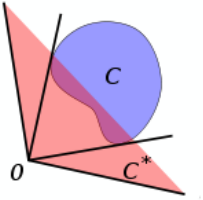
Dual and normal cones

- Let $C \subset \mathbb{R}^n$ be a set. We define its dual cone by

$$C^\oplus := \{x \mid x^\top c \geq 0, \quad \forall c \in C\}$$

- For any set C, $C^\oplus$ is a closed convex cone.
- The normal cone of C at x_0 is

$$N_C(x_0) := \{\lambda \in \mathbb{R}^n \mid \lambda^\top (x - x_0) \leq 0, \forall x \in C\}$$



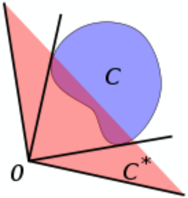
Dual and normal cones

- Let $C \subset \mathbb{R}^n$ be a set. We define its dual cone by

$$C^{\oplus} := \{x \mid x^\top c \geq 0, \quad \forall c \in C\}$$

- For any set C, $C^{\oplus}$ is a closed convex cone.
- The normal cone of C at x_0 is

$$N_C(x_0) := \{\lambda \in \mathbb{R}^n \mid \lambda^\top(x - x_0) \leq 0, \forall x \in C\}$$



Examples

- The positive orthant $K = \mathbb{R}^n_+$ is a self dual cone, that is $K^\oplus = K$.
- In the space of symmetric matrices $S_n(\mathbb{R})$, with the scalar product $\langle A, B\rangle = \mathrm{tr}(AB)$, the set of positive semidefinite matrices $K = S_n^+(\mathbb{R})$ is self dual.
- Let $\|\cdot\|$ be a norm. The cone $K = \{(x,t) \mid \|x\| \leq t\}$ has for dual $K^\oplus = \{(\lambda, z) \mid \|\lambda\|_\star \leq z\}$, where $\|\lambda\|_\star := \sup_{x:\|x\|\leq 1} \lambda^\top x$.

♠ Exercise: prove these results

Some basic properties

Let $K \subset \mathbb{R}^n$ be a cone.

- $K^\oplus$ is closed convex.
- $K_1 \subset K_2$ implies $K_2^\oplus \subset K_1^\oplus$
- $K^{\oplus\oplus} = \overline{\mathrm{conv}}\, K$

♣ Exercise: Prove these results

Video resources

https://www.youtube.com/watch?v=P3W_wFZ2kUo

Contents

Contents

Functions with non finite values

- It is very useful in optimization to allow functions to take non-finite values, that is to take values in $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.
- If both $-\infty$ and $+\infty$ are allowed be very careful of each addition[1]!
- Let $f : X \to \bar{\mathbb{R}}$. We define
  - The epigraph of f as
$$\mathrm{epi}(f) := \{(x,t) \in X \times \mathbb{R} \quad | \quad f(x) \leq t\}$$
  - the domain of f as
$$\mathrm{dom}(f) := \{x \in X \quad | \quad f(x) < +\infty\}.$$
  - The sublevel set of level α
$$lev_\alpha(f) := \{x \in X \quad | \quad f(x) \leq \alpha\}.$$
- f is said to be lower semicontinuous (l.s.c.) if $\mathrm{epi}(f)$ is closed.
- f is said to be proper if it never takes value $-\infty$, has a non-empty domain (at least one finite value).

[1] Which is why we usually consider proper functions.

Functions with non finite values

- It is very useful in optimization to allow functions to take non-finite values, that is to take values in $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.
- If both $-\infty$ and $+\infty$ are allowed be very careful of each addition[1]!
- Let $f : X \to \bar{\mathbb{R}}$. We define
  - The epigraph of f as
    $$\mathrm{epi}(f) := \{(x,t) \in X \times \mathbb{R} \quad | \quad f(x) \leq t \}$$
  - the domain of f as
    $$\mathrm{dom}(f) := \{x \in X \quad | \quad f(x) < +\infty\}.$$
  - The sublevel set of level α
    $$lev_\alpha(f) := \{x \in X \quad | \quad f(x) \leq \alpha\}.$$
- f is said to be lower semicontinuous (l.s.c.) if $\mathrm{epi}(f)$ is closed.
- f is said to be proper if it never takes value $-\infty$, has a non-empty domain (at least one finite value).

