CentraleSupelec
ST7 – Optimization
Part III : Convexity

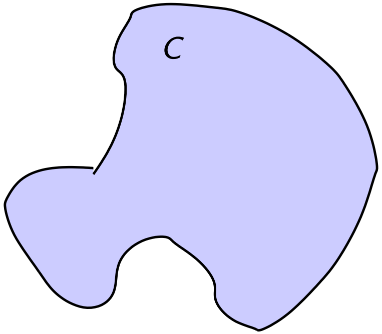
jean-christophe@pesquet.eu

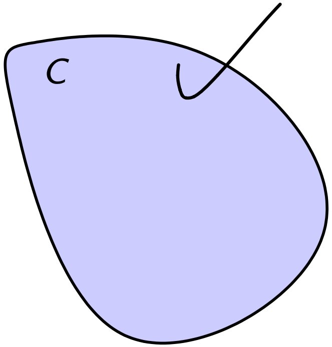
Convex set : definition

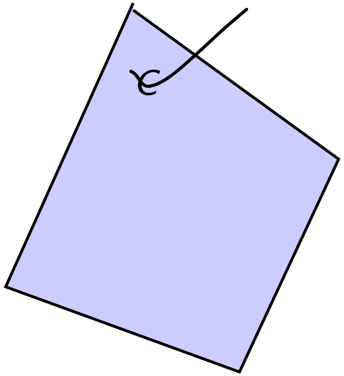
Let $\mathcal{H}$ be a Hilbert space. $C \subset \mathcal{H}$ is a convex set if

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

Convex sets ?





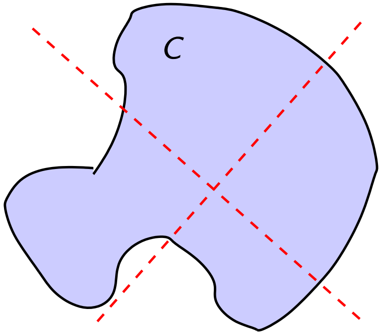


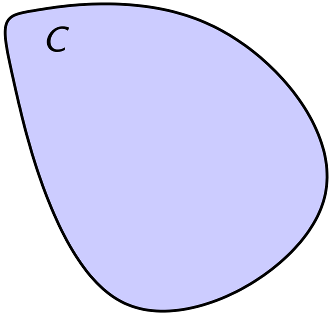
Convex set : definition

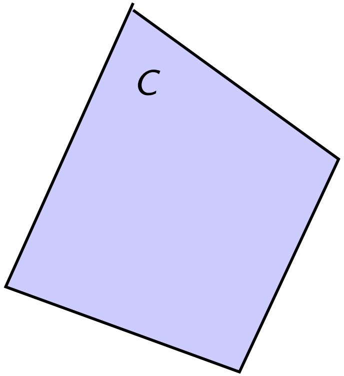
Let $\mathcal{H}$ be a Hilbert space. $C \subset \mathcal{H}$ is a convex set if

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

Convex sets ?







Convex set : properties

凸集 .

- $\varnothing$ is considered as a convex set.
- If C is a convex set, then $(\forall n \in \mathbb{N}^*)$ $(\forall (x_1, \dots, x_n) \in C^n)$ $(\forall (\alpha_1, \dots, \alpha_n) \in [0, +\infty[^n)$ 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 $\operatorname{int}(C)$ and $\overline{C}$ are convex sets.

Convex set : properties

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

  $$\alpha C = \left\{ \alpha x \;\middle|\; x \in C \right\}$$

  is a convex set.
- If C_1 and C_2 are convex sets, then

  $$\left\{ \begin{array}{l} C_1 \boxed{\times} C_2 \\ C_1 \boxed{+} C_2 = \left\{ x_1 + x_2 \;\middle|\; (x_1, x_2) \in C_1 \times C_2 \right\} \end{array} \right.$$

  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 凸函数.

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

$$\underline{f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)}$$

Convex function : definitions

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

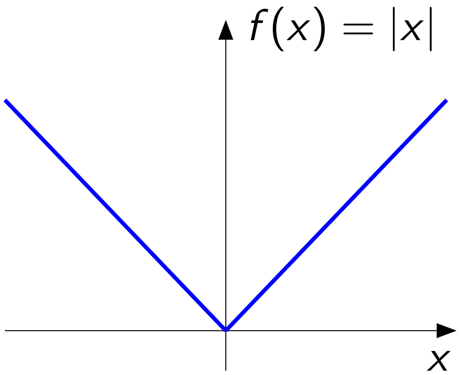
$$(\forall (x,y) \in (\mathrm{dom}\, f)^2)(\forall \alpha \in]0,1[) \quad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

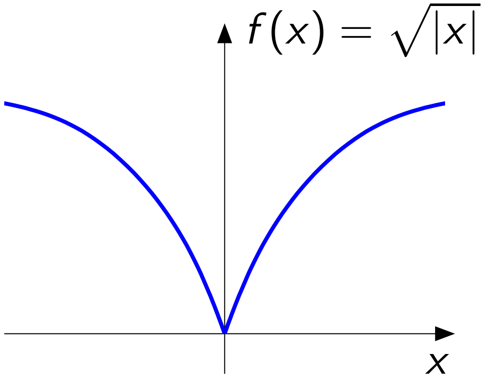
Convex function : definitions

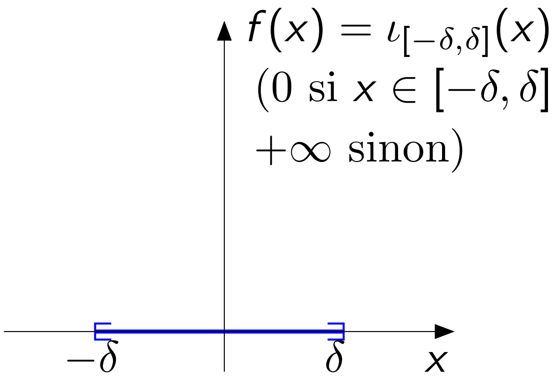
$f : \mathcal{H} \to [-\infty, +\infty]$ is a convex function if

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

Convex functions ?







