Monotone operator splitting methods and proximal algorithms. Applications in image recovery.

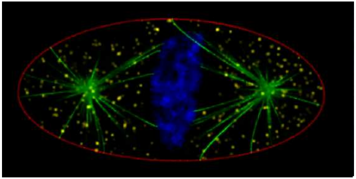
N. Pustelnik[1]
in collaboration with J.-C. Pesquet[2]

[1] ENS Lyon – Laboratoire de Physique – CNRS UMR 5672
[2] Univ. Paris-Est – LIGM – CNRS UMR 8049

Summer school BigOptim 2015
June, 29-30 2015

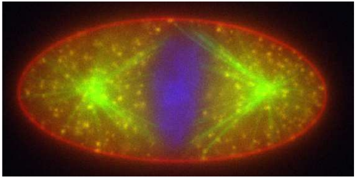
Monotone operators and inverse problems

[Microscopy, ISBI Challenge 2013, F. Soulez]

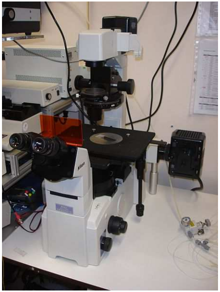


Original image

?

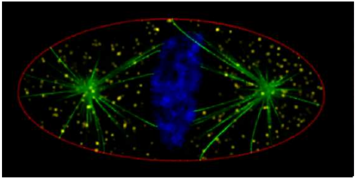


Degraded image

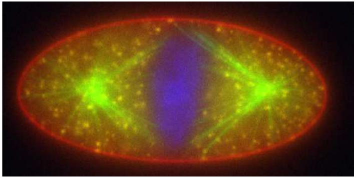


Monotone operators and inverse problems

[Microscopy, ISBI Challenge 2013, F. Soulez]



?



Original image

$\overline{x} \in \mathbb{R}^N$

Degraded image

$z = \mathcal{P}_\alpha(H\overline{x}) \in \mathbb{R}^M$

- $H \in \mathbb{R}^{M \times N}$: matrix associated with the degradation operator.
- $\mathcal{P}_\alpha \colon \mathbb{R}^M \to \mathbb{R}^M$: noise degradation with parameter α (e.g. Poisson noise).

$\to$ Find a good estimate of $\overline{x}$ from the observations z, using some a priori knowledge on H and on the noise statistics.

Monotone operators and inverse problems

Inverse problem:

Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{P}_\alpha(H\overline{x})$.

- Inverse filtering (if $M = N$ and H is invertible)

$$\begin{aligned}\widehat{x} &= H^{-1}z \\ &= H^{-1}(H\overline{x} + b) \quad \leftarrow \quad \text{if } b \in \mathbb{R}^M \text{ is an additive noise} \\ &= \overline{x} + H^{-1}b\end{aligned}$$

$\rightarrow$ Closed form expression, but amplification of the noise if H is ill-conditioned (*ill-posed problem*).

Monotone operators and inverse problems

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{P}_\alpha(H\overline{x})$.

- Inverse filtering (if $M \geq N$ and the rank of H is N)

$$\begin{aligned}\widehat{x} &= (H^\top H)^{-1} H^\top z \\ &= (H^\top H)^{-1} H^\top (H\overline{x} + b) \quad \leftarrow \quad \text{if } b \in \mathbb{R}^M \text{ is an additive noise} \\ &= \overline{x} + (H^\top H)^{-1} H^\top b\end{aligned}$$

$\rightarrow$ Closed form expression, but amplification of the noise if H is ill-conditioned (*ill-posed problem*).

Monotone operators and inverse problems

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{P}_\alpha(H\overline{x})$.

- Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \; \underbrace{f_1(x)}_{\substack{\text{Data fidelity term} \\ \text{e.g.} \|Hx - z\|_2^2}} + \underbrace{f_2(x)}_{\substack{\text{Regularization term} \\ \text{e.g. } \lambda\|x\|_p^p \text{ with } \begin{cases} p \geq 1 \\ \lambda \in \,]0,+\infty[\end{cases}}}$$

$\rightarrow$ Often no closed form expression (e.g. if $p \neq 2$ and $H \neq \mathrm{Id}$)
or solution expensive to compute (e.g. if $p = 2$, $H \neq \mathrm{Id}$ and $N \gg 1$)
$\rightarrow$ Iterative strategy.

Monotone operators and inverse problems

Inverse problem:
Find an estimate $\widehat{x}$ close to $\overline{x}$ from the observations $z = \mathcal{P}_\alpha(H\overline{x})$.

- Variational approach (more general context)

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathbb{R}^N} \sum_{i=1}^{m} f_i(x)$$

where f_i may denote a data fidelity term / a (hybrid) regularization term / constraint.

$\rightarrow$ Iterative strategy.

Monotone operators and inverse problems

Iterative strategy = Optimization algorithm:

Construct a sequence $(x_n)_{n\in\mathbb{N}}$ that converges to $\widehat{x} \in \underset{x\in\mathbb{R}^N}{\mathrm{Argmin}} \sum_{i=1}^{m} f_i(x)$.

- Sequence such that $(\forall n \in \mathbb{N})\ x_{n+1} = Tx_n$ where T denotes an operator from $\mathbb{R}^N$ to $\mathbb{R}^N$.

  $\rightarrow$ How can we build T from the functionals $(f_i)_{1\leq i\leq m}$ involved in the minimization problem ?

  $\rightarrow$ Which properties are required by T in order to ensure the convergence of $(x_n)_{n\in\mathbb{N}}$ to $\widehat{x}$?

Naive answer

Fixed point theorem (E. Picard, 1856-1941)

If

- $\widehat{x}$ is a fixed point of T, i.e. $\widehat{x} = T\widehat{x}$
- T is a strict contraction, i.e. there exists $\rho \in [0,1[$ such that

$$\big(\forall (x,x') \in \mathbb{R}^N \times \mathbb{R}^N\big) \qquad \|Tx - Tx'\| \leq \rho \|x - x'\|$$

then $(x_n)_{n\in\mathbb{N}}$ converges to $\widehat{x}$.

Proof: For all $n \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+1} - \widehat{x}\| &= \|Tx_n - T\widehat{x}\| \\ &\leq \rho \|x_n - \widehat{x}\|. \end{aligned}$$

Consequently, $\|x_n - \widehat{x}\| \leq \rho^n \|x_0 - \widehat{x}\|$. Hence, we have proved that $(x_n)_{n\in\mathbb{N}}$ converges linearly to $\widehat{x}$.

Why do we need to go further ?

Limitations:

- It is difficult (even sometimes impossible) to build a *strictly* contractive operator T.
- One may prefer iterations built as $(\forall n \in \mathbb{N})\ x_{n+1} = T_n x_n$ where T_n denotes an operator from $\mathbb{R}^N$ to $\mathbb{R}^N$.
- It is often intricate to build T_n, while it may be easier to write T_n as a composition of simpler operators (*splitting techniques*).
- T_n can be multivalued, i.e. $(\forall n \in \mathbb{N})\ x_{n+1} \in T_n x_n$.

Tutorial philosophy

- Provide a modern vision of convex optimization in order to deal with nonsmooth problems (sparsity)
  $\rightarrow$ possibly non-finite functions, monotone operators,...