[1] Which is why we usually consider proper functions.

Functions with non finite values

- It is very useful in optimization to allow functions to take non-finite values, that is to take values in $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$.
- If both $-\infty$ and $+\infty$ are allowed be very careful of each addition[1]!
- Let $f : X \to \bar{\mathbb{R}}$. We define
  - The epigraph of f as
  $$\mathrm{epi}(f) := \{(x,t) \in X \times \mathbb{R} \quad | \quad f(x) \leq t\}$$
  - the domain of f as
  $$\mathrm{dom}(f) := \{x \in X \quad | \quad f(x) < +\infty\}.$$
  - The sublevel set of level α
  $$lev_\alpha(f) := \{x \in X \quad | \quad f(x) \leq \alpha\}.$$
- f is said to be lower semicontinuous (l.s.c.) if $\mathrm{epi}(f)$ is closed.
- f is said to be proper if it never takes value $-\infty$, has a non-empty domain (at least one finite value).

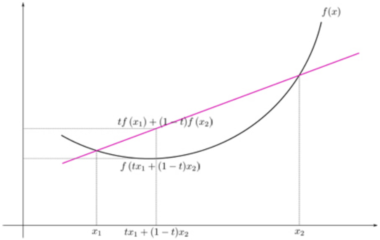
[1] Which is why we usually consider proper functions.

Convex function

- A function $f : X \to \bar{\mathbb{R}}$ is convex if its epigraph is convex.
- $f : X \to \mathbb{R} \cup \{+\infty\}$ is convex iff

$$\forall t \in [0,1],\ \forall x, y \in X,\quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

- f is concave if $-f$ is convex.



Basic properties

♡

- If f, g convex, $t > 0$, then $tf + g$ is convex.
- If f convex non-decreasing, g convex, then $f \circ g$ convex.
- If f convex and a affine, then $f \circ a$ is convex.
- If $(f_i)_{i \in I}$ is a family of convex functions, then $\sup_{i \in I} f_i$ is convex.
- The domain and the sublevel sets of a convex function are convex.
- A convex function is always above its tangents (supporting hyperplanes).

♣ Exercise: Prove these results.

Basic properties

♡

- If f, g convex, $t > 0$, then $tf + g$ is convex.
- If f convex non-decreasing, g convex, then $f \circ g$ convex.
- If f convex and a affine, then $f \circ a$ is convex.
- If $(f_i)_{i \in I}$ is a family of convex functions, then $\sup_{i \in I} f_i$ is convex.
- The domain and the sublevel sets of a convex function are convex.
- A convex function is always above its tangents (supporting hyperplanes).

♣ Exercise: Prove these results.

Theorem (Jensen inequality)

Let f be a convex function and $\boldsymbol{X}$ an integrable random variable. Then we have

$$f(\mathbb{E}[\boldsymbol{X}]) \leq \mathbb{E}[f(\boldsymbol{X})].$$

Convey function: regularity ♡

Consider a convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$.

- f is continuous (on $\mathbb{R}^n$) if and only if $\mathrm{dom}(f) = \mathbb{R}^n$ (i.e., if it is finite everywhere).
- f is continuous on the interior of its domain.
- f is Lipschitz on any compact convex set contained in the interior of its domain.

Convex functions: strict and strong convexity ♡

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is strictly convex iff

$$\forall t \in]0,1[, \quad \forall x, y \in X, \qquad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-strongly convex iff $\forall t \in]0,1[, \quad \forall x, y \in X,$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x - y\|^2$$

- If $f \in C^1(\mathbb{R}^n)$
  - $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ iff f convex
  - if strict inequality holds, then f strictly convex
  - $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-strongly convex iff $\forall x, y \in X$

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$$

- If $f \in C^2(\mathbb{R}^n)$,
  - $\nabla^2 f \succeq 0$ iff f convex
  - if $\nabla^2 f \succ 0$ then f strictly convex
  - if $\nabla^2 f \succeq \alpha I$ then f is α-strongly convex

Convex functions: strict and strong convexity

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is strictly convex iff

$$\forall t \in]0,1[, \quad \forall x, y \in X, \qquad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-strongly convex iff $\forall t \in]0,1[, \quad \forall x, y \in X,$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x-y\|^2$$

