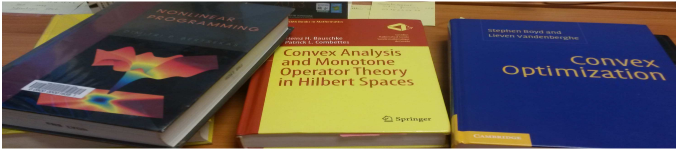
Master MVA
Optimization Reminders
Part I

Jean-Christophe Pesquet

Center for Visual Computing
Inria, CentraleSupélec
Institut Universitaire de France
jean-christophe@pesquet.eu

Reference books



- **D. Bertsekas**, Nonlinear programming, Athena Scientic, Belmont, Massachussets, 1996.
- **Y. Nesterov**, Introductory Lectures on Convex Optimization : A Basic Course, Springer, 2004.
- **S. Boyd and L. Vandenberghe**, Convex optimization, Cambridge University Press, 2004.
- **H. H. Bauschke and P. L. Combettes**, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, Springer, New York, 2017.

A few mathematical background

Hilbert spaces

A (real) pre-Hilbertian space $\mathcal{H}$ is a real vector space endowed with a function $\langle \cdot \mid \cdot \rangle : \mathcal{H}^2 \to \mathbb{R}$, called inner product such that

① $(\forall (x,y) \in \mathcal{H}^2)\ \langle x \mid y \rangle = \langle y \mid x \rangle$

② $(\forall (x,y,z) \in \mathcal{H}^3)\ \langle x + y \mid z \rangle = \langle x \mid z \rangle + \langle y \mid z \rangle$

③ $(\forall (x,y) \in \mathcal{H}^2)\ (\forall \alpha \in \mathbb{R})\ \langle \alpha x \mid y \rangle = \alpha \langle y \mid x \rangle$

④ $(\forall x \in \mathcal{H})\ \langle x \mid x \rangle \geq 0$ and
$\langle x \mid x \rangle = 0 \Leftrightarrow x = 0.$

- The associated norm is

$$(\forall x \in \mathcal{H}) \qquad \|x\| = \sqrt{\langle x \mid x \rangle}.$$

Hilbert spaces

A (real) **Hilbert space $\mathcal{H}$** is a complete pre-Hilbertian space.

- Particular case : $\mathcal{H} = \mathbb{R}^N$ (Euclidean space with dimension N).

Hilbert spaces

A (real) **Hilbert space $\mathcal{H}$** is a complete pre-Hilbertian space.

- Particular case : $\mathcal{H} = \mathbb{R}^N$ (Euclidean space with dimension N).

$2^{\mathcal{H}}$ is the power set of $\mathcal{H}$, i.e. the family of all subsets of $\mathcal{H}$.

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
A linear operator $L\colon \mathcal{H} \to \mathcal{G}$ is **bounded** if

$$\|L\| = \sup_{\|x\|_{\mathcal{H}} \leq 1} \|Lx\|_{\mathcal{G}} < +\infty$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
A linear operator $L\colon \mathcal{H} \to \mathcal{G}$ is **bounded** if

$$\|L\| = \sup_{\|x\| \leq 1} \|Lx\| < +\infty$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
A linear operator $L\colon \mathcal{H} \to \mathcal{G}$ is bounded if

$$\|L\| = \sup_{\|x\|=1} \|Lx\| < +\infty$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
A linear operator $L\colon \mathcal{H} \to \mathcal{G}$ is **bounded** if

$$\|L\| = \sup_{x \in \mathcal{H} \setminus \{0\}} \frac{\|Lx\|}{\|x\|} < +\infty$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
A linear operator $L\colon \mathcal{H} \to \mathcal{G}$ is bounded if

$$\|L\| = \sup_{x \in \mathcal{H} \setminus \{0\}} \frac{\|Lx\|}{\|x\|} < +\infty$$

- A linear operator from $\mathcal{H}$ to $\mathcal{G}$ is continuous if and only if it is bounded.
- In finite dimension, every linear operator is bounded.
- In the following, it will be assumed that all the underlying Hilbert spaces are finite dimensional.

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
A linear operator $L\colon \mathcal{H} \to \mathcal{G}$ is **bounded** if

$$\|L\| = \sup_{x \in \mathcal{H}\setminus\{0\}} \frac{\|Lx\|}{\|x\|} < +\infty$$

- A linear operator from $\mathcal{H}$ to $\mathcal{G}$ is continuous if and only if it is bounded.
- In finite dimension, every linear operator is bounded.
- In the following, it will be assumed that all the underlying Hilbert spaces are finite dimensional.

$\mathcal{B}(\mathcal{H},\mathcal{G})$: Banach space of (bounded) linear operators from $\mathcal{H}$ to $\mathcal{G}$.

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Its **adjoint L^*** is the operator in $\mathcal{B}(\mathcal{G}, \mathcal{H})$ defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \qquad \langle y \mid Lx \rangle_{\mathcal{G}} = \langle L^* y \mid x \rangle_{\mathcal{H}} .$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Its adjoint L^* is the operator in $\mathcal{B}(\mathcal{G}, \mathcal{H})$ defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \qquad \langle y \mid Lx \rangle = \langle L^* y \mid x \rangle .$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Its **adjoint L^*** is the operator in $\mathcal{B}(\mathcal{G}, \mathcal{H})$ defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \qquad \langle Lx \mid y \rangle = \langle x \mid L^* y \rangle \,.$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Its **adjoint L^*** is the operator in $\mathcal{B}(\mathcal{G}, \mathcal{H})$ defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \qquad \langle Lx \mid y \rangle = \langle x \mid L^* y \rangle .$$

Example :

$$\text{If} \qquad L \colon \mathcal{H} \to \mathcal{H}^n \colon x \mapsto (x, \ldots, x)$$

$$\text{then} \qquad L^* \colon \mathcal{H}^n \to \mathcal{H} \colon y = (y_1, \ldots, y_n) \mapsto \sum_{i=1}^{n} y_i$$

Proof :

$$\langle Lx \mid y \rangle = \langle (x, \ldots, x) \mid (y_1, \ldots, y_n) \rangle = \sum_{i=1}^{n} \langle x \mid y_i \rangle = \left\langle x \mid \sum_{i=1}^{n} y_i \right\rangle$$

Hilbert spaces

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Its adjoint L^* is the operator in $\mathcal{B}(\mathcal{G}, \mathcal{H})$ defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \qquad \langle Lx \mid y \rangle = \langle x \mid L^* y \rangle \,.$$

- We have $\|L^*\| = \|L\|$.
- If L is bijective (i.e. an isomorphism) then $L^{-1} \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ and $(L^{-1})^* = (L^*)^{-1}$.
- If $\mathcal{H} = \mathbb{R}^N$ and $\mathcal{G} = \mathbb{R}^M$ then $L^* = L^\top$.