Convex function : definitions

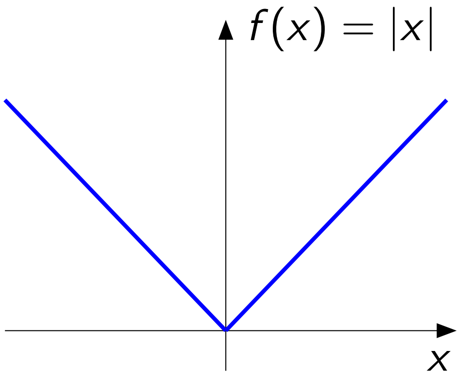
$f : \mathcal{H} \to [-\infty, +\infty]$ is a **convex function** if

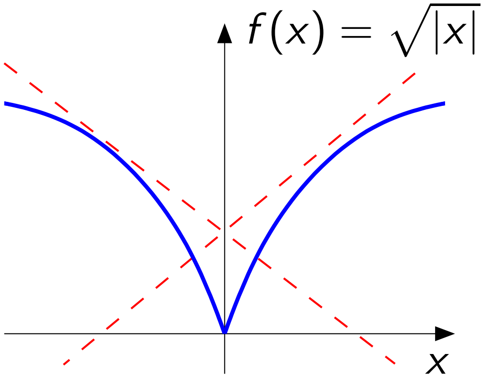
$$(\forall (x,y) \in (\underline{\mathrm{dom}\, f})^2)(\forall \alpha \in]0,1[)$$

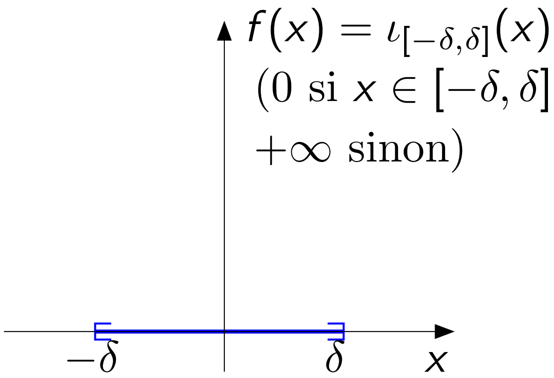
$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

domain of f.

Convex functions ?







Convex function : definitions

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

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

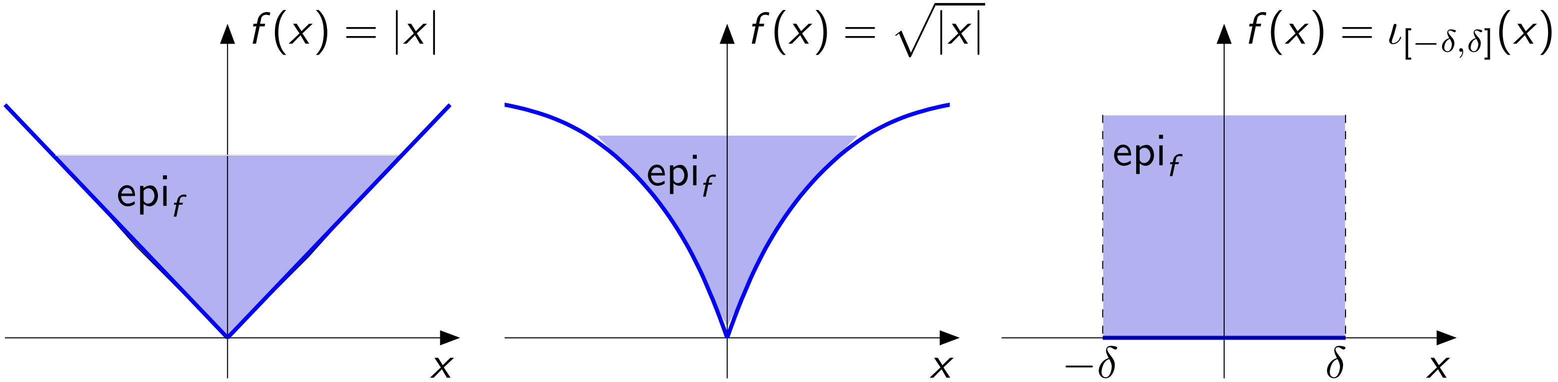
- $f : \mathcal{H} \to [-\infty, +\infty]$ is concave if $-f$ is convex.

Convex functions : definition

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

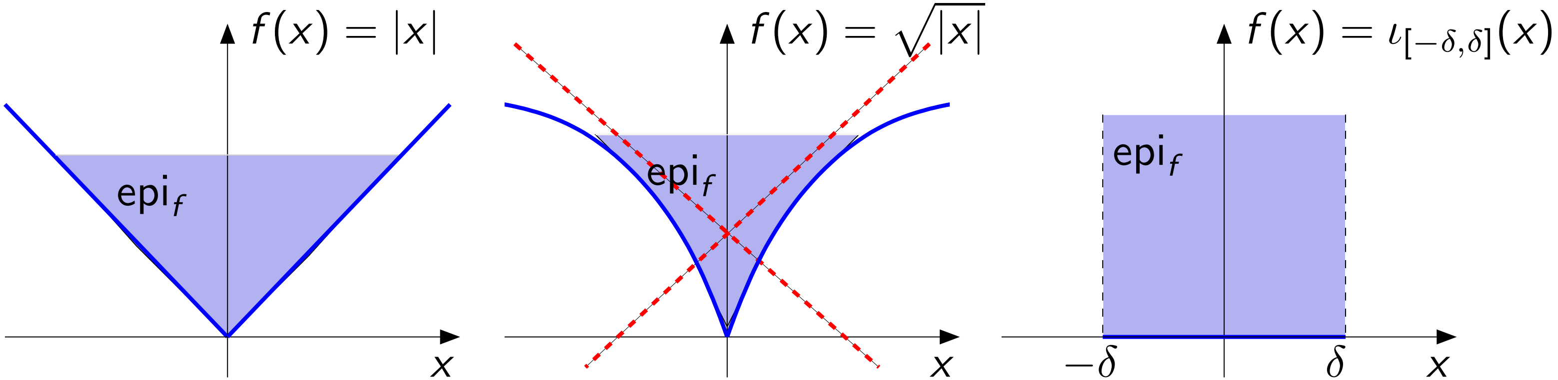
Convex functions : definition

$f : \mathcal{H} \to [-\infty, +\infty]$ 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 : definition

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

Proof : Assume that epi f is a convex set. For every $(x, y) \in (\mathrm{dom}\, f)^2$ and $(\eta, \rho) \in \mathbb{R}^2$ such that $\eta \geq f(x)$ and $\rho \geq f(y)$, $(x, \eta) \in \mathrm{epi}\, f$ and $(y, \rho) \in \mathrm{epi}\, f$.
Then, for every $\alpha \in]0, 1[$,

$$\underline{\alpha(x, \eta) + (1-\alpha)(y, \rho) = (\alpha x + (1-\alpha)y, \alpha\eta + (1-\alpha)\rho) \in \mathrm{epi}\, f,}$$

which means that

$$f(\alpha x + (1-\alpha)y) \leq \alpha\eta + (1-\alpha)\rho.$$

By letting η tend to $f(x)$ and ρ tend to $f(y)$,

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

Convex functions : definition

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

Proof : Assume that f is convex. Let $(x, \eta) \in \operatorname{epi} f$ and $(y, \rho) \in \operatorname{epi} f$. Then, $f(x) \leq \eta$ and $f(y) \leq \rho$ and, for every $\alpha \in]0, 1[$,

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \leq \alpha\eta + (1 - \alpha)\rho.$$

Therefore, $\alpha(x, \eta) + (1 - \alpha)(y, \rho) \in \operatorname{epi} f$ and $\operatorname{epi} f$ is a convex set.