- If $f \in C^1(\mathbb{R}^n)$
  - $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ iff f convex
  - if strict inequality holds, then f strictly convex
  - $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-strongly convex iff $\forall x, y \in X$

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$$

- If $f \in C^2(\mathbb{R}^n)$,
  - $\nabla^2 f \succcurlyeq 0$ iff f convex
  - if $\nabla^2 f \succ 0$ then f strictly convex
  - if $\nabla^2 f \succcurlyeq \alpha I$ then f is α-strongly convex

Convex functions: strict and strong convexity

♡

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is strictly convex iff
$$\forall t \in]0,1[, \quad \forall x, y \in X, \qquad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$
- $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-strongly convex iff $\forall t \in]0,1[, \quad \forall x, y \in X,$
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x - y\|^2$$
- If $f \in C^1(\mathbb{R}^n)$
  - $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ iff f convex
  - if strict inequality holds, then f strictly convex
  - $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-strongly convex iff $\forall x, y \in X$
$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$$
- If $f \in C^2(\mathbb{R}^n)$,
  - $\nabla^2 f \succcurlyeq 0$ iff f convex
  - if $\nabla^2 f \succ 0$ then f strictly convex
  - if $\nabla^2 f \succcurlyeq \alpha I$ then f is α-strongly convex

Three convexity tests you will actually use

Test 1 — Definition (segment / Jensen)

To prove f convex on a convex domain $\mathrm{dom}(f)$, show

$$\forall x, y \in \mathrm{dom}(f),\ \forall t \in [0,1], \qquad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Test 2 — Epigraph / sublevel sets (geometric)

f is convex $\iff$ its epigraph is convex:

$$\mathrm{epi}(f) = \{(x,s) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq s\} \text{ is convex.}$$

Test 3 — Smooth test (first/second order)

If $f \in C^1$ on an open convex set:

$$f \text{ convex } \iff \langle \nabla f(x) - \nabla f(y), x - y\rangle \geq 0 \quad \forall x, y.$$

If $f \in C^2$:

Important examples

- The indicator function of a set $C \subset X$,
$$\mathbb{I}_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \textit{otherwise} \end{cases}$$
is convex iff C is convex.
- $x \mapsto e^{ax}$ is convex for any $a \in \mathbb{R}$
- $x \mapsto \|x\|^q$ is convex for $q \geq 1$ and any norm
- $x \mapsto \ln(x)$ is concave
- $x \mapsto x\ln(x)$ is convex
- $x \mapsto \ln(\sum_{i=1}^n e^{x_i})$ is convex

Contents

Convex optimization problem ♡

$$\underset{x\in C}{\text{Min}} \quad f(x)$$

Where C is non-empty, closed convex and f convex finite valued[2], is a convex optimization problem.

- If C is compact and f proper lsc, then there exists an optimal solution.
- If f is proper lsc and coercive, then there exists an optimal solution.
- The set of optimal solutions is convex.
- If f is strictly convex the minimum (if it exists) is unique.
- If f is α-strongly convex (and proper lsc) the minimum exists and is unique.

♣ Exercise: Prove these results.

[2]Alternatively, f can be proper lower semicontinuous. Here I assumed $C \subset \mathrm{dom}(f)$.

Optimality conditions

Note that minimizing f over C or minimizing $f + \mathbb{I}_C$ over X is the same thing.

We consider the (unconstrained) optimization problem

$$\underset{x \in X}{\mathrm{Min}} \quad f(x),$$

with $x^\sharp$ an optimal solution and f not necessarily convex.

- If f is differentiable, then $\nabla f(x^\sharp) = 0$.
- If f is twice differentiable, then $\nabla^2 f(x^\sharp) \succeq 0$.
- If f is twice differentiable and $\nabla f(x_0) = 0$ and $\nabla^2 f(x_0) \succ 0$ then x_0 is a local minimum.

If, in addition, f is convex then $\nabla f(x) = 0$ is a necessary and sufficient optimality condition.

Optimality conditions

♡

Note that minimizing f over C or minimizing $f + \mathbb{I}_C$ over X is the same thing.
We consider the (unconstrained) optimization problem

$$\operatorname*{Min}_{x \in X} \quad f(x),$$

with $x^\sharp$ an optimal solution and f not necessarily convex.