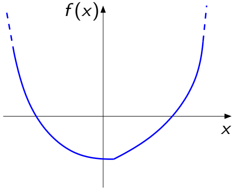
Infinite values functions

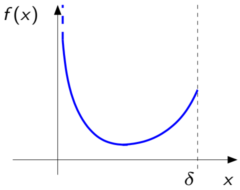
Functional analysis : definitions

Let S be a nonempty set of a Hilbert space $\mathcal{H}$.
Let $f\colon S \to]-\infty, +\infty]$.

- The **domain** of f is $\operatorname{dom} f = \{x \in S \mid f(x) < +\infty\}$.
- The function f is **proper** if $\operatorname{dom} f \neq \varnothing$.

Domains of the functions ?



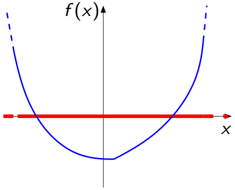


Functional analysis : definitions

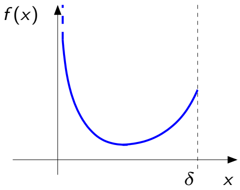
Let S be a nonempty set of a Hilbert space $\mathcal{H}$.
Let $f : S \to]-\infty, +\infty]$.

- The **domain** of f is $\operatorname{dom} f = \{x \in S \mid f(x) < +\infty\}$.
- The function f is **proper** if $\operatorname{dom} f \neq \varnothing$.

Domains of the functions ?



$\operatorname{dom} f = \mathbb{R}$
(proper)

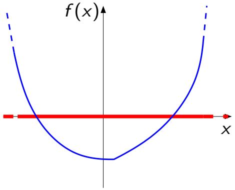


Functional analysis : definitions

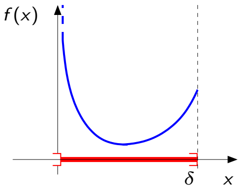
Let S be a nonempty set of a Hilbert space $\mathcal{H}$.
Let $f : S \to]-\infty, +\infty]$.

- The domain of f is $\operatorname{dom} f = \{x \in S \mid f(x) < +\infty\}$.
- The function f is proper if $\operatorname{dom} f \neq \varnothing$.

Domains of the functions ?



$\operatorname{dom} f = \mathbb{R}$
(proper)



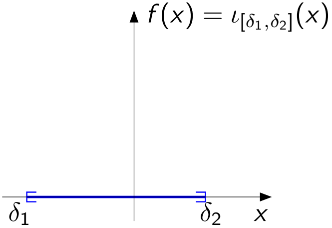
$\operatorname{dom} f =]0, \delta]$
(proper)

Functional analysis : definitions

Let $C \subset \mathcal{H}$.
The **indicator function of C** is

$$(\forall x \in \mathcal{H}) \qquad \iota_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise.} \end{cases}$$

Example : $C = [\delta_1, \delta_2]$



Limits inf and sup

Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of elements in $[-\infty, +\infty]$.
Its infimum limit is $\liminf \xi_n = \lim_{n\to+\infty} \inf \{\xi_k \mid k \geq n\} \in [-\infty, +\infty]$ and its supremum limit is $\limsup \xi_n = \lim_{n\to+\infty} \sup \{\xi_k \mid k \geq n\} \in [-\infty, +\infty]$.

Limits inf and sup

Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of elements in $[-\infty,+\infty]$.
Its infimum limit is $\liminf \xi_n = \lim_{n\to+\infty} \inf\left\{\xi_k \;\middle|\; k \geq n\right\} \in [-\infty,+\infty]$ and its supremum limit is $\limsup \xi_n = \lim_{n\to+\infty} \sup\left\{\xi_k \;\middle|\; k \geq n\right\} \in [-\infty,+\infty]$.

- $\limsup \xi_n = -\liminf(-\xi_n)$
- $\liminf \xi_n \leq \limsup \xi_n$
- $\lim_{n\to+\infty} \xi_n = \overline{\xi} \in [-\infty,+\infty]$ if and only if $\liminf \xi_n = \limsup \xi_n = \overline{\xi}$.

Epigraph

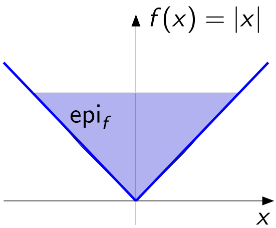
Let $f : \mathcal{H} \to \,]-\infty, +\infty]$. The **epigraph** of f is

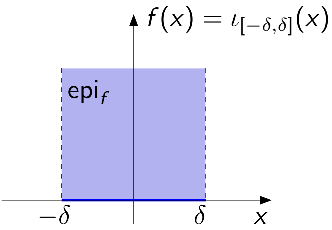
$$\operatorname{epi} f = \left\{ (x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \;\middle|\; f(x) \leq \zeta \right\}$$

Epigraph

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$. The **epigraph** of f is

$$\operatorname{epi} f = \left\{(x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \;\middle|\; f(x) \leq \zeta\right\}$$





Lower semi-continuity

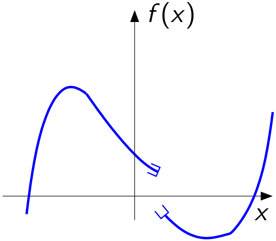
Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a **lower semi-continuous** (l.s.c.) function at $x \in \mathcal{H}$ if, for every sequence $(x_n)_{n\in\mathbb{N}}$ of $\mathcal{H}$,

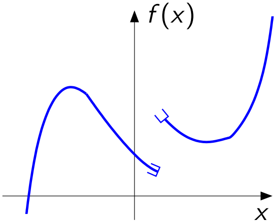
$$x_n \to x \quad \Rightarrow \quad \liminf f(x_n) \geq f(x).$$

Lower semi-continuity

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous function on $\mathcal{H}$ if and only if epi f is closed

- l.s.c. functions ?