- Provide a powerful framework to capture many convex optimization algorithms (forward-backward, Douglas-Rachford, ADMM,...) in a unifying form.

- Introduce the technical literature on this topic
  $\rightarrow$ deal with infinite dimensional Hilbert spaces ... even if most of the signal/image processing applications are in finite dimension.

Tutorial philosophy

- Illustrate the performance of the algorithms on inverse problems examples (without giving too many details).
- Do not explore all the applications of monotone operators.

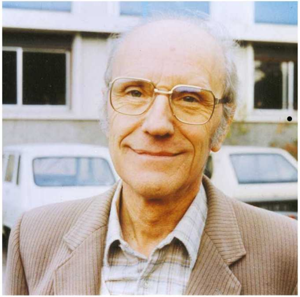
Tutorial philosophy

- Illustrate the performance of the algorithms on inverse problems examples (without giving too many details).
- Do not explore all the applications of monotone operators. (Denoising, restoration, reconstruction, machine learning, ressource allocation, networking, communications,...)

Tutorial philosophy

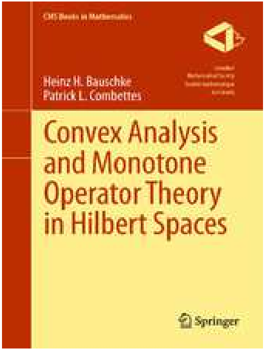
- Illustrate the performance of the algorithms on inverse problems examples (without giving too many details).
- Do not explore all the applications of monotone operators. (Denoising, restoration, reconstruction, machine learning, ressource allocation, networking, communications,...)
- Focus on the convergence of the iterates $(x_n)_{n\in\mathbb{N}}$ rather than on the convergence of criterion $\left(\sum_{i=1}^m f_i(x_n)\right)_{n\in\mathbb{N}}$.

A pioneer



Jean-Jacques Moreau
(1923–2014)

Reference book



– H.H. Bauschke and P.L. Combettes –

Outline

1. Background on monotone and maximally monotone operators

   → *Inversion, subdifferential, conjugate of a convex function*

2. Nonexpansive operators

   → *Taxinomy, resolvent, and proximity operator*

3. Search for a zero of a maximally monotone operator

   → *Fixed points, Fejér monotonicity, Douglas-Rachford, Forward-Backward*

4. Duality

   → *Main theorems, ADMM, primal-dual methods*

Part 1: Background

1. Monotone operators
   - Definition
   - Properties
   - Basic operations
   - Inversion
   - Maximality
   - Usefulness for convex optimization (subdifferential)

2. Maximally monotone operators
   - Properties
   - Basic operations
   - Inversion
   - Usefulness of inversion for convex optimization

Hilbert spaces

A (real) Hilbert space $\mathcal{H}$ is a complete real vector space endowed with an inner product $\langle \cdot \mid \cdot \rangle$ whose associated norm is

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

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

Hilbert spaces

A (real) **Hilbert space $\mathcal{H}$** is a complete real vector space endowed with an inner product $\langle \cdot \mid \cdot \rangle$ whose associated norm is

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

- 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** (or continuous) 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** (or continuous) 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** (or continuous) if

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

- In finite dimension, every linear operator is bounded.

Hilbert spaces

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

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

- In finite dimension, every linear operator is bounded.

$\mathcal{B}(\mathcal{H},\mathcal{G})$: 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:

If $\qquad L \colon \mathcal{H} \to \mathcal{H}^n \colon y \mapsto (y, \dots, y)$

then $\qquad L^* \colon \mathcal{H}^n \to \mathcal{H} \colon x = (x_1, \dots, x_n) \mapsto \sum_{i=1}^{n} x_i$

Proof:

$$\langle Ly \mid x \rangle = \langle (y, \dots, y) \mid (x_1, \dots, x_n) \rangle = \sum_{i=1}^{n} \langle y \mid x_i \rangle = \left\langle y \;\middle|\; \sum_{i=1}^{n} x_i \right\rangle$$

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 $L^* = L^\top$.

Hilbert spaces

Let $\mathcal{H}$ be a Hilbert space and $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$.

- L is self-adjoint if $L^* = L$.
- L is positive if $(\forall x \in \mathcal{H})\ \langle x \mid Lx \rangle \geq 0$.
- L is strictly positive if L is positive and if $(\forall x \in \mathcal{H})\ \langle x \mid Lx \rangle = 0 \Leftrightarrow x = 0$.

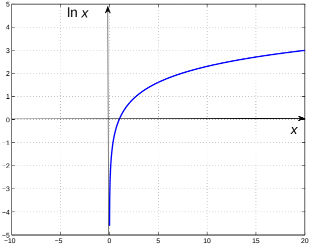
Mappings versus multivalued operators

Let $\mathcal{H}$ be a real Hilbert space.
A is an $\mathcal{H}$-valued mapping defined on $D \subset \mathcal{H}$ if

$$A\colon D \to \mathcal{H}$$
$$x \mapsto A(x)$$

- Example:

$$A\colon \ ]0,+\infty[\to \mathbb{R}$$
$$x \mapsto \ln x$$



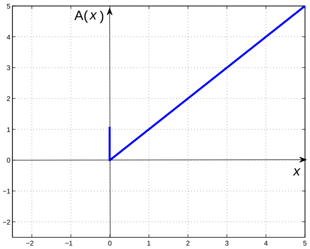
Mappings versus multivalued operators

Let $\mathcal{H}$ be a real Hilbert space.
A is a (multivalued) operator if

$$A\colon \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \mapsto \left\{A_i(x) \;\middle|\; i \in I_x \subset \mathbb{R}\right\}$$

- Example:

$$A\colon \mathbb{R} \to 2^{\mathbb{R}}$$
$$x \mapsto \begin{cases} \{x\} & \text{if } x > 0 \\ [0,1] & \text{if } x = 0 \\ \varnothing & \text{if } x < 0 \end{cases}$$

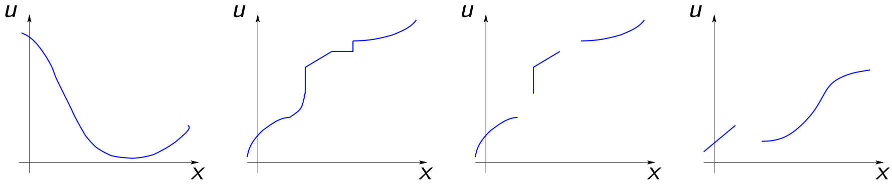


Graph

Let $\mathcal{H}$ be a real Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The graph of A is

$$\mathrm{gra} A = \{(x, u) \in \mathcal{H}^2 \mid u \in Ax\}.$$

- Graph examples:



Graph

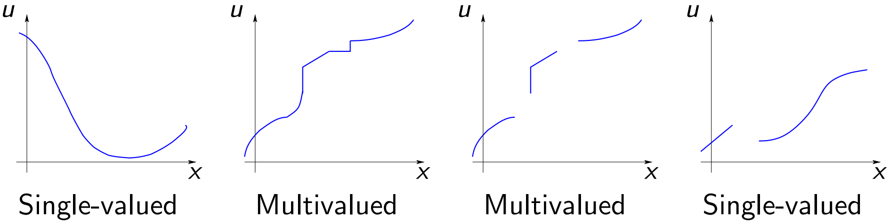
Let $\mathcal{H}$ be a real Hilbert space.

Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.

The graph of A is

$$\mathrm{gra} A = \{(x, u) \in \mathcal{H}^2 \mid u \in Ax\}.$$

- Graph examples:

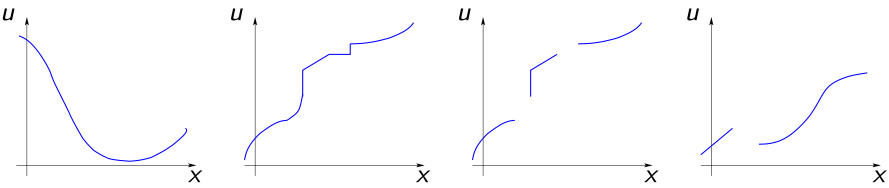


Monotone operator: definition

Let $\mathcal{H}$ be a real Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is **monotone** if

$$(\forall (x_1, u_1) \in \operatorname{gra} A)\,(\forall (x_2, u_2) \in \operatorname{gra} A) \qquad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq 0.$$

- Monotone operators ?

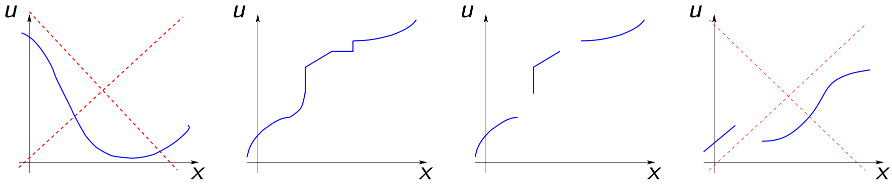


Monotone operator: definition

Let $\mathcal{H}$ be a real Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is **monotone** if

$$\big(\forall (x_1,u_1) \in \operatorname{gra} A\big)\big(\forall (x_2,u_2) \in \operatorname{gra} A\big) \qquad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq 0.$$

- Monotone operators ?



Monotone operator: example

Let $\mathcal{H}$ be a real Hilbert space.
Let $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$.

- A is monotone $\Leftrightarrow$ A is positive
- A monotone $\Leftrightarrow$ $A + A^*$ monotone $\Leftrightarrow$ A^* monotone.

Monotone operator: example

Let $\mathcal{H}$ be a real Hilbert space.
Let $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$.

- A is monotone $\Leftrightarrow$ A is positive
- A monotone $\Leftrightarrow$ $A + A^*$ monotone $\Leftrightarrow$ A^* monotone.

Proof:

$$\begin{aligned}
A \text{ monotone} \quad &\Leftrightarrow \quad (\forall (x_1, x_2) \in \mathcal{H}^2) \qquad \langle x_1 - x_2 \mid Ax_1 - Ax_2 \rangle \geq 0 \\
&\Leftrightarrow \quad (\forall x \in \mathcal{H}) \qquad 2 \langle x \mid Ax \rangle \geq 0 \\
&\Leftrightarrow \quad (\forall x \in \mathcal{H}) \qquad \langle x \mid Ax \rangle + \langle A^* x \mid x \rangle \geq 0 \\
&\Leftrightarrow \quad (\forall x \in \mathcal{H}) \qquad \langle x \mid (A + A^*) x \rangle \geq 0 \\
&\Leftrightarrow \quad A + A^* \text{ monotone}
\end{aligned}$$

Monotone operator: example

Let $\mathcal{H}$ be a real Hilbert space.
Let $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$.

- A is monotone $\Leftrightarrow$ A is positive
- A monotone $\Leftrightarrow$ $A + A^*$ monotone $\Leftrightarrow$ A^* monotone.

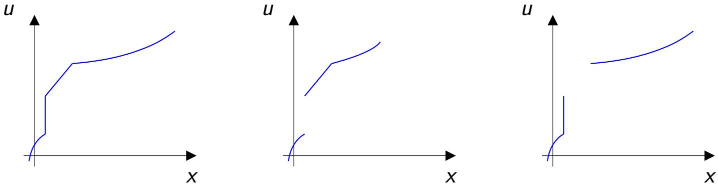
- For $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ to be monotone, A is not required to be self-adjoint.
  Example : $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ skewed (i.e. $A^* = -A$) is monotone.

Domain

Let $\mathcal{H}$ be a real Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The **domain** of A is

$$\operatorname{dom} A = \{x \in \mathcal{H} \mid Ax \neq \varnothing\}.$$

- Which domain ?

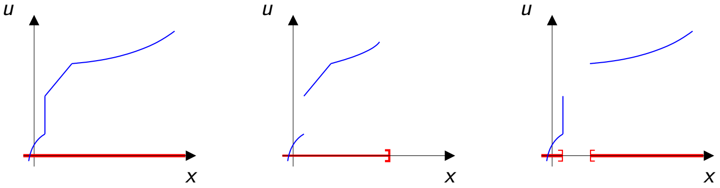


Domain

Let $\mathcal{H}$ be a real Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The domain of A is

$$\operatorname{dom} A = \left\{ x \in \mathcal{H} \;\middle|\; Ax \neq \varnothing \right\}.$$

- Which domain ?

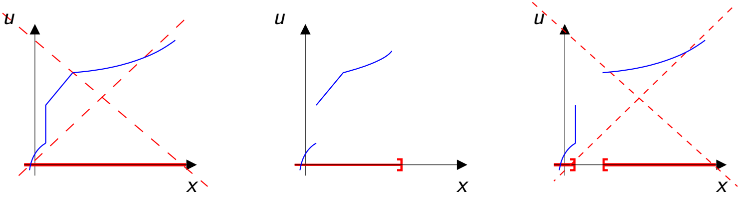


Domain

Let $\mathcal{H}$ be a real Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The domain of A is

$$\operatorname{dom} A = \{x \in \mathcal{H} \mid Ax \neq \varnothing\}.$$

- Let $C \subset \mathcal{H}$. If $\operatorname{dom} A = C$ and for every $x \in C$, Ax is a singleton, we view A as a mapping from C to $\mathcal{H}$.



Monotone operator: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two monotone operators.
The following operators are monotone:

- $x \mapsto y + \gamma\rho A(\rho x + z) = \{y + \gamma\rho u \mid u \in A(\rho x + z)\}$
  where $(y, z) \in \mathcal{H}^2$, $\gamma \in [0, +\infty[$ and $\rho \in \mathbb{R}$.
- $A \times B : \mathcal{H} \times \mathcal{G} \to 2^{\mathcal{H}\times\mathcal{G}}$
  $(x, y) \mapsto Ax \times Ay = \{(u, v) \mid u \in Ax, v \in Bx\}$.
- $A + B : x \mapsto \{u + v \mid u \in Ax, v \in Bx\}$ if $\mathcal{G} = \mathcal{H}$.
- $L^*BL : x \mapsto \{L^*v \mid v \in B(Lx)\}$ if $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Monotone operator: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two monotone operators.
The following operators are monotone:

- $x \mapsto y + \gamma\rho A(\rho x + z)$ where $(y, z) \in \mathcal{H}^2$, $\gamma \in [0, +\infty[$ and $\rho \in \mathbb{R}$.
- $A \times B$.
- $A + B$ if $\mathcal{G} = \mathcal{H}$.
- $L^* BL$ if $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Proof : Let $(x_1, u_1) \in \operatorname{gra}(L^* BL)$ and $(x_2, u_2) \in \operatorname{gra}(L^* BL)$.
We have $u_1 = L^* v_1$ and $u_2 = L^* v_2$ where $v_1 \in B(Lx_1)$ and $v_2 \in B(Lx_2)$.
Moreover,