- If f is differentiable, then $\nabla f(x^\sharp) = 0$.
- If f is twice differentiable, then $\nabla^2 f(x^\sharp) \succeq 0$.
- If f is twice differentiable and $\nabla f(x_0) = 0$ and $\nabla^2 f(x_0) \succ 0$ then x_0 is a local minimum.

If, in addition, f is convex then $\nabla f(x) = 0$ is a necessary and sufficient optimality condition.

Optimality conditions

Note that minimizing f over C or minimizing $f + \mathbb{I}_C$ over X is the same thing.
We consider the (unconstrained) optimization problem

$$\underset{x \in X}{\mathrm{Min}} \qquad f(x),$$

with $x^\sharp$ an optimal solution and f not necessarily convex.

- If f is differentiable, then $\nabla f(x^\sharp) = 0$.
- If f is twice differentiable, then $\nabla^2 f(x^\sharp) \succeq 0$.
- If f is twice differentiable and $\nabla f(x_0) = 0$ and $\nabla^2 f(x_0) \succ 0$ then x_0 is a local minimum.

If, in addition, f is convex then $\nabla f(x) = 0$ is a necessary and sufficient optimality condition.

Optimality conditions

Note that minimizing f over C or minimizing $f + \mathbb{I}_C$ over X is the same thing.
We consider the (unconstrained) optimization problem

$$\underset{x \in X}{\mathrm{Min}} \qquad f(x),$$

with $x^\sharp$ an optimal solution and f not necessarily convex.

- If f is differentiable, then $\nabla f(x^\sharp) = 0$.
- If f is twice differentiable, then $\nabla^2 f(x^\sharp) \succeq 0$.
- If f is twice differentiable and $\nabla f(x_0) = 0$ and $\nabla^2 f(x_0) \succ 0$ then x_0 is a local minimum.

If, in addition, f is convex then $\nabla f(x) = 0$ is a necessary and sufficient optimality condition.

Contents

Partial infimum

♡

Let f be a convex function and C a convex set. The function

$$g : x \mapsto \inf_{y \in C} f(x, y)$$

is convex.

♠ Exercise: Prove this result.

♣ Exercise: Prove that the function distance to a convex set C defined by

$$d_C(x) := \inf_{c \in C} \|c - x\|$$

is convex.

Partial infimum ♡

Let f be a convex function and C a convex set. The function

$$g : x \mapsto \inf_{y \in C} f(x, y)$$

is convex.

♠ Exercise: Prove this result.

♣ Exercise: Prove that the function distance to a convex set C defined by

$$d_C(x) := \inf_{c \in C} \|c - x\|$$

is convex.

Perspective function

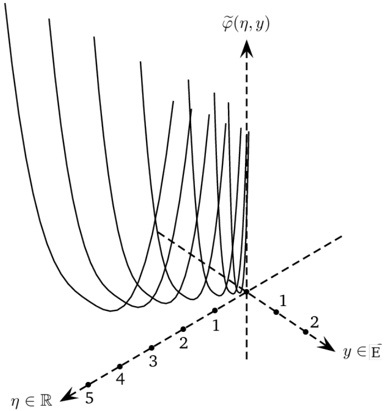
Let $\phi : E \to \overrightarrow{\mathbb{R}}$. The perspective of ϕ is defined as $\tilde{\phi} : \mathbb{R}_+^* \times E \to \mathbb{R}$ by

$$\tilde{\phi}(\eta, y) := \eta \phi(y/\eta).$$

Theorem

ϕ is convex iff $\tilde{\phi}$ is convex.

♠ Exercise: prove this result



Inf-Convolution

Let f and g be proper function from X to $\mathbb{R} \cup \{+\infty\}$. We define

$$f \Box g : x \mapsto \inf_{y \in X} f(y) + g(x-y)$$

♣ Exercise: Show that

- $f \Box g = g \Box f$
- If f and g are convex then so is $f \Box g$

Contents

Contents

Subdifferential of convex function

Let X be an Hilbert space, $f : X \to \bar{\mathbb{R}}$ convex.