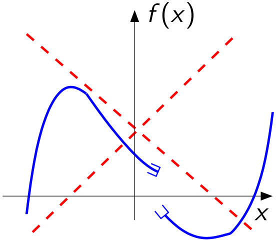


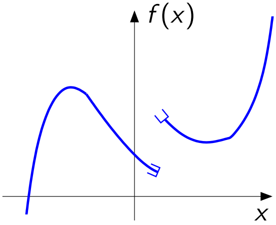


Lower semi-continuity

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous function on $\mathcal{H}$ if and only if $\operatorname{epi} f$ is closed

- l.s.c. functions ?

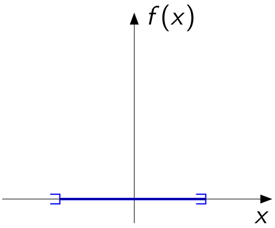


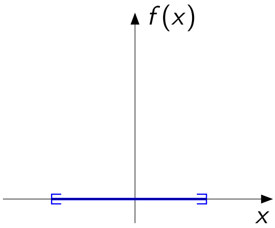


Lower semi-continuity

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous function on $\mathcal{H}$ if and only if epi f is closed

- l.s.c. functions ?

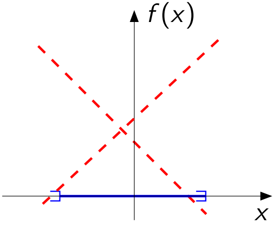


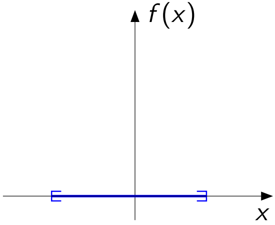


Lower semi-continuity

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous function on $\mathcal{H}$ if and only if epi f is closed

- l.s.c. functions ?

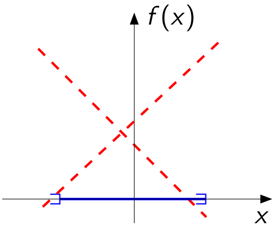


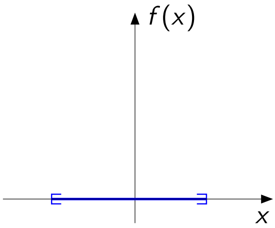


Lower semi-continuity

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$.
f is a lower semi-continuous function on $\mathcal{H}$ if and only if $\operatorname{epi} f$ is closed

- l.s.c. functions ?





Lower semi-continuity

- Every continuous function on $\mathcal{H}$ is l.s.c.
- Every finite sum of l.s.c. functions is l.s.c.
- Let $(f_i)_{i \in I}$ be a family of l.s.c functions.
  Then, $\sup_{i \in I} f_i$ is l.s.c.

Does a minimum exist ?

Minimizers

Let S be a nonempty set of a Hilbert space $\mathcal{H}$.
Let $f\colon S \to \,]-\infty,+\infty]$ be a proper function and let $\widehat{x} \in S$.

- $\widehat{x}$ is a local minimizer of f if $\widehat{x} \in \operatorname{dom} f$ and there exists an open neigborhood O of $\widehat{x}$ such that

$$(\forall x \in O \cap S) \qquad f(\widehat{x}) \leq f(x).$$

- $\widehat{x}$ is a (global) minimizer of f if

$$(\forall x \in S) \qquad f(\widehat{x}) \leq f(x).$$

Minimizers

Let S be a nonempty set of a Hilbert space $\mathcal{H}$.
Let $f\colon S \to \,]-\infty, +\infty]$ be a proper function and let $\widehat{x} \in S$.

- $\widehat{x}$ is a strict local minimizer of f if there exists an open neigborhood O of $\widehat{x}$ such that

$$(\forall x \in (O \cap S) \setminus \{\widehat{x}\}) \qquad f(\widehat{x}) < f(x).$$

- $\widehat{x}$ is a strict (global and unique) minimizer of f if

$$(\forall x \in S \setminus \{\widehat{x}\}) \qquad f(\widehat{x}) < f(x).$$

Existence of a minimizer

Weierstrass theorem

Let S be a nonempty compact set of a Hilbert space $\mathcal{H}$.
Let $f : S \to \,]-\infty, +\infty]$ be a proper l.s.c function.
Then, there exists $\widehat{x} \in S$ such that

$$f(\widehat{x}) = \inf_{x \in S} f(x).$$

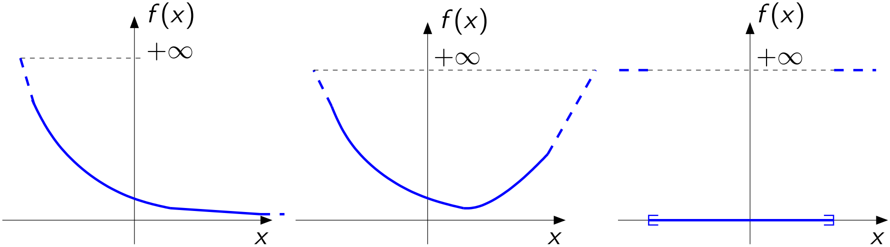
Existence of a minimizer

Let $\mathcal{H}$ be a Hilbert space. Let $f \colon \mathcal{H} \to \,]-\infty, +\infty]$.
f is coercive if $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

Existence of a minimizer

Let $\mathcal{H}$ be a Hilbert space. Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is **coercive** if $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

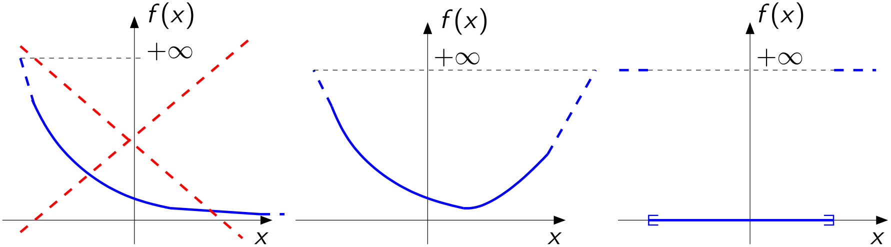
Coercive functions ?



Existence of a minimizer

Let $\mathcal{H}$ be a Hilbert space. Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is **coercive** if $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

Coercive functions ?