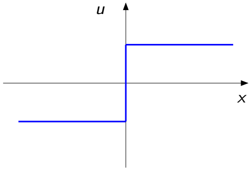
$$\begin{aligned}\langle u_1 - u_2 \mid x_1 - x_2\rangle &= \langle v_1 - v_2 \mid L(x_1 - x_2)\rangle \\ &= \langle v_1 - v_2 \mid Lx_1 - Lx_2\rangle \geq 0.\end{aligned}$$

Monotone operator: inversion

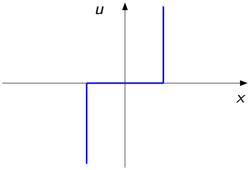
Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$.
A^{-1} is the operator from $\mathcal{H}$ to $2^{\mathcal{H}}$ the graph of which is

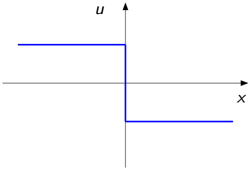
$$\operatorname{gra}(A^{-1}) = \big\{(u,x) \;\big|\; (x,u) \in \operatorname{gra} A\big\}.$$

Graph of A



Graph of A^{-1} ?



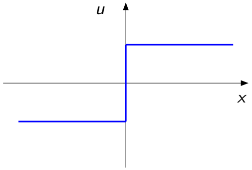


Monotone operator: inversion

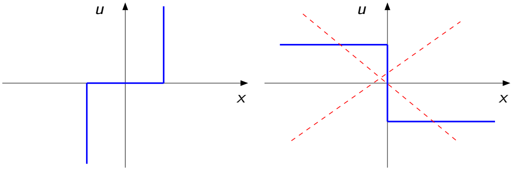
Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$.
A^{-1} is the operator from $\mathcal{H}$ to $2^{\mathcal{H}}$ the graph of which is

$$\mathrm{gra}(A^{-1}) = \big\{(u,x) \;\big|\; (x,u) \in \mathrm{gra}A\big\}.$$

Graph of A



Graph of A^{-1} ?



Monotone operator: inversion

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$.
A^{-1} is the operator from $\mathcal{H}$ to $2^{\mathcal{H}}$ the graph of which is

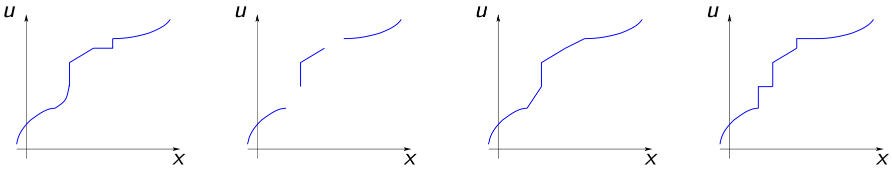
$$\mathrm{gra}(A^{-1}) = \big\{(u, x) \;\big|\; (x, u) \in \mathrm{gra} A\big\}.$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator.
A^{-1} is monotone.

Maximally monotone operator: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is maximally monotone if A is monotone and if there exists no monotone operator $B : \mathcal{H} \to 2^{\mathcal{H}}$ (different from A) such that $\mathrm{gra} B$ properly contains $\mathrm{gra} A$.

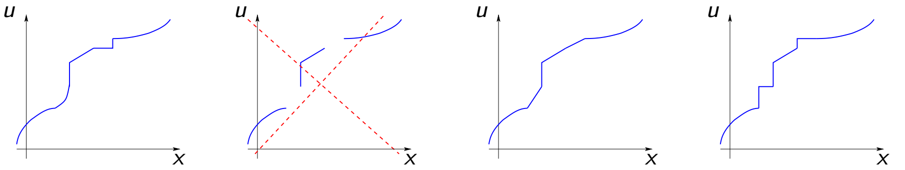
Maximally monotone operator ?



Maximally monotone operator: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is **maximally monotone** if A is monotone and if there exists no monotone operator $B \colon \mathcal{H} \to 2^{\mathcal{H}}$ (different from A) such that $\operatorname{gra} B$ properly contains $\operatorname{gra} A$.

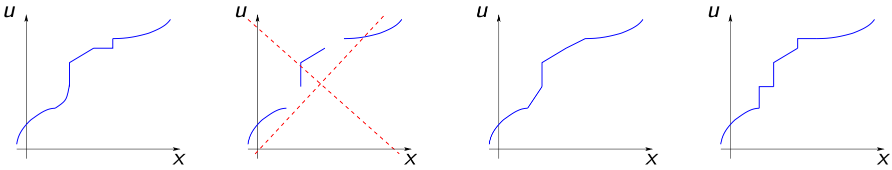
Maximally monotone operator ?



Maximally monotone operator: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is **maximally monotone** if A is monotone and if there exists no monotone operator $B : \mathcal{H} \to 2^{\mathcal{H}}$ (different from A) such that $\mathrm{gra} B$ properly contains $\mathrm{gra} A$.

Maximally monotone operator ?



Maximally monotone operator: second definition

Let $\mathcal{H}$ be a Hilbert space.
$A : \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone if one if the following equivalent conditions is satisfied:

(i) A is monotone and there exists no monotone operator $B \colon \mathcal{H} \to 2^{\mathcal{H}}$ such that $\mathrm{gra} B$ properly contains $\mathrm{gra} A$.

(ii) For every $(x_1, u_1) \in \mathcal{H}^2$,

$$(x_1, u_1) \in \mathrm{gra} A \quad \Leftrightarrow \quad \big(\forall (x_2, u_2) \in \mathrm{gra} A\big) \quad \langle x_1 - x_2 \mid u_1 - u_2 \rangle \geq 0.$$

Equivalence between (i) and (ii) :

$(ii) \Rightarrow (i)$: Condition (ii) insures A to be monotone.
Moreover, if B monotone and $\mathrm{gra} A \subset \mathrm{gra} B$ then $\big(\forall (x_1, u_1) \in \mathrm{gra} B\big)$

$$\big(\forall (x_2, u_2) \in \mathrm{gra} A\big) \quad \langle x_1 - x_2 \mid u_1 - u_2 \rangle \geq 0$$

Condition (ii) leads to $(x_1, u_1) \in \mathrm{gra} A$. Thus $B = A$.

Maximally monotone operator: second definition

Let $\mathcal{H}$ be a Hilbert space.
$A : \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone if one if the following equivalent conditions is satisfied:

- (*i*) A is monotone and there exists no monotone operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ such that $\mathrm{gra} B$ properly contains $\mathrm{gra} A$.
- (*ii*) For every $(x_1, u_1) \in \mathcal{H}^2$,

$$(x_1, u_1) \in \mathrm{gra} A \quad \Leftrightarrow \quad \big(\forall (x_2, u_2) \in \mathrm{gra} A\big) \quad \langle x_1 - x_2 \mid u_1 - u_2 \rangle \geq 0.$$

Equivalence between (i) and (ii) :

$(i) \Rightarrow (ii)$: Let $(x_1, u_1) \in \mathcal{H}^2$ such that the inequality is satisfied. Let B such that $\mathrm{gra} B = \mathrm{gra} A \cup \{(x_1, u_1)\}$. If A is monotone, B is monotone and $\mathrm{gra} A \subset \mathrm{gra} B$. From Condition (i), $B = A \Rightarrow (x_1, u_1) \in \mathrm{gra} A$.