- The subdifferential of f at $x \in \mathrm{dom}(f)$ is the set of slopes of all affine minorants of f exact at x:
$$\partial f(x) := \Big\{\lambda \in X \quad | \quad f(\cdot) \geq \langle \lambda, \cdot - x \rangle + f(x)\Big\}.$$
- If f is differentiable at x then
$$\partial f(x) = \big\{\nabla f(x)\big\}.$$

Examples

$\diamondsuit$

- If $f : x \mapsto |x|$, then

$$\partial f(x) = \begin{cases} -1 & \text{if } x < 0 \\ [-1, 1] & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$$

- If C is convex then, for $x \in C$, $\partial(\mathbb{I}_C)(x) = N_C(x)$
  ♣ Exercise: Prove it.
- If f_1 and f_2 are convex and differentiable. Define $f = \max(f_1, f_2)$. Then
  - if $f_1(x) > f_2(x)$, $\partial f(x) = \{\nabla f_1(x)\}$
  - if $f_1(x) < f_2(x)$, $\partial f(x) = \{\nabla f_2(x)\}$;
  - if $f_1(x) = f_2(x)$, $\partial f(x) = \overline{\mathrm{conv}}(\{\nabla f_1(x), \nabla f_2(x)\})$.

Subdifferential calculus

Let f_1 and f_2 be proper convex functions.

Theorem (Moreau-Rockafellar)

We have

$$\partial(f_1)(x) + \partial(f_2)(x) \subset \partial(f_1 + f_2)(x), \qquad \forall x$$

Further if $\mathrm{ri}(\mathrm{dom}(f_1)) \cap \mathrm{ri}(\mathrm{dom}(f_2)) \neq \emptyset$ *then*

$$\partial(f_1)(x) + \partial(f_2)(x) = \partial(f_1 + f_2)(x), \qquad \forall x$$

When f_i *is polyhedral you can replace* $\mathrm{ri}(\mathrm{dom}(f_i))$ *by* $\mathrm{dom}(f_i)$ *in the condition.*

Subdifferential calculus

Let f_1 and f_2 be proper convex functions.

Theorem (Moreau-Rockafellar)

We have

$$\partial(f_1)(x) + \partial(f_2)(x) \subset \partial(f_1 + f_2)(x), \qquad \forall x$$

Further if $\mathrm{ri}(\mathrm{dom}(f_1)) \cap \mathrm{ri}(\mathrm{dom}(f_2)) \neq \emptyset$ *then*

$$\partial(f_1)(x) + \partial(f_2)(x) = \partial(f_1 + f_2)(x), \qquad \forall x$$

When f_i *is polyhedral you can replace* $\mathrm{ri}(\mathrm{dom}(f_i))$ *by* $\mathrm{dom}(f_i)$ *in the condition.*

Theorem

If f *is convex and* $a : x \mapsto Ax + b$ *with* $Im(a) \cap \mathrm{ri}(\mathrm{dom}(f)) \neq \emptyset$*, then*

$$\partial(f \circ a)(x) = A^\top \partial f(Ax + b).$$

Counterexample to the qualification in Moreau–Rockafellar

$$f(x) = \begin{cases} -\sqrt{x}, & x \geq 0, \\ +\infty, & x < 0, \end{cases} \qquad g(x) = f(-x).$$

Then $\mathrm{dom}\, f = [0, \infty)$, $\mathrm{dom}\, g = (-\infty, 0]$, hence $\mathrm{dom}\, f \cap \mathrm{dom}\, g = \{0\}$ and

$$(f+g)(x) = \begin{cases} 0, & x = 0, \\ +\infty, & x \neq 0, \end{cases} \quad \Rightarrow \quad f + g = \delta_{\{0\}}.$$

Therefore $\partial(f+g)(0) = \mathbb{R}$. On the other hand, $\partial f(0) = \partial g(0) = \emptyset$, so

$$\partial f(0) + \partial g(0) = \emptyset \subsetneq \mathbb{R} = \partial(f+g)(0).$$

First order optimality conditions

Theorem

Let $f : X \mapsto \mathbb{R} \cup \{+\infty\}$ be a convex function (not necessarily) differentiable. $x^\sharp$ is a minimizer of f if and only if $0 \in \partial f(x^\sharp)$.

First order optimality conditions

Theorem

Let $f : X \mapsto \mathbb{R} \cup \{+\infty\}$ be a convex function (not necessarily) differentiable. $x^\sharp$ is a minimizer of f if and only if $0 \in \partial f(x^\sharp)$.