Existence of a minimizer

Theorem

Let $\mathcal{H}$ be a (finite dimensional) Hilbert space.
Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper l.s.c. coercive function.
Then, the set of minimizers of f is a nonempty compact set.

Existence of a minimizer

Theorem

Let $\mathcal{H}$ be a (finite dimensional) Hilbert space.
Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper l.s.c. coercive function.
Then, the set of minimizers of f is a nonempty compact set.

<u>Proof :</u> Since f is proper, there exists $x_0 \in \mathcal{H}$ such that $f(x_0) \in \mathbb{R}$. The coercivity of f implies that there exists $\eta \in \,]0, +\infty[$ such that, for every $x \in \mathcal{H}$ satisfying $\|x - x_0\| > \eta$, $f(x) > f(x_0)$.
Let $S = \big\{x \in \mathcal{H} \;\big|\; \|x - x_0\| \leq \eta\big\}$, $S \cap \operatorname{dom} f \neq \varnothing$ and S is compact.
Then, there exists $\widehat{x} \in S$ such that $f(\widehat{x}) = \inf_{x \in S} f(x) \leq f(x_0)$. Thus, $f(\widehat{x}) = \inf_{x \in \mathcal{H}} f(x)$.
$\mathrm{Argmin} f \subset S$ is bounded. In addition, if $(x_n)_{n \in \mathbb{N}}$ is a sequence of minimizers converging to $\widehat{x} \in \mathcal{H}$. Then, $f(\widehat{x}) \leq \liminf f(x_n) = \inf_{x \in \mathcal{H}} f(x)$ and, consequently, $\widehat{x} \in \mathrm{Argmin} f$. Therefore, $\mathrm{Argmin} f$ is closed.

Exercise 1

Let f be the Shannon entropy function defined as

$$f(x) = \begin{cases} \sum_{i=1}^{N} x^{(i)} \ln x^{(i)} & \text{if } x = (x^{(i)})_{1 \le i \le N} \in \,]0, +\infty[^N \\ +\infty & \text{if } (\exists j \in \{1, \ldots, N\}) \; x^{(j)} < 0. \end{cases}$$

1. How can we extend the definition of function f so that it is l.s.c. on $\mathbb{R}^N$?
2. What can be said about the existence of a minimizer of this function on a nonempty closed subset of the set $C = \left\{(x^{(i)})_{1 \le i \le N} \in [0, +\infty[^N \;\middle|\; \sum_{i=1}^{N} x^{(i)} = 1\right\}$?

Hints from differential calculus

Necessary condition for the existence of a minimizer (Euler's inequality)

Theorem

Let D be an open subset of a Hilbert space $\mathcal{H}$ and let $C \subset D$. Let $f : D \to]-\infty, +\infty]$ be differentiable at $\widehat{x} \in C$. If $\widehat{x}$ is a local minimizer of f on C then, for every $y \in \mathcal{H}$ such that $[\widehat{x}, y] \subset C$,

$$\langle \nabla f(\widehat{x}) \mid y - \widehat{x} \rangle \geq 0.$$

If $\widehat{x} \in \operatorname{int}(C)$, then the condition reduces to

$$\nabla f(\widehat{x}) = 0.$$

Necessary condition for the existence of a minimizer (Euler's inequality)

Theorem

Let D be an open subset of a Hilbert space $\mathcal{H}$ and let $C \subset D$. Let $f\colon D \to \,]-\infty, +\infty]$ be differentiable at $\widehat{x} \in C$. If $\widehat{x}$ is a local minimizer of f on C then, for every $y \in \mathcal{H}$ such that $[\widehat{x}, y] \subset C$,

$$\langle \nabla f(\widehat{x}) \mid y - \widehat{x} \rangle \geq 0.$$

If $\widehat{x} \in \mathrm{int}\,(C)$, then the condition reduces to

$$\nabla f(\widehat{x}) = 0.$$

<u>Remark</u> : A zero of the gradient ∇f is called a critical point of f.

Necessary condition for the existence of a minimizer

Theorem

Let C be an open subset of a Hilbert space $\mathcal{H}$. Let $f\colon C \to \mathbb{R}$ be differentiable on C. Let f be twice differentiable at $\widehat{x} \in C$. If If $\widehat{x}$ is a local minimizer of f on C, then

(i) $\nabla f(\widehat{x}) = 0$;

(ii) the Hessian $\nabla^2 f(\widehat{x})$ of f at $\widehat{x}$ is positive semi-definite, i.e.

$$(\forall z \in \mathcal{H}) \qquad \langle z \mid \nabla^2 f(\widehat{x}) z \rangle \geq 0.$$

Sufficient conditions for the existence of a minimizer

Theorem

Let C be an open subset of a Hilbert space $\mathcal{H}$. Let $f\colon C \to \mathbb{R}$ be differentiable on C.

(i) If f is twice differentiable at $\widehat{x} \in C$, $\nabla f(\widehat{x}) = 0$ and the Hessian $\nabla^2 f(\widehat{x})$ of f at $\widehat{x}$ is positive definite, i.e.

$$(\forall z \in \mathcal{H} \setminus \{0\}) \qquad \left\langle z \mid \nabla^2 f(\widehat{x}) z \right\rangle > 0.$$

then f has a strict local minimum at $\widehat{x}$.

Sufficient conditions for the existence of a minimizer

Theorem

Let C be an open subset of a Hilbert space $\mathcal{H}$. Let $f\colon C \to \mathbb{R}$ be differentiable on C.

(i) If f is twice differentiable at $\widehat{x} \in C$, $\nabla f(\widehat{x}) = 0$ and the Hessian $\nabla^2 f(\widehat{x})$ of f at $\widehat{x}$ is positive definite, i.e.

$$(\forall z \in \mathcal{H} \setminus \{0\}) \qquad \left\langle z \mid \nabla^2 f(\widehat{x}) z \right\rangle > 0.$$

then f has a strict local minimum at $\widehat{x}$.

(ii) If f is twice differentiable on an open neighborhood $D \subset C$ of $\widehat{x}$, $\nabla f(\widehat{x}) = 0$ and the Hessian of f is positive semi-definite on D, i.e.

$$(\forall x \in D)(\forall z \in \mathcal{H}) \qquad \left\langle z \mid \nabla^2 f(x) z \right\rangle \geq 0,$$

then f has a local minimum at $\widehat{x}$.