Continuous functions

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to \mathcal{H}$ be monotone and continuous. Then A is maximally monotone.

Example :
If $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ is positive, then L is maximally monotone.

Maximally monotone operator

Maximally monotone operator

Maximally monotone operator

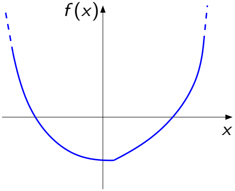
Usefulness in convex optimization

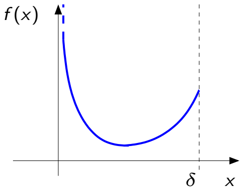
Convex analysis: definitions

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

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

Domains of the functions ?



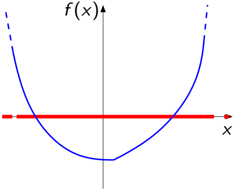


Convex analysis: definitions

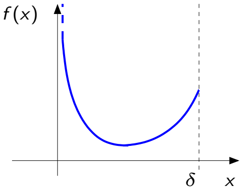
Let $f : \mathcal{H} \to]-\infty, +\infty]$ where $\mathcal{H}$ is a Hilbert space.

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

Domains of the functions ?



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

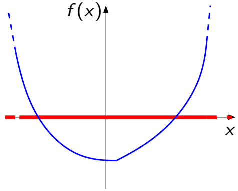


Convex analysis: definitions

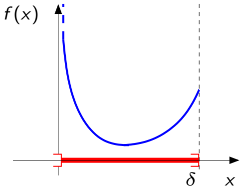
Let $f : \mathcal{H} \to]-\infty, +\infty]$ where $\mathcal{H}$ is a Hilbert space.

- The **domain** of f is $\mathrm{dom}\, f = \{x \in \mathcal{H} \,|\, f(x) < +\infty\}$.
- The function f is **proper** if $\mathrm{dom}\, f \neq \varnothing$.

Domains of the functions ?



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



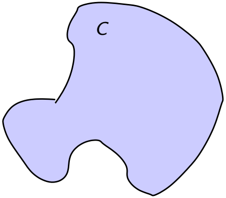
$\mathrm{dom}\, f =]0, \delta]$
(proper)

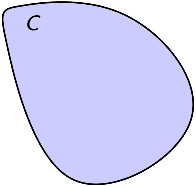
Convex analysis: definitions

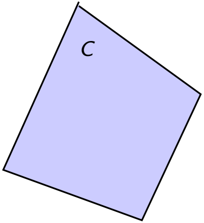
$C \subset \mathcal{H}$ is a convex set if

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

Convex sets ?





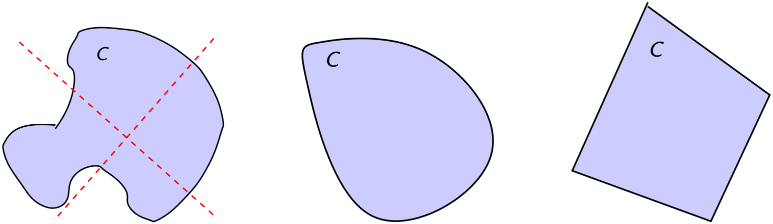


Convex analysis: definitions

$C \subset \mathcal{H}$ 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 analysis: definitions

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

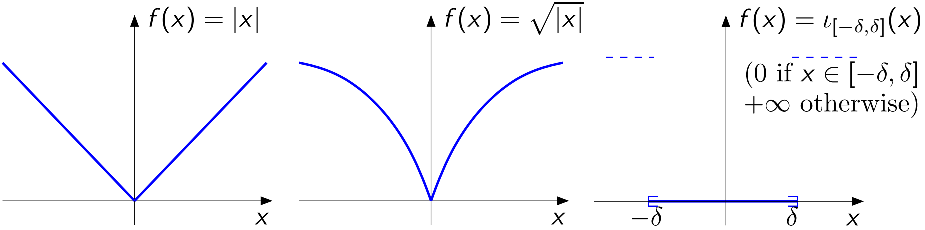
$$(\forall (x,y) \in \mathcal{H}^2)(\forall \alpha \in]0,1[)$$
$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

Convex analysis: definitions

$f : \mathcal{H} \to \,]-\infty, +\infty]$ is a **convex fonction** 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 functions ?

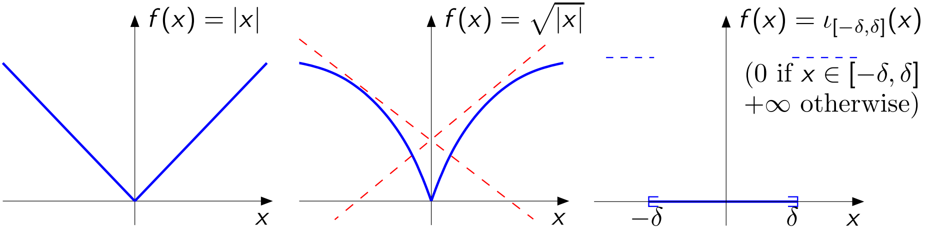


Convex analysis: definitions

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

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

Convex functions ?



Convex analysis: definitions

$f : \mathcal{H} \to]-\infty, +\infty]$ is convex $\Leftrightarrow$ the epigraph of f, i.e.

$$\operatorname{epi} f = \big\{(x,\zeta) \in \operatorname{dom} f \times \mathbb{R} \;\big|\; f(x) \le \zeta\big\}$$

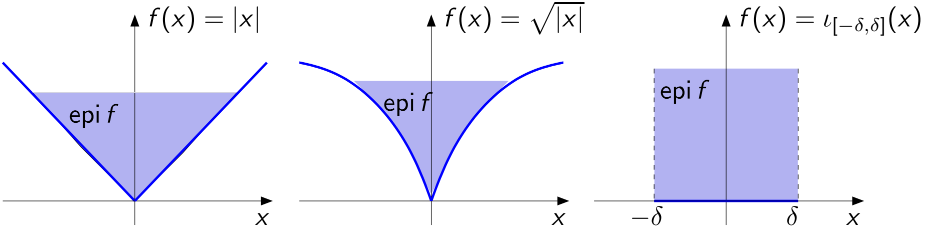
is convex.

Convex analysis: definitions

$f : \mathcal{H} \to]-\infty, +\infty]$ is convex $\Leftrightarrow$ the epigraph of f, i.e.

$$\operatorname{epi} f = \{(x,\zeta) \in \operatorname{dom} f \times \mathbb{R} \mid f(x) \leq \zeta\}$$

is convex.