Convex functions : properties

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

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

  is a convex set.

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

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

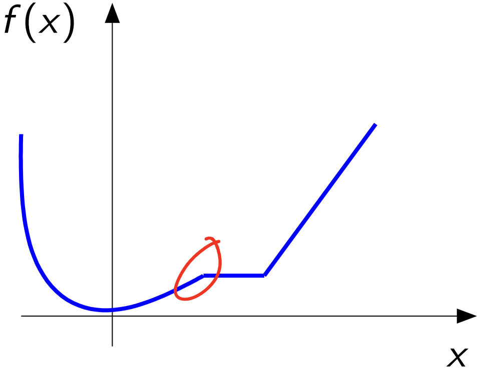
Strictly convex functions

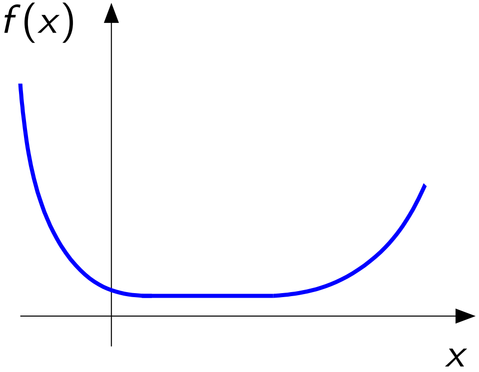
Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$.

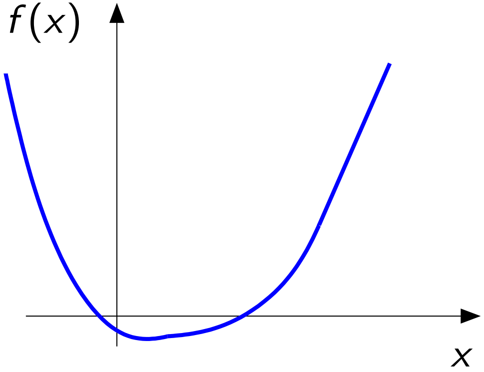
f is **strictly convex** if

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

Strictly convex functions ?





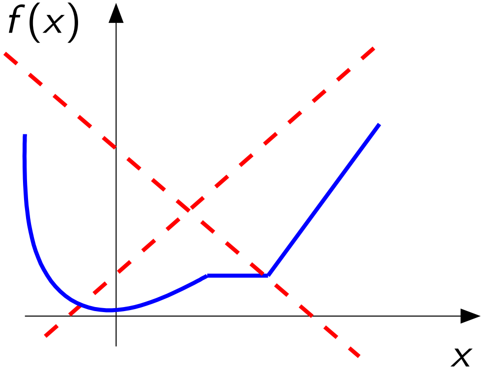


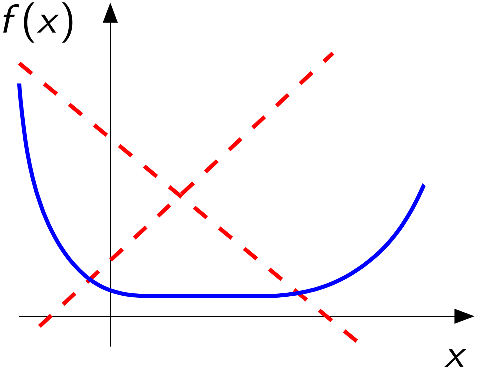
Strictly convex functions

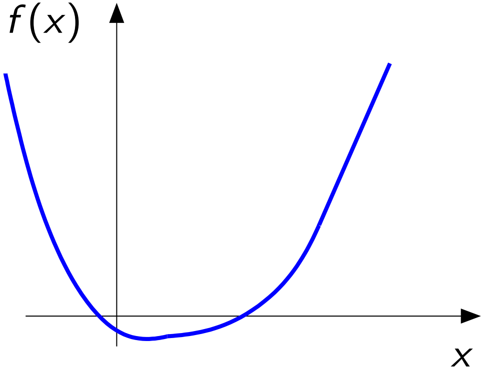
Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$.
f is **strictly convex** if

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

Strictly convex functions ?







Minimizers of a convex function

Theorem

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

- $\left\{x \in \mathcal{H} \;\middle|\; f(x) = \mu\right\}$ 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 $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ be a proper convex function such that $\mu = \inf f > -\infty$.

- $\left\{x \in \mathcal{H} \;\middle|\; f(x) = \mu\right\}$ 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 $\Omega = \left\{x \in \mathcal{H} \;\middle|\; f(x) = \mu\right\}$. 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.

Proof : Although the result is valid for infinite dimensional spaces, we will prove it only when $\mathcal{H}$ is finite dimensional.
Since C is a nonempty closed and convex set, $\iota_C \in \Gamma_0(\mathcal{H})$. Then, $f + \iota_C$ is l.s.c. and convex. Since $\operatorname{dom} f \cap C \neq \varnothing$, $f + \iota_C \in \Gamma_0(\mathcal{H})$. In addition, as f is coercive or C is bounded, $f + \iota_C$ is coercive. We are thus guaranteed that there exists $\widehat{x} \in C$ such that $f(\widehat{x}) = \inf_{x \in \mathcal{H}} f(x) + \iota_C(x) = \inf_{x \in C} f(x)$.
If f is strictly convex, then $f + \iota_C$ is strictly convex and there is a unique minimizer.

Exercise 1

Let $(a_k)_{1\le k\le K}$ be of vectors of $\mathbb{R}^N$. Show that

$$C = \{\sum_{k=1}^{K} \xi_k a_k \mid (\forall k \in \{1,\ldots,K\})\ \xi_k \ge 0\}$$

is a nonempty convex cone of $\mathbb{R}^N$.
When the vectors $(a_k)_{1\le k\le K}$ are independent, show that C is closed.
(Actually, the latter result remains true for any family $(a_k)_{1\le k\le K}$ of vectors of $\mathbb{R}^N$.)