Theorem

Let f be a proper convex function and C a closed non-empty convex set such that $\mathrm{ri}(C) \cap \mathrm{ri}(\mathrm{dom}(f)) \neq \emptyset$ then $x^\sharp$ is an optimal solution to

$$\min_{x \in C} \quad f(x)$$

iff

$$0 \in \partial f(x^\sharp) + N_C(x^\sharp),$$

iff

$$\exists \lambda \in \partial f(x^\sharp), \quad \lambda \in -N_C(x^\sharp).$$

Normal cone, Tangent cone and optimality

Let C be a convex set. We define the tangent cone of $C \subset \mathbb{R}^n$ at point $x \in C$, as the set of directions in which you can move from x while staying in C for some time, that is

$$T_C(x) := \overline{\left\{\lambda(y - x) \;\middle|\; y \in C, \quad \lambda \in \mathbb{R}^+\right\}}$$

In particular, $T_C(x) = \mathbb{R}^n$ iff $x \in int(C)$.

♣ Exercise: Prove that $[T_C(x)]^\oplus = -N_C(x)$.

Normal cone, Tangent cone and optimality

Let C be a convex set. We define the tangent cone of $C \subset \mathbb{R}^n$ at point $x \in C$, as the set of directions in which you can move from x while staying in C for some time, that is

$$T_C(x) := \overline{\left\{ \lambda(y - x) \;\middle|\; y \in C, \quad \lambda \in \mathbb{R}^+ \right\}}$$

In particular, $T_C(x) = \mathbb{R}^n$ iff $x \in int(C)$.

♣ Exercise: Prove that $[T_C(x)]^{\oplus} = -N_C(x)$.

Partial infimum

Let $f : X \times Y \to \bar{\mathbb{R}}$ be a jointly convex and proper function, and define

$$v(x) = \inf_{y \in Y} f(x, y)$$

then v is convex.

If v is proper, and $v(x) = f(x, y^\sharp(x))$ then

$$\partial v(x) = \{g \in X \quad | \quad (g, 0) \in \partial f(x, y^\sharp(x))\}$$

proof:

$$\begin{aligned} g \in \partial v(x) \quad &\Leftrightarrow \quad \forall x', \qquad v(x') \geq v(x) + \langle g, x' - x \rangle \\ &\Leftrightarrow \quad \forall x', y' \quad f(x', y') \geq f(x, y^\sharp(x)) + \left\langle \begin{pmatrix} g \\ 0 \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y^\sharp(x) \end{pmatrix} \right\rangle \\ &\Leftrightarrow \quad \begin{pmatrix} g \\ 0 \end{pmatrix} \in \partial f(x, y^\sharp(x)) \end{aligned}$$

Partial infimum

Let $f : X \times Y \to \bar{\mathbb{R}}$ be a jointly convex and proper function, and define

$$v(x) = \inf_{y \in Y} f(x, y)$$

then v is convex.
If v is proper, and $v(x) = f(x, y^\sharp(x))$ then

$$\partial v(x) = \left\{ g \in X \quad | \quad (g, 0) \in \partial f(x, y^\sharp(x)) \right\}$$

proof:

$$\begin{aligned} g \in \partial v(x) \quad & \Leftrightarrow \quad \forall x', \qquad v(x') \geq v(x) + \langle g, x' - x \rangle \\ & \Leftrightarrow \quad \forall x', y' \quad f(x', y') \geq f(x, y^\sharp(x)) + \left\langle \begin{pmatrix} g \\ 0 \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y^\sharp(x) \end{pmatrix} \right\rangle \\ & \Leftrightarrow \quad \begin{pmatrix} g \\ 0 \end{pmatrix} \in \partial f(x, y^\sharp(x)) \end{aligned}$$

Partial infimum

Let $f : X \times Y \to \bar{\mathbb{R}}$ be a jointly convex and proper function, and define

$$v(x) = \inf_{y \in Y} f(x, y)$$

then v is convex.
If v is proper, and $v(x) = f(x, y^\sharp(x))$ then

$$\partial v(x) = \left\{ g \in X \quad \middle| \quad (g, 0) \in \partial f(x, y^\sharp(x)) \right\}$$

proof:

$$\begin{aligned}
g \in \partial v(x) \quad & \Leftrightarrow \quad \forall x', \qquad v(x') \geq v(x) + \langle g, x' - x \rangle \\
& \Leftrightarrow \quad \forall x', y' \quad f(x', y') \geq f(x, y^\sharp(x)) + \left\langle \begin{pmatrix} g \\ 0 \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} x \\ y^\sharp(x) \end{pmatrix} \right\rangle \\
& \Leftrightarrow \quad \begin{pmatrix} g \\ 0 \end{pmatrix} \in \partial f(x, y^\sharp(x))
\end{aligned}$$