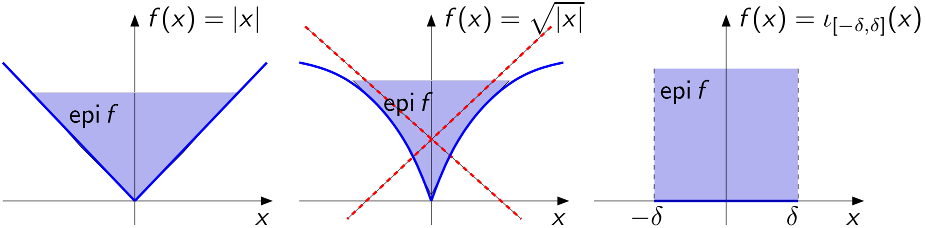


Convex analysis: definitions

$f : \mathcal{H} \to]-\infty, +\infty]$ is convex $\Leftrightarrow$ the epigraph of f, i.e.

$$\operatorname{epi} f = \{(x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \mid f(x) \leq \zeta\}$$

is convex.

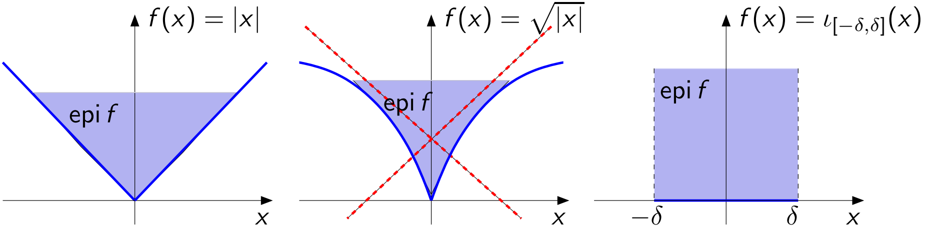


Convex analysis: definitions

$f : \mathcal{H} \to \,]-\infty, +\infty]$ is convex $\Leftrightarrow$ the **epigraph** of f, i.e.

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

is convex.

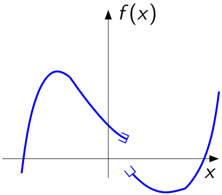


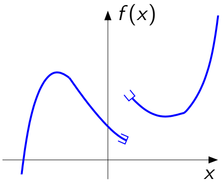
- $f : \mathcal{H} \to [-\infty, +\infty[$ is concave if $-f$ is convex.

Convex analysis: definitions

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous (l.s.c.) function on $\mathcal{H}$ if epi f is closed

- l.s.c. functions ?

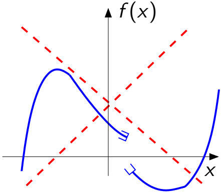


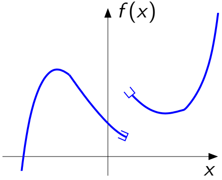


Convex analysis: definitions

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous (l.s.c.) function on $\mathcal{H}$ if epi f is closed

- l.s.c. functions ?

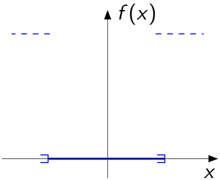


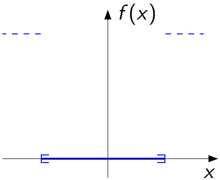


Convex analysis: definitions

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous (l.s.c.) function on $\mathcal{H}$ if epi f is closed

- l.s.c. functions ?

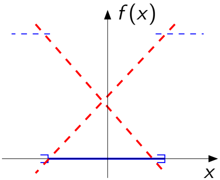


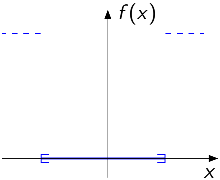


Convex analysis: definitions

Let $f : \mathcal{H} \to]-\infty, +\infty]$.
f is a lower semi-continuous (l.s.c.) function on $\mathcal{H}$ if epi f is closed

- l.s.c. functions ?

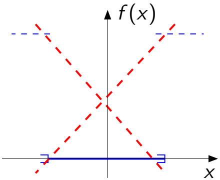


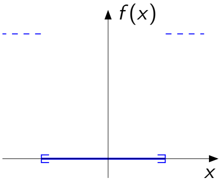


Convex analysis: definitions

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$.
f is a lower semi-continuous (l.s.c.) function on $\mathcal{H}$ if epi f is closed

- l.s.c. functions ?





Convex analysis: definitions/properties

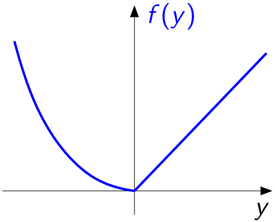
- $\Gamma_0(\mathcal{H})$: class of convex, l.s.c., and proper functions from $\mathcal{H}$ to $]-\infty, +\infty]$.
- Every continuous function on $\mathcal{H}$ is l.s.c.
- Every finite sum of l.s.c. (convex) functions is l.s.c. (convex).
- Let $(f_i)_{i\in I}$ be a family of l.s.c. (convex) functions. $\sup_{i\in I} f_i$ is l.s.c. (convex).

Subdifferential of a convex function: definition

The (Moreau) subdifferential of f, denoted by ∂f,

Subdifferential of a convex function: definition

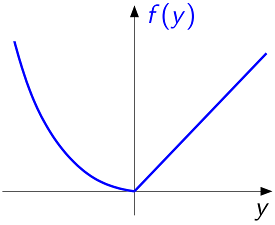
Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f,

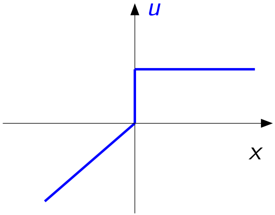


Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \le f(y)\} \end{aligned}$$

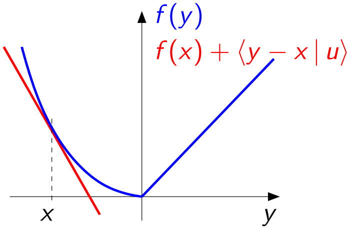


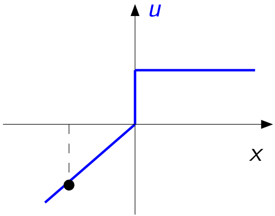


Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u\rangle + f(x) \le f(y)\} \end{aligned}$$

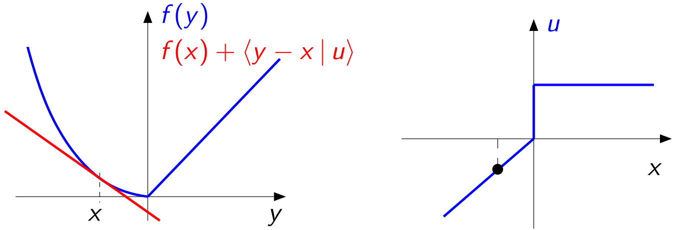




Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

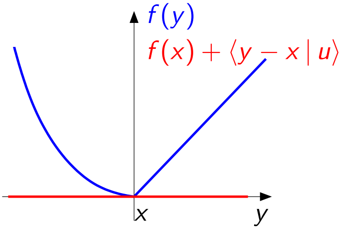
$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u\rangle + f(x) \le f(y)\} \end{aligned}$$