Exercise 2

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

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

Characterization of differentiable convex functions

Let $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ be Gâteaux 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 convex functions

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

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

Proof :
Assume that f is convex. Let $x \in \operatorname{dom} f$. For every $\alpha \in [0,1]$ and $y \in \mathcal{H}$,

$$\begin{aligned} & f\big(x + \alpha(y-x)\big) \leq (1-\alpha) f(x) + \alpha f(y) \\ \Rightarrow \quad & \langle \nabla f(x) \mid y - x \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha > 0}} \frac{f\big(x + \alpha(y-x)\big) - f(x)}{\alpha} \leq f(y) - f(x). \end{aligned}$$

Characterization of differentiable convex functions

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

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

Proof :
Conversely, if the gradient inequality is satisfied, we have, for every $(x,y) \in (\operatorname{dom} f)^2$ and $\alpha \in [0,1]$, $\alpha x + (1-\alpha) y \in \operatorname{dom} f$, and

$$f(x) \geq f(\alpha x + (1-\alpha)y) + (1-\alpha)\left\langle \nabla f(\alpha x + (1-\alpha)y) \mid x-y \right\rangle$$
$$f(y) \geq f(\alpha x + (1-\alpha)y) + \alpha\left\langle \nabla f(\alpha x + (1-\alpha)y) \mid y-x \right\rangle .$$

By multiplying the first inequality by α and the second one by $1-\alpha$ and summing them, we get

$$\alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha)y).$$

Characterization of differentiable strictly convex functions

Let $f\colon \mathcal{H} \to\ ]-\infty,+\infty]$ be Gâteaux 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 strictly convex functions

Let $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ be Gâteaux 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 $x \neq y$,

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

Proof :
Assume that f is strictly convex. Let $(x,y) \in \operatorname{dom} f$ with $x \neq y$. Let $\alpha \in]0,1[$ and let $z = \alpha x + (1-\alpha) y \in \operatorname{dom} f$. Since f is convex,

$$\begin{aligned} & f(z) \geq f(x) + \langle \nabla f(x) \mid z - x \rangle \\ \Leftrightarrow \quad & f(\alpha x + (1-\alpha)y) \geq f(x) + (1-\alpha) \langle \nabla f(x) \mid y - x \rangle . \end{aligned}$$

By using the inequality $\alpha f(x) + (1-\alpha) f(y) > f(\alpha x + (1-\alpha) y)$, we deduce that

$$\begin{aligned} & \alpha f(x) + (1-\alpha) f(y) > f(x) + (1-\alpha) \langle \nabla f(x) \mid y - x \rangle \\ \Rightarrow \quad & f(y) > f(x) + \langle \nabla f(x) \mid y - x \rangle . \end{aligned}$$

Characterization of differentiable strictly convex functions

Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be Gâteaux 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 $x \neq y$,

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

Proof :

Conversely, if the gradient inequality is satisfied, we have, for every $(x, y) \in (\operatorname{dom} f)^2$ with $x \neq y$ and $\alpha \in]0, 1[$, we have $\alpha x + (1-\alpha)y \in \operatorname{dom} f \setminus \{x, y\}$, and

$$\begin{aligned} f(x) &> f(\alpha x + (1-\alpha)y) + (1-\alpha)\langle \nabla f(\alpha x + (1-\alpha)y) \mid x - y \rangle \\ f(y) &> f(\alpha x + (1-\alpha)y) + \alpha\langle \nabla f(\alpha x + (1-\alpha)y) \mid y - x \rangle\,. \end{aligned}$$

By multiplying the first inequality by α and the second one by $1 - \alpha$ and summing them, we get

$$\alpha f(x) + (1-\alpha) f(y) > f(\alpha x + (1-\alpha)y).$$

Characterization of differentiable convex functions

Let $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ be Gâteaux 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 differentiable convex functions

Let $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ be Gâteaux 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.

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

Proof :
Assume that f is convex. For every $(x,y) \in (\operatorname{dom} f)^2$,

$$\begin{aligned} f(x) - f(y) &\geq \langle \nabla f(y) \mid x - y\rangle \\ f(y) - f(x) &\geq \langle \nabla f(x) \mid y - x\rangle . \end{aligned}$$

By summing, $\langle \nabla f(y) - \nabla f(x) \mid x - y\rangle \leq 0$.

Characterization of differentiable convex functions

Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ be Gâteaux 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.

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

Proof :

Conversely, assume that ∇f is monotone on $\operatorname{dom} f$. For every $(x,y) \in (\operatorname{dom} f)^2$, let $\varphi\colon [0,1] \to \mathbb{R}\colon \alpha \mapsto f(x+\alpha(y-x))$.
φ is continuous on $[0,1]$, differentiable on $]0,1[$ and its derivative is

$$(\forall \alpha \in]0,1[) \qquad \dot{\varphi}(\alpha) = \langle \nabla f(x+\alpha(y-x)) \mid y-x \rangle\,.$$

According to the mean value theorem, there exists $\alpha \in]0,1[$ such that

$$\left\{\begin{aligned} & & \varphi(1) - \varphi(0) &= \dot{\varphi}(\alpha) \\ &\Rightarrow & f(y) - f(x) &= \langle \nabla f(x+\alpha(y-x)) \mid y-x \rangle\,. \end{aligned}\right.$$

On the other hand, as $x+\alpha(y-x) \in \operatorname{dom} f$,

$$? \quad \langle \nabla f(x+\alpha(y-x)) - \nabla f(x) \mid y-x \rangle \geq 0$$

which leads to $f(y) - f(x) \geq \langle \nabla f(x) \mid y-x \rangle$.

Characterization of strictly differentiable convex functions

Let $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ be Gâteaux 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 $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ be a twice Fréchet differentiable function on $\operatorname{dom} f$, which is a nonempty open convex set.

- f is convex if and only if, for every $x \in \operatorname{dom} f$,
$$(\forall z \in \mathcal{H}) \qquad \langle z \mid \nabla^2 f(x) z\rangle \geq 0.$$
- If, for every $x \in \operatorname{dom} f$,
$$(\forall z \in \mathcal{H} \setminus \{0\}) \qquad \langle z \mid \nabla^2 f(x) z\rangle > 0,$$
then f is strictly convex.

Characterization of twice differentiable convex functions

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

- f is convex if and only if, for every $x \in \operatorname{dom} f$,
$$(\forall z \in \mathcal{H}) \qquad \left\langle z \mid \nabla^2 f(x) z \right\rangle \geq 0.$$
- If, for every $x \in \operatorname{dom} f$,
$$(\forall z \in \mathcal{H} \setminus \{0\}) \qquad \left\langle z \mid \nabla^2 f(x) z \right\rangle > 0,$$
then f is strictly convex.