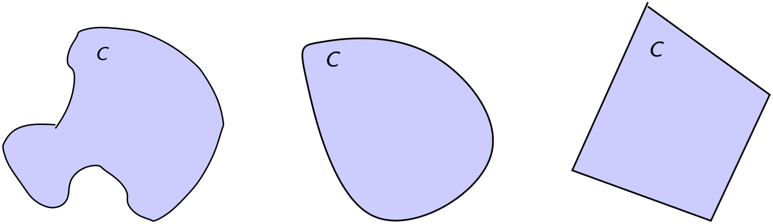
Magic of convexity

Convex set : definition

Let $\mathcal{H}$ be a Hilbert space. $C \subset \mathcal{H}$ is a **convex set** if

$$(\forall (x, y) \in C^2)(\forall \alpha \in]0,1[) \qquad \alpha x + (1-\alpha) y \in C$$

Convex sets ?

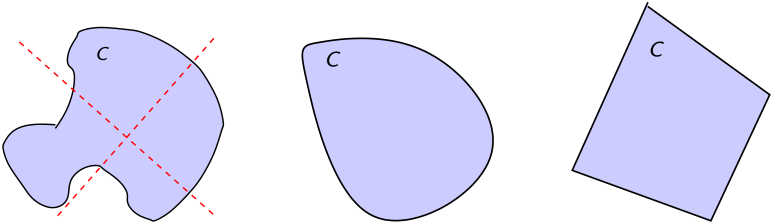


Convex set : definition

Let $\mathcal{H}$ be a Hilbert space. $C \subset \mathcal{H}$ is a convex set if

$$(\forall (x,y) \in C^2)(\forall \alpha \in]0,1[) \qquad \alpha x + (1-\alpha) y \in C$$

Convex sets ?



Convex set : properties

- $\varnothing$ is considered as a convex set.
- If C is a convex set, then $(\forall n \in \mathbb{N}^*)$ $(\forall (x_1, \ldots, x_n) \in C^n)$ $(\forall (\alpha_1, \ldots, \alpha_n) \in [0, +\infty[^n)$ with $\sum_{i=1}^{n} \alpha_i = 1$,
$$\sum_{i=1}^{n} \alpha_i x_i \in C.$$
- Every vector (affine) space is convex.
- If C is a convex set, then $\mathrm{int}\,(C)$ and $\overline{C}$ are convex sets.

Convex set : properties

- If C is a convex set then, for every $\alpha \in \mathbb{R}$,

$$\alpha C = \{\alpha x \mid x \in C\}$$

  is a convex set.

- If C_1 and C_2 are convex sets, then

$$C_1 \times C_2$$
$$C_1 + C_2 = \{x_1 + x_2 \mid (x_1, x_2) \in C_1 \times C_2\}$$

  are convex sets.

- If $(C_i)_{i\in\mathcal{I}}$ is a family of convex sets of $\mathcal{H}$, then $\bigcap_{i\in I} C_i$ is convex.

Convex hull

Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$. The **convex hull** of C is the smallest convex set including C. It is denoted by $\operatorname{conv}(C)$.

- $\operatorname{conv}(C)$ is the intersection of all the convex sets including C.
- Let $x \in \mathcal{H}$. $x \in \operatorname{conv}(C)$ if and only if $(\exists n \in \mathbb{N}^*)$ $(\exists (x_1, \ldots, x_n) \in C^n)$ $(\exists (\alpha_1, \ldots, \alpha_n) \in \,]0, +\infty[^n$ with $\sum_{i=1}^{n} \alpha_i = 1$ such that

$$x = \sum_{i=1}^{n} \alpha_i x_i.$$

Convex function : definitions

$f : \mathcal{H} \to]-\infty, +\infty]$ is a convex function if

$$\big(\forall (x, y) \in \mathcal{H}^2\big)(\forall \alpha \in]0, 1[)$$
$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Convex function : definitions

$f : \mathcal{H} \to \,]-\infty, +\infty]$ is a convex function if

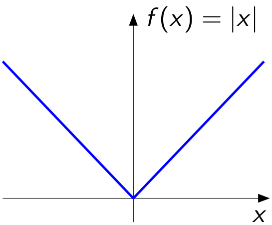
$$\big(\forall (x, y) \in (\operatorname{dom} f)^2\big)(\forall \alpha \in]0, 1[)$$
$$f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y)$$

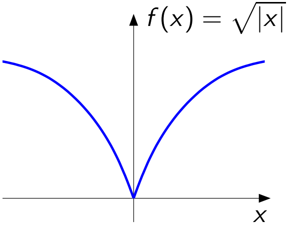
Convex function : definitions

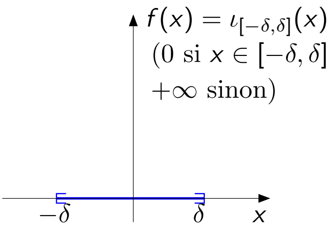
$f : \mathcal{H} \to \,]-\infty, +\infty]$ is a **convex function** if

$$\big(\forall (x,y) \in (\mathrm{dom}\, f)^2\big)(\forall \alpha \in]0,1[)$$
$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Convex functions ?





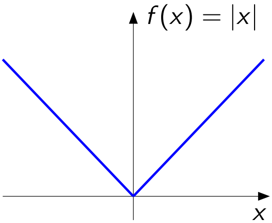


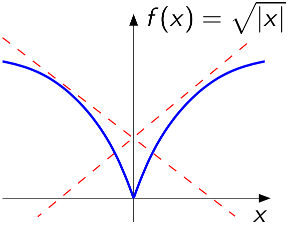
Convex function : definitions

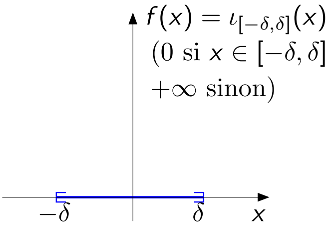
$f : \mathcal{H} \to \,]-\infty, +\infty]$ is a **convex function** if

$$\big(\forall (x,y) \in (\operatorname{dom} f)^2\big)(\forall \alpha \in]0,1[)$$
$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Convex functions ?