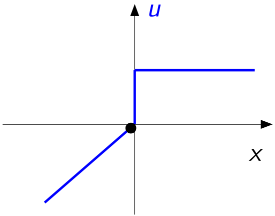


Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

$$\partial f : \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\}$$

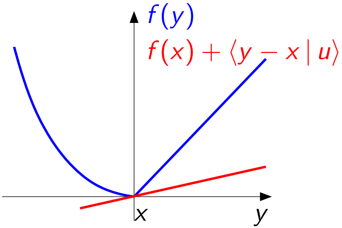


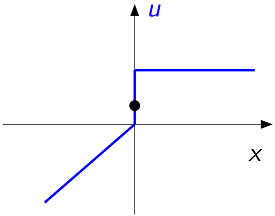


Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\} \end{aligned}$$

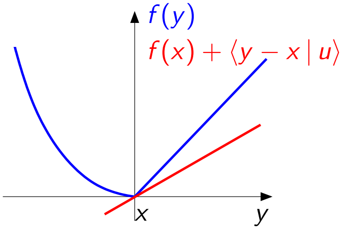


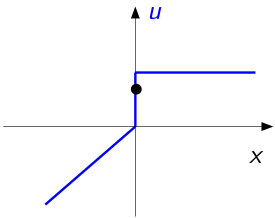


Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\} \end{aligned}$$

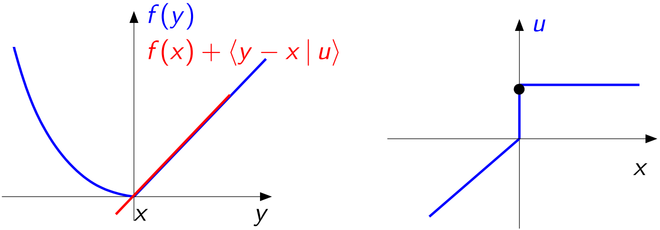




Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

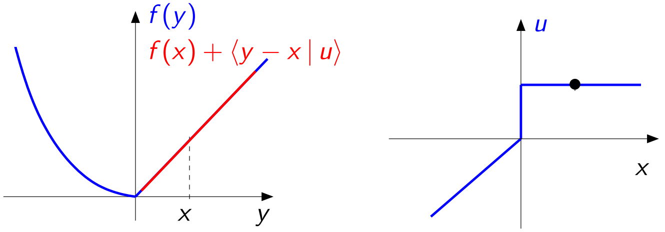
$$\partial f : \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\}$$



Subdifferential of a convex function: definition

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

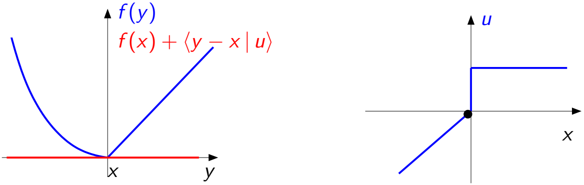
$$\partial f : \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\}$$



Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted by ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\} \end{aligned}$$



- Fermat rule:

$$0 \in \partial f(x) \Leftrightarrow (\forall y \in \mathcal{H})\ \langle y - x \,|\, 0 \rangle + f(x) \leq f(y)$$
$$\Leftrightarrow x \in \operatorname{Argmin} f$$

Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u\rangle + f(x) \leq f(y)\} \end{aligned}$$

- $u \in \partial f(x)$ is a subgradient of f at x.

Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\} \end{aligned}$$

- $u \in \partial f(x)$ is a subgradient of f at x.
- If $x \notin \operatorname{dom} f$, then $\partial f(x) = \varnothing$.

Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\} \end{aligned}$$

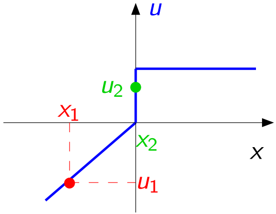
- $u \in \partial f(x)$ is a subgradient of f at x.
- If $x \notin \operatorname{dom} f$, then $\partial f(x) = \varnothing$.
- For every $x \in \operatorname{dom} f$, $\partial f(x)$ is a closed and convex set.

Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\} \end{aligned}$$

- ∂f is a monotone operator :

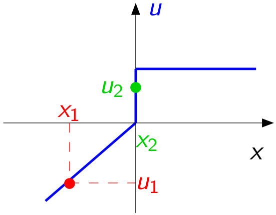


Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\begin{aligned} \partial f : \mathcal{H} &\to 2^{\mathcal{H}} \\ x &\to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u\rangle + f(x) \leq f(y)\} \end{aligned}$$

- ∂f is a monotone operator :
  Let $u_1 \in \partial f(x_1)$ and $u_2 \in \partial f(x_2)$.

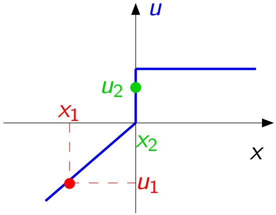


Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\partial f : \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\}$$

- ∂f is a monotone operator :
  Let $u_1 \in \partial f(x_1)$ and $u_2 \in \partial f(x_2)$.
  By using the subdifferential definition:
  $$\langle x_2 - x_1 | u_1 \rangle + f(x_1) \leq f(x_2)$$
  $$\langle x_1 - x_2 | u_2 \rangle + f(x_2) \leq f(x_1)$$



Subdifferential of a convex function: properties

Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
The (Moreau) subdifferential of f, denoted ∂f, is such that

$$\partial f : \mathcal{H} \to 2^{\mathcal{H}}$$
$$x \to \{u \in \mathcal{H} \,|\, (\forall y \in \mathcal{H})\ \langle y - x | u \rangle + f(x) \leq f(y)\}$$

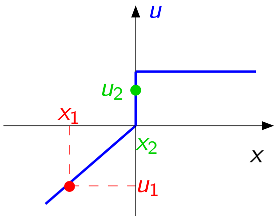
- ∂f is a monotone operator :

  Let $u_1 \in \partial f(x_1)$ and $u_2 \in \partial f(x_2)$.

  By using the subdifferential definition:

  $$\langle x_2 - x_1 | u_1 \rangle + f(x_1) \leq f(x_2)$$
  $$\langle x_1 - x_2 | u_2 \rangle + f(x_2) \leq f(x_1)$$

  and thus $\langle x_1 - x_2 \,|\, u_1 - u_2 \rangle \geq 0$.



Subdifferential of a convex function: properties

- The subdifferential of a convex and proper function is:
  - Monotone
  - If f is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$

Subdifferential of a convex function: properties

- The subdifferential of a convex and proper function is:
  - Monotone
  - If f is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$

$$(\forall y \in \mathcal{H}) \qquad \langle \nabla f(x) \mid y \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x+\alpha y) - f(x)}{\alpha}.$$

<u>Proof</u>:
For all $\alpha \in [0,1]$ and $y \in \mathcal{H}$,

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

$$\Rightarrow \quad \langle \nabla f(x) \mid y - x \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f\big(x+\alpha(y-x)\big) - f(x)}{\alpha} \leq f(y) - f(x)$$

Thus $\nabla f(x) \in \partial f(x)$.

Subdifferential of a convex function: properties

- The subdifferential of a convex and proper function is:
  - Monotone
  - If f is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$

$$(\forall y \in \mathcal{H}) \qquad \langle \nabla f(x) \mid y \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x+\alpha y) - f(x)}{\alpha}.$$

Proof:

Reciprocally, $u \in \partial f(x)$, then, for all $\alpha \in [0, +\infty[$ and $y \in \mathcal{H}$,

$$f(x+\alpha y) \geq f(x) + \langle u \mid x + \alpha y - x \rangle$$

$$\Rightarrow \quad \langle \nabla f(x) \mid y \rangle = \lim_{\substack{\alpha \to 0 \\ \alpha \neq 0}} \frac{f(x+\alpha y) - f(x)}{\alpha} \geq \langle u \mid y \rangle$$

Selecting $y = u - \nabla f(x)$, we deduce than $\|u - \nabla f(x)\|^2 \leq 0$. Thus $u = \nabla f(x)$.