Proof : Let $(x, y) \in (\operatorname{dom} f)^2$ with $x \neq y$. According to Taylor-Mc Laurin formula, there exists $u \in]x, y[\subset \operatorname{dom} f$ such that

$$f(y) = f(x) + \langle \nabla f(x) \mid y - x \rangle + \frac{1}{2} \left\langle y - x \mid \nabla^2 f(u)(y - x) \right\rangle .$$

If $\left\langle y - x \mid \nabla^2 f(u)(y - x) \right\rangle \geq 0$, then $f(y) \geq f(x) + \langle \nabla f(x) \mid y - x \rangle$.
If $\left\langle y - x \mid \nabla^2 f(u)(y - x) \right\rangle > 0$, then $f(y) > f(x) + \langle \nabla f(x) \mid y - x \rangle$.

Characterization of twice differentiable convex functions

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

- f is convex if and only if, for every $x \in \operatorname{dom} f$,

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

- If, for every $x \in \operatorname{dom} f$,

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

  then f is strictly convex.

Proof : Conversely, if f is convex. Since $\operatorname{dom} f$ is open, for every $x \in \operatorname{dom} f$ and $z \in \mathcal{H}$, there exists $\delta \in \left]0, +\infty\right[$ such that $(\forall \alpha \in]0, \delta[)\ x + \alpha z \in \operatorname{dom} f$ and

$$\begin{aligned} &\langle \nabla f(x + \alpha z) - \nabla f(x) \mid \alpha z \rangle \geq 0 \\ \Leftrightarrow &\langle \alpha^{-1} (\nabla f(x + \alpha z) - \nabla f(x)) \mid z \rangle \geq 0 \end{aligned}$$

By letting $\alpha \to 0$, $\langle \nabla^2 f(x) z \mid z \rangle \geq 0$.

存在最小值/极小值.

Condition for the existence of a minimizer

Theorem

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

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

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 $\widehat{x} \in \operatorname{int}(C)$.

$f(y) \geq f(\hat{x}) + \langle \nabla f(\hat{x}) | y - \hat{x} \rangle$

Condition for the existence of a minimizer

Theorem

Let $\mathcal{H}$ be Hilbert space.
Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ be a Gâteaux 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 : Conversely, assume that the inequality holds. Let $y \in C$. Since f is convex and Gâteaux differentiable,

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

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

投影

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

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

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

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

Proof : (i) Let $x \in \mathcal{H}$. The function $f \colon \mathcal{H} \to \mathbb{R} \colon y \mapsto \|y - x\|^2$ is continuous, strictly convex and coercive. Since C is nonempty $\operatorname{dom} f \cap C = C \neq \varnothing$. Then, there exists a unique minimizer $\widehat{x}$ of f on C.

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

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

Proof : (ii) f is Gâteaux differentiable and

$$(\forall y \in \mathcal{H}) \qquad \nabla f(y) = 2(y - x).$$

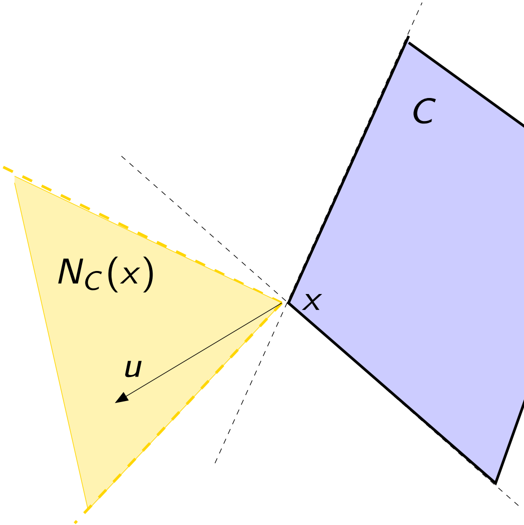
Therefore, $\widehat{x}$ is the minimizer of f onto C if and only if

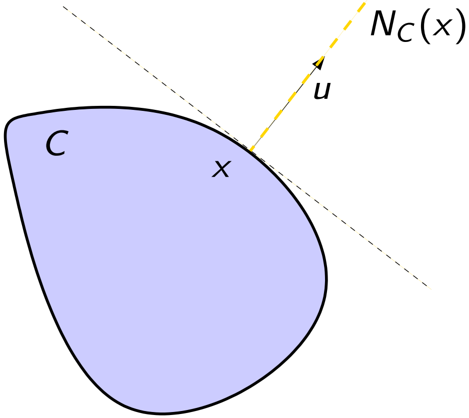
$$(\forall y \in C) \qquad \langle \nabla f(\widehat{x}) \mid y - \widehat{x} \rangle = 2 \left\langle \widehat{x} - x \mid y - \widehat{x} \right\rangle \geq 0.$$

Geometrical interpretation 几何.

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

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

Recall : A set $S \subset \mathcal{H}$ is a cone (with vertex 0) if

$$(\forall x \in S)(\forall \alpha \in]0, +\infty[) \quad \alpha x \in S.$$

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 $x \in \mathcal{H}$,

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

- Let $\mathcal{H}$ be a Hilbert space, let $u \in \mathcal{H} \setminus \{0\}$, let $\delta \in \mathbb{R}$, and let C be the closed affine hyperplane defined as

$$C = \big\{ x \in \mathcal{H} \;\big|\; \langle u \mid x \rangle = \delta \big\}.$$

  The projection onto C is

$$(\forall x \in \mathcal{H}) \qquad P_C(x) = x + \frac{\delta - \langle u \mid x \rangle}{\|u\|^2} u.$$

Examples of projections

- Let $\mathcal{H}$ be a Hilbert space, let $u \in \mathcal{H} \setminus \{0\}$, let $\delta \in \mathbb{R}$, and let C be the closed affine hyperplane defined as

$$C = \big\{ x \in \mathcal{H} \;\big|\; \langle u \mid x \rangle = \delta \big\}.$$

The projection onto C is

$$(\forall x \in \mathcal{H}) \qquad P_C(x) = x + \frac{\delta - \langle u \mid x \rangle}{\|u\|^2} u.$$

Proof : C is a nonempty closed convex set.
For every $x \in \mathcal{H}$,

$$\langle u \mid P_C(x) \rangle = \left\langle u \mid x + \frac{\delta - \langle u \mid x \rangle}{\|u\|^2} u \right\rangle = \delta.$$

Then $P_C(x) \in C$.