Convex function : definitions

$f : \mathcal{H} \to \,]-\infty, +\infty]$ is a **convex function** if

$$\big(\forall (x, y) \in (\mathrm{dom}\, f)^2\big)(\forall \alpha \in]0, 1[)$$
$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

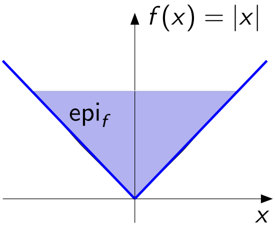
- $f : \mathcal{H} \to [-\infty, +\infty[$ is concave if $-f$ is convex.

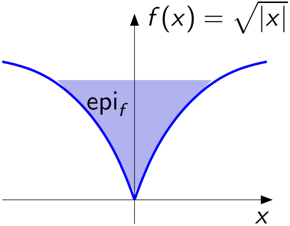
Convex functions : definition

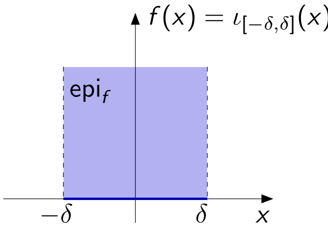
$f : \mathcal{H} \to \,]-\infty, +\infty]$ is convex $\Leftrightarrow$ its epigraph is a convex set.

Convex functions : definition

$f : \mathcal{H} \to \,]-\infty, +\infty]$ is convex $\Leftrightarrow$ its epigraph is a convex set.

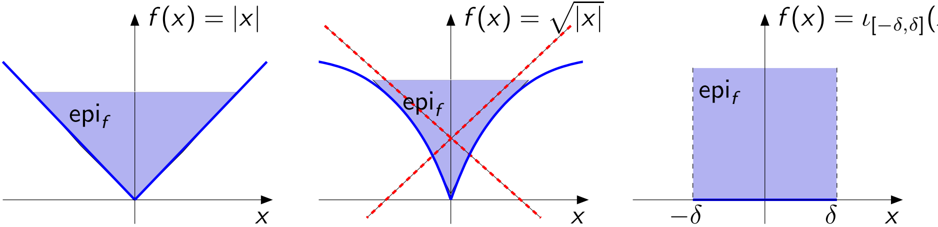






Convex functions : definition

$f : \mathcal{H} \to \,]-\infty, +\infty]$ is convex $\Leftrightarrow$ its epigraph is a convex set.



Convex functions : properties

- If $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ is convex, then $\operatorname{dom} f$ is convex and its lower level set at height $\eta \in \mathbb{R}$

$$\operatorname{lev}_{\leq\eta} f = \left\{x \in \mathcal{H} \;\middle|\; f(x) \leq \eta\right\}$$

  is a convex set.

- $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ is convex if and only if $(\forall(x,y) \in (\operatorname{dom} f)^2)$ $\varphi_{x,y}\colon [0,1] \to \left]-\infty,+\infty\right] : \alpha \mapsto f(\alpha x + (1-\alpha)y)$ is convex.

Convex functions : properties

- Every finite sum of convex functions is convex.
- Let $(f_i)_{i\in I}$ be a family of convex functions. Then, $\sup_{i\in I} f_i$ is convex.
- $\Gamma_0(\mathcal{H})$: class of convex, l.s.c., and proper functions from $\mathcal{H}$ to $]-\infty,+\infty]$.
- Let $C \subset \mathcal{H}$.
  $\iota_C \in \Gamma_0(\mathcal{H}) \Leftrightarrow C$ is a nonempty closed convex set.
  Proof : $\text{epi}_{\iota_C} = C \times [0,+\infty[$.

Strictly convex functions

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to\ ]-\infty,+\infty]$.
f is strictly convex if

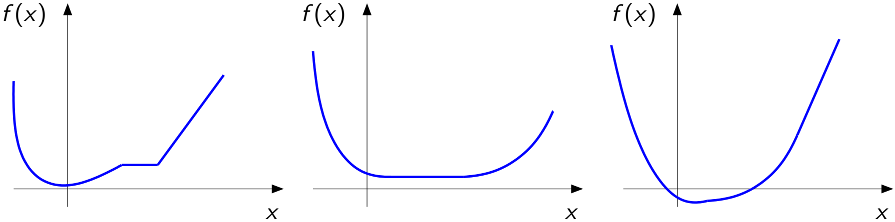
$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{dom} f)(\forall \alpha \in]0,1[)$$
$$x \neq y \quad \Rightarrow \quad f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y).$$

Strictly convex functions

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$.
f is **strictly convex** if

$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{dom} f)(\forall \alpha \in]0,1[)$$
$$x \neq y \quad \Rightarrow \quad f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y).$$

Strictly convex functions ?

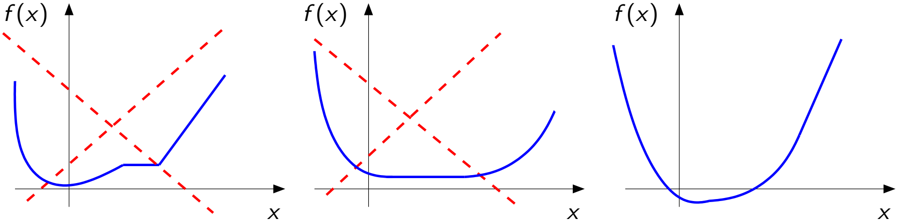


Strictly convex functions

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$.
f is strictly convex if

$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{dom} f)(\forall \alpha \in]0,1[)$$
$$x \neq y \quad \Rightarrow \quad f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y).$$

Strictly convex functions ?



Minimizers of a convex function

Theorem

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a proper convex function such that $\mu = \inf f > -\infty$.

- $\{x \in \mathcal{H} \mid f(x) = \mu\}$ is convex.
- Every local minimizer of f is a global minimizer.
- If f is strictly convex, then there exists at most one minimizer.

Minimizers of a convex function

Theorem

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a proper convex function such that $\mu = \inf f > -\infty$.

- $\{x \in \mathcal{H} \mid f(x) = \mu\}$ is convex.
- Every local minimizer of f is a global minimizer.
- If f is strictly convex, then there exists at most one minimizer.

Proof : Let $\Omega = \{x \in \mathcal{H} \mid f(x) = \mu\}$. Let $(x, y) \in \Omega^2$ and let $\alpha \in [0, 1]$. We have

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) = \mu$$

which shows that $\alpha x + (1-\alpha)y \in \Omega$.

Minimizers of a convex function

Theorem

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a proper convex function such that $\mu = \inf f > -\infty$.