Convex function: regularity

- Assume f convex, then f is continuous on the relative interior of its domain, and Lipschitz on any compact contained in the relative interior of its domain.
- A proper convex function is subdifferentiable on the relative interior of its domain.
- If f is convex, it is L-Lipschitz iff $\partial f(x) \subset B(0, L), \quad \forall x \in \mathrm{dom}(f)$

Contents

Fenchel transform

$\diamondsuit$

Let X be a Hilbert space, $f : X \to \bar{\mathbb{R}}$ be a proper function.

- The Fenchel transform of f, is $f^\star : X \to \bar{\mathbb{R}}$ with

$$f^\star(\lambda) := \sup_{x \in X} \langle \lambda, x \rangle - f(x).$$

- $f^\star$ is convex lsc as the supremum of affine functions.
- $f \leq g$ implies that $f^\star \geq g^\star$.
- If f is proper convex lsc, then $f^{\star\star} = f$, otherwise $f^{\star\star} \leq f$.

$\clubsuit$ Exercise: Prove the first two points

Fenchel transform

Let X be a Hilbert space, $f : X \to \bar{\mathbb{R}}$ be a proper function.

- The Fenchel transform of f, is $f^\star : X \to \bar{\mathbb{R}}$ with

$$f^\star(\lambda) := \sup_{x \in X} \langle \lambda, x \rangle - f(x).$$

- $f^\star$ is convex lsc as the supremum of affine functions.
- $f \leq g$ implies that $f^\star \geq g^\star$.
- If f is proper convex lsc, then $f^{\star\star} = f$, otherwise $f^{\star\star} \leq f$.

♣ Exercise: Prove the first two points

Fenchel transform

Let X be a Hilbert space, $f : X \to \bar{\mathbb{R}}$ be a proper function.

- The Fenchel transform of f, is $f^\star : X \to \bar{\mathbb{R}}$ with

$$f^\star(\lambda) := \sup_{x \in X} \langle \lambda, x \rangle - f(x).$$

- $f^\star$ is convex lsc as the supremum of affine functions.
- $f \leq g$ implies that $f^\star \geq g^\star$.
- If f is proper convex lsc, then $f^{\star\star} = f$, otherwise $f^{\star\star} \leq f$.

♣ Exercise: Prove the first two points

Fenchel transform

Let X be a Hilbert space, $f : X \to \bar{\mathbb{R}}$ be a proper function.

- The Fenchel transform of f, is $f^\star : X \to \bar{\mathbb{R}}$ with

$$f^\star(\lambda) := \sup_{x \in X} \langle \lambda, x \rangle - f(x).$$

- $f^\star$ is convex lsc as the supremum of affine functions.
- $f \leq g$ implies that $f^\star \geq g^\star$.
- If f is proper convex lsc, then $f^{\star\star} = f$, otherwise $f^{\star\star} \leq f$.

♣ Exercise: Prove the first two points

Fenchel transform and subdifferential

- By definition $f^\star(\lambda) \geq \langle \lambda, x \rangle - f(x)$ for all x,
- thus we always have (Fenchel-Young) $f(x) + f^\star(\lambda) \geq \langle \lambda, x \rangle$.
- Recall that $\lambda \in \partial f(x)$ iff,

$$f(x') \geq f(x) + \langle \lambda, x' - x \rangle, \qquad \forall x'$$

iff

$$\langle \lambda, x \rangle - f(x) \geq \langle \lambda, x' \rangle - f(x') \qquad \forall x'$$

that is

$$\lambda \in \partial f(x) \Leftrightarrow x \in \arg\max_{x' \in X} \left\{ \langle \lambda, x' \rangle - f(x') \right\} \Leftrightarrow f(x) + f^\star(\lambda) = \langle \lambda, x \rangle$$

- If f proper convex lsc

$$\lambda \in \partial f(x) \quad \Longleftrightarrow \quad x \in \partial f^\star(\lambda).$$