Examples of projections

- Let $\mathcal{H}$ be a Hilbert space, let $u \in \mathcal{H} \setminus \{0\}$, let $\delta \in \mathbb{R}$, and let C be the closed affine hyperplane defined as

$$C = \big\{x \in \mathcal{H} \;\big|\; \langle u \mid x\rangle = \delta\big\}.$$

The projection onto C is

$$(\forall x \in \mathcal{H}) \qquad P_C(x) = x + \frac{\delta - \langle u \mid x\rangle}{\|u\|^2} u.$$

Proof : For every $x \in \mathcal{H}$ and $y \in C$,

$$\begin{aligned}\langle x - P_C(x) \mid y - P_C(x)\rangle &= \frac{\langle u \mid x\rangle - \delta}{\|u\|^2} \langle u \mid y - P_C(x)\rangle\\ &= \frac{\langle u \mid x\rangle - \delta}{\|u\|^2}(\langle u \mid y\rangle - \langle u \mid P_C(x)\rangle)\\ &= \frac{\langle u \mid x\rangle - \delta}{\|u\|^2}(\langle u \mid y\rangle - \delta) = 0.\end{aligned}$$

This shows that P_C is the projection onto C.

Examples of projections

- Let $\mathcal{H}$ be a Hilbert space, let $u \in \mathcal{H} \setminus \{0\}$, let $\delta \in \mathbb{R}$, and let C be the closed affine hyperplane defined as

  $$C = \big\{x \in \mathcal{H} \;\big|\; \langle u \mid x\rangle = \delta\big\}.$$

  The projection onto C is

  $$(\forall x \in \mathcal{H}) \qquad P_C(x) = x + \frac{\delta - \langle u \mid x\rangle}{\|u\|^2} u.$$

- Let $\mathcal{H}$ be a Hilbert space, let $u \in \mathcal{H} \setminus \{0\}$, let $\delta \in \mathbb{R}$, and let C be the closed half space defined as

  $$C = \big\{x \in \mathcal{H} \;\big|\; \langle u \mid x\rangle \leq \delta\big\}.$$

  The projection onto C is

  $$(\forall x \in \mathcal{H}) \qquad P_C(x) = \begin{cases} x + \dfrac{\delta - \langle u \mid x\rangle}{\|u\|^2} u & \text{if } \langle u \mid x\rangle > \delta \\ x & \text{otherwise.} \end{cases}$$

Properties of the projection 投影的性质.

- 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.

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

Proof : For every $(x,y) \in \mathcal{H}^2$,

$$\begin{aligned} \langle x - y \mid P_C(x) - P_C(y)\rangle = &\langle x - P_C(x) \mid P_C(x) - P_C(y)\rangle \\ &+ \|P_C(x) - P_C(y)\|^2 \\ &+ \langle P_C(y) - y \mid P_C(x) - P_C(y)\rangle \,. \end{aligned}$$

Since $P_C(x) \in C$ and $P_C(y) \in C$

$$\begin{aligned} \langle x - P_C(x) \mid P_C(y) - P_C(x)\rangle &\leq 0 \\ \langle y - P_C(y) \mid P_C(x) - P_C(y)\rangle &\leq 0. \end{aligned}$$

Then, $\langle x - y \mid P_C(x) - P_C(y)\rangle \geq \|P_C(x) - P_C(y)\|^2$.

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.

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

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

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

- 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.

适定性问题,

Feasibility problem

Problem

Let $\mathcal{H}$ be a Hilbert space. Let $m \in \mathbb{N} \setminus \{0, 1\}$.

Let $(C_i)_{1 \le i \le m}$ be closed convex subsets of $\mathcal{H}$ such that $\bigcap_{i=1}^{m} C_i \neq \varnothing$.

minizer.

We want to

$$\text{Find } \widehat{x} \in \bigcap_{i=1}^{m} C_i.$$

POCS (Projection Onto Convex Sets) algorithm

For every $n \in \mathbb{N}$, let $i_n - 1$ denote the remainder after division of n by m.

```
Set $x_0 \in \mathcal{H}$
For $n = 0, \ldots$
    $x_{n+1} = P_{C_{i_n}}(x_n)$.
```

Feasibility problem

Problem

Let $\mathcal{H}$ be a Hilbert space. Let $m \in \mathbb{N} \setminus \{0, 1\}$.

Let $(C_i)_{1 \leq i \leq m}$ be closed convex subsets of $\mathcal{H}$ such that $\bigcap_{i=1}^{m} C_i \neq \varnothing$.

We want to

$$\text{Find } \widehat{x} \in \bigcap_{i=1}^{m} C_i.$$

POCS (Projection Onto Convex Sets) algorithm

Let $(\lambda_n)_{n \geq 0}$ be a sequence of $[\epsilon_1, 2 - \epsilon_2]$ with $(\epsilon_1, \epsilon_2) \in \,]0, +\infty[^2$ such that $\epsilon_1 + \epsilon_2 < 2$.
For every $n \in \mathbb{N}$, let $i_n - 1$ denote the remainder after division of n by m.

```
Set $x_0 \in \mathcal{H}$
For $n = 0, \dots$
    $x_{n+1} = x_n + \lambda_n (P_{C_{i_n}}(x_n) - x_n).$
```

Convergence of POCS

Theorem

The sequence $(x_n)_{n\in\mathbb{N}}$ generated by the POCS algorithm converges weakly to a point in $\bigcap_{i=1}^{m} C_i$.

Remark : The result remains valid if the cyclic rule for activating the projections is replaced by a quasi-cyclic rule, i.e. the indice $(i_n)_{n\geq 0}$ are now such that

$$(\exists K \in \mathbb{N})(\forall n \in \mathbb{N}) \qquad \{1, \dots, m\} \subset \{i_n, \dots, i_{n+K}\}.$$

Exercice 3

Let $\mathcal{H}$ be a Hilbert space. Let $f : \mathcal{H} \to \mathbb{R}$ be Gâteaux differentiable and concave. Let $C \subset \mathcal{H}$ and let $\widehat{x} \in \mathrm{int}\,(C)$.

Show that f admits a global minimizer on C at $\widehat{x}$ if and only if f is constant over C.

Exercice 4

Let

$$\begin{aligned} f: \quad \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto \ln(\exp(x) + \exp(y)). \end{aligned}$$

Show that f is convex. Is it strictly convex ?

Exercise 5

1. What is the projection onto the positive orthant $[0,+\infty[^N$ in $\mathbb{R}^N$?
2. Let $\overline{x} \in \mathbb{R}^N$ and let $\rho \in \,]0,+\infty[$. What is the projection onto the closed ball
$$B(\overline{x},\rho) = \big\{x \in \mathbb{R}^N \;\big|\; \|x-\overline{x}\| \leq \rho\big\}?$$

Appendix

Fejér-monotonicity

Let D be a nonempty subset of a Hilbert space $\mathcal{H}$.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$.
$(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to D if

$$(\forall x \in D)(\forall n \in \mathbb{N}) \qquad \|x_{n+1} - x\| \leq \|x_n - x\|.$$

Let $D \subset \mathcal{H}$.
Let $(x_n)_{n\in\mathbb{N}}$ be Fejér-monotone with respect to D then

- for every $x \in D$, $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges,
- $(x_n)_{n\in\mathbb{N}}$ is bounded.