- $\{x \in \mathcal{H} \mid f(x) = \mu\}$ is convex.
- Every local minimizer of f is a global minimizer.
- If f is strictly convex, then there exists at most one minimizer.

Proof : Let $\widehat{x}$ be a local minimizer of f. For every $y \in \mathcal{H} \setminus \{\widehat{x}\}$, there exists $\alpha \in]0,1[$ such that

$$\begin{aligned} & f(\widehat{x}) \le f\big(\widehat{x} + \alpha(y - \widehat{x})\big) \le (1-\alpha) f(\widehat{x}) + \alpha f(y) \\ \Rightarrow \quad & f(\widehat{x}) \le f(y) \end{aligned}$$

If f is strictly convex, the inequality is strict.

Existence and uniqueness of a minimizer

Theorem

Let $\mathcal{H}$ be a Hilbert space and C a closed convex subset of $\mathcal{H}$. Let $f \in \Gamma_0(\mathcal{H})$ such that $\operatorname{dom} f \cap C \neq \varnothing$.

If f is coercive or C is bounded, then there exists $\widehat{x} \in C$ such that

$$f(\widehat{x}) = \inf_{x \in C} f(x).$$

If, moreover, f is strictly convex, this minimizer $\widehat{x}$ is unique.

Exercise 2

Let $\mathcal{H}$ be a Hilbert space.

1. Show that the function $x \mapsto \|x\|^2$ is strictly convex.
2. A function $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ is *strongly convex* with modulus $\beta \in \,]0, +\infty[$ if there exists a convex function $g\colon \mathcal{H} \to \,]-\infty, +\infty]$ such that
$$f = g + \frac{\beta}{2}\|\cdot\|^2.$$
Show that that every strongly convex function is strictly convex.
3. Show that a function $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ is strongly convex with modulus $\beta \in \,]0, +\infty[$ if and only if
$$(\forall (x, y) \in \mathcal{H}^2)(\forall \alpha \in]0, 1[)$$
$$f(\alpha x + (1-\alpha)y) + \alpha(1-\alpha)\frac{\beta}{2}\|x - y\|^2 \leq \alpha f(x) + (1-\alpha)f(y).$$

Convex + smooth

Characterization of differentiable convex functions

Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be differentiable on $\operatorname{dom} f$, which is a nonempty open convex set.
Then, f is convex if and only if

$$(\forall (x, y) \in (\operatorname{dom} f)^2) \quad f(y) \geq f(x) + \langle \nabla f(x) \mid y - x \rangle \,.$$

Characterization of differentiable strictly convex functions

Let $f\colon \mathcal{H} \to\ ]-\infty,+\infty]$ be differentiable on $\operatorname{dom} f$, which is a nonempty open convex set.
Then, f is strictly convex if and only if, for every $(x,y) \in (\operatorname{dom} f)^2$ with $x \neq y$,

$$f(y) > f(x) + \langle \nabla f(x) \mid y - x \rangle \,.$$

Characterization of differentiable convex functions

Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be differentiable on $\operatorname{dom} f$, which is a nonempty open convex set.
Then, f is convex if and only if ∇f is monotone on $\operatorname{dom} f$, i.e.

$$(\forall (x, y) \in (\operatorname{dom} f)^2) \quad \langle \nabla f(y) - \nabla f(x) \mid y - x \rangle \geq 0.$$

Characterization of strictly differentiable convex functions

Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ be differentiable on $\operatorname{dom} f$, which is a nonempty open convex set.
Then, f is strictly convex if and only if ∇f is strictly monotone on $\operatorname{dom} f$, i.e. for every $(x,y) \in (\operatorname{dom} f)^2$ with $x \neq y$,

$$\langle \nabla f(y) - \nabla f(x) \mid y - x \rangle > 0.$$

Characterization of twice differentiable convex functions

Let $\mathcal{H}$ be a Hilbert space.
Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a twice differentiable function on $\operatorname{dom} f$, which is a nonempty open convex set.

- f is convex if and only if, for every $x \in \operatorname{dom} f$, $\nabla^2 f(x)$ is positive semi-definite.
- If, for every $x \in \operatorname{dom} f$, $\nabla^2 f(x)$ is positive definite, then f is strictly convex.

Condition for the existence of a minimizer

Theorem

Let $\mathcal{H}$ be Hilbert space.
Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a differentiable convex function on $\operatorname{dom} f$, which is an open set. Let $C \subset \operatorname{dom} f$ be a nonempty convex set. $\widehat{x} \in C$ is a (global) minimizer of f on C if and only if

$$(\forall y \in C) \qquad \langle \nabla f(\widehat{x}) \mid y - \widehat{x} \rangle \geq 0.$$

If $\widehat{x} \in \operatorname{int}(C)$, then the condition reduces to

$$\nabla f(\widehat{x}) = 0.$$

Condition for the existence of a minimizer

Theorem

Let $\mathcal{H}$ be Hilbert space.
Let $f\colon \mathcal{H} \to\ ]-\infty,+\infty]$ be a differentiable convex function on $\operatorname{dom} f$, which is an open set. Let $C \subset \operatorname{dom} f$ be a nonempty convex set. $\widehat{x} \in C$ is a (global) minimizer of f on C if and only if

$$(\forall y \in C) \qquad \langle \nabla f(\widehat{x}) \mid y - \widehat{x}\rangle \geq 0.$$

If $\widehat{x} \in \operatorname{int}(C)$, then the condition reduces to

$$\nabla f(\widehat{x}) = 0.$$

Proof : We have already seen that the inequality is a necessary condition for $\widehat{x}$ to be a local minimizer of f on C and that it reduces to the vanishing condition on the gradient if $\widehat{x} \in \operatorname{int}(C)$.

Condition for the existence of a minimizer

Theorem

Let $\mathcal{H}$ be Hilbert space.
Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a differentiable convex function on $\operatorname{dom} f$, which is an open set. Let $C \subset \operatorname{dom} f$ be a nonempty convex set. $\widehat{x} \in C$ is a (global) minimizer of f on C if and only if

$$(\forall y \in C) \qquad \langle \nabla f(\widehat{x}) \mid y - \widehat{x} \rangle \geq 0.$$

If $\widehat{x} \in \operatorname{int}(C)$, then the condition reduces to

$$\nabla f(\widehat{x}) = 0.$$

<u>Proof</u> : Conversely, assume that the inequality holds. Let $y \in C$. Since f is convex and Gâteaux differentiable,

$$f(y) \geq f(\widehat{x}) + \langle \nabla f(\widehat{x}) \mid y - \widehat{x} \rangle \geq f(\widehat{x}).$$

Hence, $\widehat{x}$ is a minimizer of f on C.