Fenchel transform and subdifferential

- By definition $f^\star(\lambda) \geq \langle \lambda, x\rangle - f(x)$ for all x,
- thus we always have (Fenchel-Young) $f(x) + f^\star(\lambda) \geq \langle \lambda, x\rangle$.
- Recall that $\lambda \in \partial f(x)$ iff,

$$f(x') \geq f(x) + \langle \lambda, x' - x\rangle, \qquad \forall x'$$

iff

$$\langle \lambda, x\rangle - f(x) \geq \langle \lambda, x'\rangle - f(x') \qquad \forall x'$$

that is

$$\lambda \in \partial f(x) \Leftrightarrow x \in \arg\max_{x' \in X} \left\{\langle \lambda, x'\rangle - f(x')\right\} \Leftrightarrow f(x) + f^\star(\lambda) = \langle \lambda, x\rangle$$

- If f proper convex lsc

$$\lambda \in \partial f(x) \quad \Longleftrightarrow \quad x \in \partial f^\star(\lambda).$$

Fenchel transform and subdifferential

- By definition $f^\star(\lambda) \geq \langle \lambda, x \rangle - f(x)$ for all x,
- thus we always have (Fenchel-Young) $f(x) + f^\star(\lambda) \geq \langle \lambda, x \rangle$.
- Recall that $\lambda \in \partial f(x)$ iff,

$$f(x') \geq f(x) + \langle \lambda, x' - x \rangle, \qquad \forall x'$$

iff

$$\langle \lambda, x \rangle - f(x) \geq \langle \lambda, x' \rangle - f(x') \qquad \forall x'$$

that is

$$\lambda \in \partial f(x) \Leftrightarrow x \in \arg\max_{x' \in X} \left\{ \langle \lambda, x' \rangle - f(x') \right\} \Leftrightarrow f(x) + f^\star(\lambda) = \langle \lambda, x \rangle$$

- If f proper convex lsc

$$\lambda \in \partial f(x) \quad \Longleftrightarrow \quad x \in \partial f^\star(\lambda).$$

Fenchel transform and subdifferential $\diamondsuit$

- By definition $f^\star(\lambda) \geq \langle \lambda, x\rangle - f(x)$ for all x,
- thus we always have (Fenchel-Young) $f(x) + f^\star(\lambda) \geq \langle \lambda, x\rangle$.
- Recall that $\lambda \in \partial f(x)$ iff,

$$f(x') \geq f(x) + \langle \lambda, x' - x\rangle, \qquad \forall x'$$

iff

$$\langle \lambda, x\rangle - f(x) \geq \langle \lambda, x'\rangle - f(x') \qquad \forall x'$$

that is

$$\lambda \in \partial f(x) \Leftrightarrow x \in \arg\max_{x' \in X} \left\{ \langle \lambda, x'\rangle - f(x') \right\} \Leftrightarrow f(x) + f^\star(\lambda) = \langle \lambda, x\rangle$$

- If f proper convex lsc

$$\lambda \in \partial f(x) \quad \Longleftrightarrow \quad x \in \partial f^\star(\lambda).$$

Contents

What you have to know

- What is a affine set, a convex set, a polyhedron, a (convex) cone
- What is a convex function, that it is above its tangents.
- Jensen inequality
- What is a convex optimization problem. That any local minimum is a global minimum.
- The necessary optimality condition $\nabla f(x^\sharp) \in [T_X(x^\sharp)]^\oplus$

What you really should know

- That you can separate convex sets with a linear function
- What is the positive dual of a cone
- Basic manipulations preserving convexity (sum, cartesian product, intersection, linear projection)
- What is the domain, the sublevel of a function f
- What is a lower semicontinuous function, a proper convex function
- Conditions of (strict, strong) convexity for differentiable functions
- The partial minimum of a convex function is convex
- The definition of the subdifferential.
- The definition of the Fenchel transform.
- The link between Fenchel transform and subdifferential.

What you have to be able to do

- Show that a set is convex
- Show that a function is (strictly, strongly) convex
- Go from constrained problem to unconstrained problem using the indicator function $\mathbb{I}_X$

What you should be able to do

- Compute dual cones
- Use advanced results (projection, partial infimum, perspective) to show that a function or a set is convex
- Compute the Fenchel transform of simple functions