Fejér-monotonicity

Fejér-monotone convergence

Let D be a nonempty subset of a Hilbert space $\mathcal{H}$.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$.
$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in D if

- $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to D

  and
- every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ lies in D.

Fejér-monotonicity

Fejér-monotone convergence

Let D be a nonempty subset of a Hilbert space $\mathcal{H}$.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$.
$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in D if

- $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to D

  and

- every weak sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ lies in D.

Proof : We will only prove this result when $\mathcal{H}$ is finite dimensional.
Since $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to D, it is a bounded sequence. There thus exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $x_{n_k} \to \widehat{x}$ and, by assumption, $\widehat{x} \in D$.
On the other hand, there exists $\alpha \in \mathbb{R}$ such that $\|x_n - \widehat{x}\| \to \alpha$. Since $\|x_{n_k} - \widehat{x}\| \to 0$, $\alpha = 0$. Hence, $x_n \to \widehat{x}$.

Proof of convergence of POCS

Step 1 :

Let $D = \bigcap_{i=1}^{m} C_i$. For every $x \in D$ and for every $n \in \mathbb{N}$,

$$\begin{aligned}\|x_{n+1} - x\|^2 &= \|x_n - x\|^2 + 2\lambda_n \left\langle x_n - x \mid P_{C_{i_n}}(x_n) - x_n \right\rangle + \lambda_n^2 \|P_{C_{i_n}}(x_n) - x_n\|^2 \\ &= \|x_n - x\|^2 + 2\lambda_n \left\langle x_n - P_{C_{i_n}}(x_n) \mid P_{C_{i_n}}(x_n) - x_n \right\rangle \\ &\quad + 2\lambda_n \left\langle P_{C_{i_n}}(x_n) - x \mid P_{C_{i_n}}(x_n) - x_n \right\rangle + \lambda_n^2 \|P_{C_{i_n}}(x_n) - x_n\|^2\end{aligned}$$

Since $x \in D \subset C_{i_n}$, $\left\langle P_{C_{i_n}}(x_n) - x \mid P_{C_{i_n}}(x_n) - x_n \right\rangle \leq 0$. By using the fact that $\lambda_n \in [\epsilon_1, 2 - \epsilon_2]$, it can be deduced that

$$\begin{aligned}\|x_{n+1} - x\|^2 &\leq \|x_n - x\|^2 - (2 - \lambda_n)\lambda_n \|P_{C_{i_n}}(x_n) - x_n\|^2 \\ &\leq \|x_n - x\|^2 - \epsilon_1\epsilon_2 \|P_{C_{i_n}}(x_n) - x_n\|^2 \leq \|x_n - x\|^2.\end{aligned}$$

This shows that $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to D.

Proof of convergence of POCS

Step 2 :

Since $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to D, for every $x \in D$, $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges . In addition, since we have proved that

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\|^2 \leq \|x_n - x\|^2 - \epsilon_1\epsilon_2\|P_{C_{i_n}}(x_n) - x_n\|^2,$$

$P_{C_{i_n}}(x_n) - x_n \to 0.$

By using now the fact that

$$x_{n+1} - x_n = \lambda_n(P_{C_{i_n}}(x_n) - x_n)$$

where $(\lambda_n)_{n\in\mathbb{N}}$ is bounded, we deduce that $x_{n+1} - x_n \to 0$.

Proof of convergence of POCS

Step 3 :

Let $j \in \{1, \ldots, m\}$ and let m_n be the smallest integer greater than or equal to n such that $i_{m_n} = j$. We have then $n \leq m_n < n + m$. In addition, since P_{C_j} is nonexpansive,

$$\begin{aligned} \|P_{C_j}(x_n) - x_n\| &\leq \|P_{C_j}(x_n) - P_{C_{i_{m_n}}}(x_{m_n}) + P_{C_{i_{m_n}}}(x_{m_n}) - x_{m_n} + x_{m_n} - x_n\| \\ &\leq \|P_{C_j}(x_n) - P_{C_j}(x_{m_n})\| + \|P_{C_{i_{m_n}}}(x_{m_n}) - x_{m_n}\| + \|x_{m_n} - x_n\| \\ &\leq \|P_{C_{i_{m_n}}}(x_{m_n}) - x_{m_n}\| + 2\|x_{m_n} - x_n\|. \end{aligned}$$

As $n \to +\infty$, $m_n \to +\infty$ and $P_{C_{i_{m_n}}}(x_{m_n}) - x_{m_n} \to 0$. In addition, for every $x \in D$,

$$\|x_{m_n} - x_n\| \leq \sum_{i=n}^{m_n - 1} \|x_{i+1} - x_i\|.$$

As there are at most $m - 1$ terms in the summation and we have proved that $x_{n+1} - x_n \to 0$, it follows that $\|x_{m_n} - x_n\| \to 0$ and $P_{C_j}(x_n) - x_n \to 0$.

Proof of convergence of POCS

Step 4 : (assuming $\mathcal{H}$ finite dimensional)
Let x be a cluster point of $(x_n)_{n\in\mathbb{N}}$. There exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $x_{n_k} \to x$.
For every $j \in \{1,\ldots,m\}$, since P_{C_j} is continuous and $P_{C_j}(x_{n_k}) - x_{n_k} \to 0$, we have $P_{C_j}(x) = x$. This shows that $x \in D$.
It follows from the Fejér-monotone convergence theorem that $x_n \to x$.

Projections onto convex sets

仿射空间

In mathematics, **projections onto convex sets** (**POCS**), sometimes known as the **alternating projection** method, is a method to find a point in the intersection of two closed convex sets. It is a very simple algorithm and has been rediscovered many times.[1] The simplest case, when the sets are affine spaces, was analyzed by John von Neumann.[2][3] The case when the sets are affine spaces is special, since the iterates not only converge to a point in the intersection (assuming the intersection is non-empty) but to the orthogonal projection of the point onto the intersection. For general closed convex sets, the limit point need not be the projection. Classical work on the case of two closed convex sets shows that the rate of convergence of the iterates is linear.[4][5] There are now extensions that consider cases when there are more than one set, or when the sets are not convex,[6] or that give faster convergence rates. Analysis of POCS and related methods attempt to show that the algorithm converges (and if so, find the rate of convergence), and whether it converges to the projection of the original point. These questions are largely known for simple cases, but a topic of active research for the extensions. There are also variants of the algorithm, such as Dykstra's projection algorithm. See the references in the further reading section for an overview of the variants, extensions and applications of the POCS method; a good historical background can be found in section III of.[7]