Exercice 3

Let $f : \mathcal{H} \to \mathbb{R}$ be differentiable.
Show that f is β-strongly convex if and only if

$$(\forall (x, y) \in \mathcal{H}^2) \quad f(y) \geq f(x) + \langle \nabla f(x) \mid y - x \rangle + \frac{\beta}{2} \|y - x\|^2.$$

Exercice 4

Let

$$\begin{aligned} f: \quad & \mathbb{R}^N \to \mathbb{R} \\ (x^{(i)})_{1\leq i\leq N} & \mapsto \ln\left(\sum_{i=1}^{N} \exp(x^{(i)})\right). \end{aligned}$$

Show that f is convex. Is it strictly convex ?

Projections

Projection onto a closed convex set

Theorem

Let C be a nonempty closed convex set of a Hilbert space $\mathcal{H}$.

(i) For every $x \in \mathcal{H}$, there exists a unique point $\widehat{x}$ in C which lies at minimum distance of x. The application $P_C \colon \mathcal{H} \to C$ which maps every $x \in \mathcal{H}$ to its associated point $\widehat{x}$ is called the projection onto C.

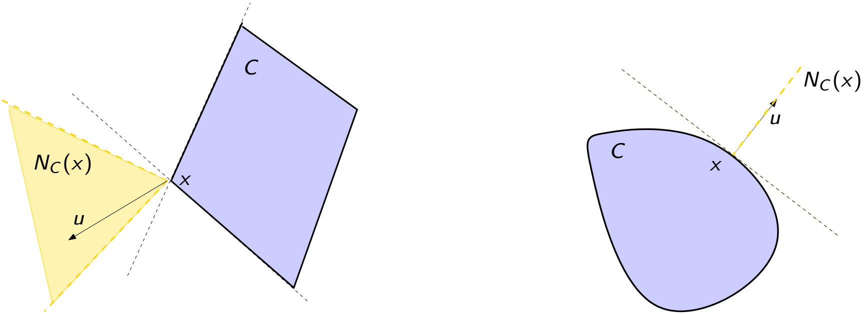
(ii) For every $x \in \mathcal{H}$, $\widehat{x} = P_C(x)$ if and only if $\widehat{x} \in C$ and

$$(\forall y \in C) \qquad \langle x - \widehat{x} \mid y - \widehat{x} \rangle \leq 0.$$

Geometrical interpretation

Let C be a nonempty subset of $\mathcal{H}$.
For every $x \in \mathcal{H}$, the normal cone to C at x is defined as

$$N_C(x) = \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \ \ \langle u \mid y - x \rangle \leq 0\} & \text{if } x \in C \\ \varnothing & \text{otherwise.} \end{cases}$$



Geometrical interpretation

Let C be a nonempty subset of $\mathcal{H}$.
For every $x \in \mathcal{H}$, the normal cone to C at x is defined as

$$N_C(x) = \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \ \ \langle u \mid y - x \rangle \leq 0\} & \text{if } x \in C \\ \varnothing & \text{otherwise.} \end{cases}$$

- If $x \in \operatorname{int} C$, then $N_C(x) = \{0\}$.
- If C is a vector space, then for every $x \in C$, $N_C(x) = C^{\perp}$.
- Let C be nonempty closed convex set of a Hilbert space $\mathcal{H}$. For every $x \in \mathcal{H}$,

$$\widehat{x} = P_C(x) \quad \Leftrightarrow \quad x - \widehat{x} \in N_C(\widehat{x}).$$

Examples of projections

- If C is a (closed) vector space of a Hilbert space $\mathcal{H}$, then P_C is the (linear) orthogonal projection onto C. Then, for every $x \in \mathcal{H}$,

$$\widehat{x} = P_C(x) \quad \Leftrightarrow \quad \begin{cases} \widehat{x} \in C \\ x - \widehat{x} \in C^{\perp}. \end{cases}$$

Properties of the projection

- Let C be a nonempty closed convex set of a Hilbert space $\mathcal{H}$. The projection onto C is a firmly nonexpansive operator, i.e.

$$(\forall (x,y) \in \mathcal{H}^2) \qquad \|P_C(x) - P_C(y)\|^2 \leq \langle x - y \mid P_C(x) - P_C(y) \rangle \,.$$

Properties of the projection

- Let C be a nonempty closed convex set of a Hilbert space $\mathcal{H}$. The projection onto C is a firmly nonexpansive operator, i.e.

$$(\forall (x,y) \in \mathcal{H}^2) \qquad \|P_C(x) - P_C(y)\|^2 \leq \langle x - y \mid P_C(x) - P_C(y) \rangle \,.$$

Properties of the projection

- Let C be a nonempty closed convex set of a Hilbert space $\mathcal{H}$. The projection onto C is a firmly nonexpansive operator, i.e.

$$(\forall (x,y) \in \mathcal{H}^2) \qquad \|P_C(x) - P_C(y)\|^2 \leq \langle x - y \mid P_C(x) - P_C(y) \rangle \, .$$

- The projection onto C is a nonexpansive operator, i.e.

$$(\forall (x,y) \in \mathcal{H}^2) \qquad \|P_C(x) - P_C(y)\| \leq \|x - y\|.$$

- The projection onto C is uniformly continuous.
- The distance to C defined as

$$(\forall x \in \mathcal{H}) \qquad d_C(x) = \inf_{y \in C} \|x - y\| = \|x - P_C(x)\|$$

is continuous.

Exercise 5

Let $\mathcal{H}$ and $\mathcal{G}$ be Hilbert spaces.
Let D be a nonempty closed convex set of $\mathcal{G}$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ be a bijective isometry and let

$$C = \left\{x \in \mathcal{H} \;\middle|\; Lx \in D\right\}.$$

1. Show that the projection onto C is well-defined.
2. Show that $P_C = L^* \circ P_D \circ L$.
3. Express P_C when $\mathcal{H} = \mathcal{G} = \mathbb{R}^N$ and $D = [0, +\infty[^N$.