Contents

Algorithm

The POCS algorithm solves the following problem:

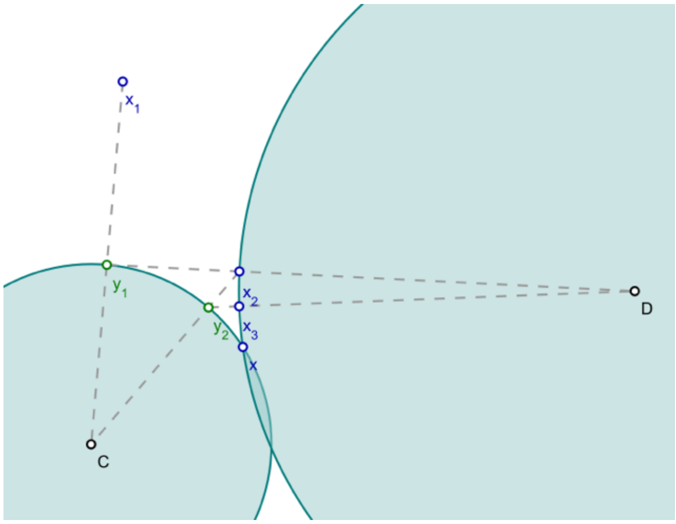
$$\text{find } x \in \mathbb{R}^n \quad \text{such that } x \in C \cap D$$

where C and D are closed convex sets.

To use the POCS algorithm, one must know how to project onto the sets C and D separately. The algorithm starts with an arbitrary value for x_0 and then generates the sequence

$$x_{k+1} = \mathcal{P}_C\left(\mathcal{P}_D(x_k)\right).$$

The simplicity of the algorithm explains some of its popularity. If the intersection of C and D is non-empty, then the sequence generated by the algorithm will converge to some point in this intersection.



Example on two circles

Unlike Dykstra's projection algorithm, the solution need not be a projection onto the intersection C and D.

Related algorithms

The method of **averaged projections** is quite similar. For the case of two closed convex sets C and D, it proceeds by

$$x_{k+1} = \frac{1}{2}(\mathcal{P}_C(x_k) + \mathcal{P}_D(x_k))$$

It has long been known to converge globally.[8] Furthermore, the method is easy to generalize to more than two sets; some convergence results for this case are in.[9]

The *averaged* projections method can be reformulated as *alternating* projections method using a standard trick. Consider the set

Example of **averaged projections** variant

$$E = \{(x,y) : x \in C,\ y \in D\}$$

which is defined in the product space $\mathbb{R}^n \times \mathbb{R}^n$. Then define another set, also in the product space:

$$F = \{(x,y) : x \in \mathbb{R}^n,\, y \in \mathbb{R}^n,\ x = y\}.$$

Thus finding $C \cap D$ is equivalent to finding $E \cap F$.

To find a point in $E \cap F$, use the alternating projection method. The projection of a vector (x, y) onto the set F is given by $(x+y, x+y)/2$. Hence

$$(x_{k+1}, y_{k+1}) = \mathcal{P}_F(\mathcal{P}_E((x_k, y_k))) = \mathcal{P}_F((\mathcal{P}_C x_k, \mathcal{P}_D y_k)) = \frac{1}{2}(\mathcal{P}_C(x_k) + \mathcal{P}_D(y_k), (\mathcal{P}_C(x_k) + \mathcal{P}_D(y_k)).$$

Since $x_{k+1} = y_{k+1}$ and assuming $x_0 = y_0$, then $x_j = y_j$ for all $j \geq 0$, and hence we can simplify the iteration to $x_{k+1} = \frac{1}{2}(\mathcal{P}_C(x_k) + \mathcal{P}_D(x_k))$.

References

1. Bauschke, H.H.; Borwein, J.M. (1996). "On projection algorithms for solving convex feasibility problems". *SIAM Review*. **38** (3): 367–426. CiteSeerX 10.1.1.49.4940 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.49.4940). doi:10.1137/S0036144593251710 (https://doi.org/10.1137%2FS0036144593251710).
2. J. von Neumann,Neumann, John Von (1949). "On rings of operators. Reduction theory". *Ann. of Math*. **50** (2): 401–485. doi:10.2307/1969463 (https://doi.org/10.2307%2F1969463). JSTOR 1969463 (https://www.jstor.org/stable/1969463). (a reprint of lecture notes first distributed in 1933)
3. J. von Neumann. Functional Operators, volume II. Princeton University Press, Princeton, NJ, 1950. Reprint of mimeographed lecture notes first distributed in 1933.
4. Gubin, L.G.; Polyak, B.T.; Raik, E.V. (1967). "The method of projections for finding the common point of convex sets". *U.S.S.R. Computational Mathematics and Mathematical Physics*. **7** (6): 1–24. doi:10.1016/0041-5553(67)90113-9 (https://doi.org/10.1016%2F0041-5553%2867%2990113-9).
5. Bauschke, H.H.; Borwein, J.M. (1993). "On the convergence of von Neumann's alternating projection algorithm for two sets". *Set-Valued Analysis*. **1** (2): 185–212. doi:10.1007/bf01027691 (https://doi.org/10.1007%2Fbf01027691).
6. Lewis, A. S.; Malick, J. (2008). "Alternating Projections on Manifolds". *Mathematics of Operations Research*. **33**: 216–234. CiteSeerX 10.1.1.416.6182 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.416.6182). doi:10.1287/moor.1070.0291 (https://doi.org/10.1287%2Fmoor.1070.0291).

7. Combettes, P. L. (1993). "The foundations of set theoretic estimation" (https://web.archive.org/web/20150614002927/http://www.ann.jussieu.fr/~plc/proc.pdf) (PDF). *Proceedings of the IEEE*. **81** (2): 182–208. doi:10.1109/5.214546 (https://doi.org/10.1109%2F5.214546). Archived from the original (http://www.ann.jussieu.fr/~plc/proc.pdf) (PDF) on 2015-06-14. Retrieved 2012-10-09.
8. A. Auslender. Methodes Numeriques pour la Resolution des Problems d'Optimisation avec Constraintes. PhD thesis, Faculte des Sciences, Grenoble, 1969
9. Local convergence for alternating and averaged nonconvex projections. A Lewis, R Luke, J Malick, 2007. arXiv (https://arxiv.org/abs/0709.0109)

Further reading

- Book from 2011: Alternating Projection Methods (https://dl.acm.org/citation.cfm?id=2077655) by René Escalante and Marcos Raydan (2011), published by SIAM.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Projections_onto_convex_sets&oldid=994018139"

This page was last edited on 13 December 2020, at 18:42 (UTC).