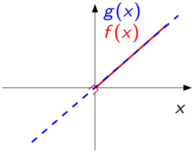
Subdifferential of a convex function: properties

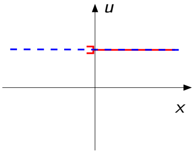
- The subdifferential of a convex and proper function is:
  - Monotone
  - If f is Gâteaux differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$
  - Non necessarily maximally monotone

Counterexample: For every $x \in \mathcal{H}$,

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ +\infty & \text{otherwise} \end{cases}, \qquad g(x) = x$$

$$\Rightarrow \quad \partial f(x) = \begin{cases} \{1\} & \text{if } x > 0 \\ \varnothing & \text{otherwise} \end{cases}, \qquad \partial g(x) = \{1\}.$$





Consequently, $\mathrm{gra}\partial f =]0, +\infty[\times \{1\} \subset \mathbb{R} \times \{1\}$
$\subset \mathrm{gra}\partial g$

Subdifferential of a convex function: properties

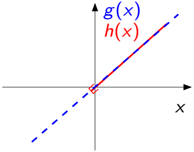
- The subdifferential of a convex, proper and l.s.c. function is
  - Maximally monotone

Subdifferential of a convex function: properties

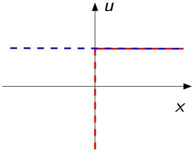
- The subdifferential of a convex, proper and l.s.c. function is
  - Maximally monotone

Example: For every $x \in \mathcal{H}$,

$$h(x) = \begin{cases} x & \text{if } x \geq 0 \\ +\infty & \text{otherwise} \end{cases}, \quad g(x) = x$$



$$\Rightarrow \quad \partial h(x) = \begin{cases} \{1\} & \text{if } x > 0 \\ ]-\infty, 1] & \text{if } x = 0 \\ \varnothing & \text{otherwise} \end{cases}, \quad \partial g(x) = \{1\}.$$



Consequently, $\mathrm{gra}\,\partial h \not\subset \mathrm{gra}\,\partial g$.

Subdifferential of a convex function: properties

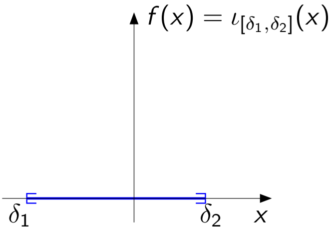
- The subdifferential of a convex, proper and l.s.c. function is
  - Maximally monotone
  - If $\mathcal{H} = \mathbb{R}$, equivalence between both properties.

Subdifferential of a convex function: example

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

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

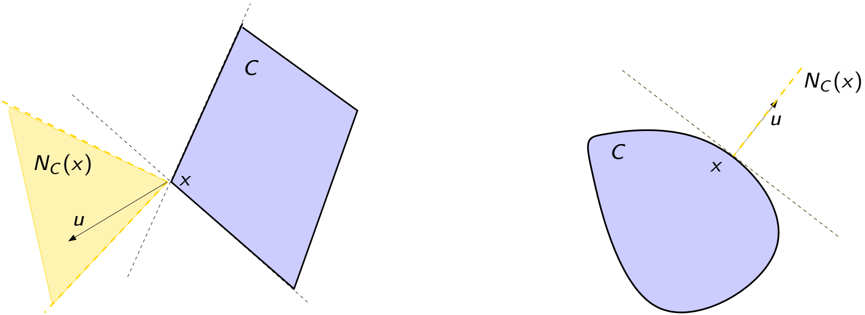
Example : $C = [\delta_1, \delta_2]$

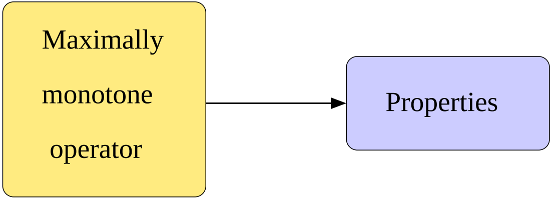


Subdifferential of a convex function: example

For every $x \in \mathcal{H}$, $\partial\iota_C(x)$ is the normal cone to C at x defined by

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




Maximally
monotone
operator
Properties

Maximally monotone operator: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
For every $x \in \mathcal{H}$, Ax is a closed convex.

Proof:

$$Ax = \bigcap_{(x',u')\in \operatorname{gra} A} \big\{ u \in \mathcal{H} \mid \langle x - x' \mid u - u' \rangle \geq 0 \big\}.$$

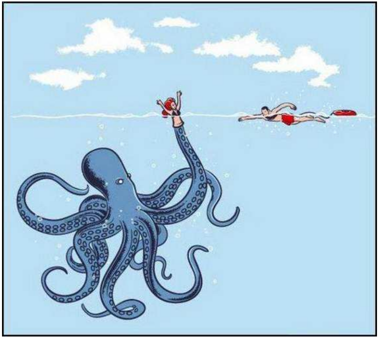
Consequently, Ax is an intersection of closed convex sets.

Maximally monotone operator: properties

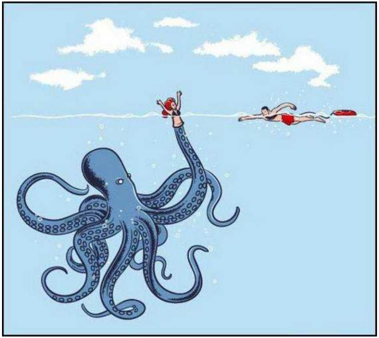
Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two maximally monotone operators.
The following operators are maximally monotone:

- $y + \gamma\rho A(\rho \cdot + z)$ where $(y, z) \in \mathcal{H}^2$, $\gamma \in [0, +\infty[$ and $\rho \in \mathbb{R}$;
- $A \times B$,
- A^{-1}.

Inverse of a maximally monotone operator



Inverse of a maximally monotone operator → Usefulness ?



Conjugate: definition

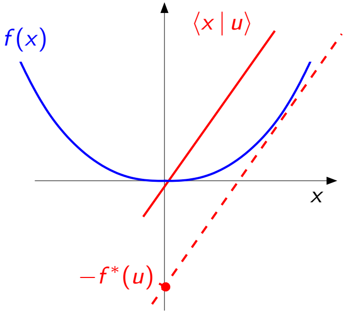
Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$.
The conjugate of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big) .$$

Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$.
The conjugate of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

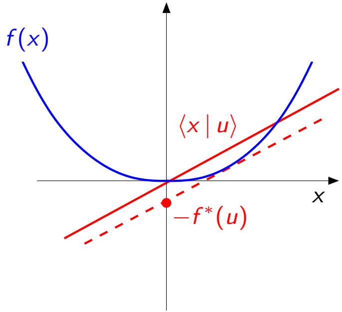
$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x \in \mathcal{H}} \big(\langle x \mid u \rangle - f(x) \big).$$



Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \,]-\infty,+\infty]$.
The **conjugate** of f is the function $f^*\colon \mathcal{H} \to [-\infty,+\infty]$ such that

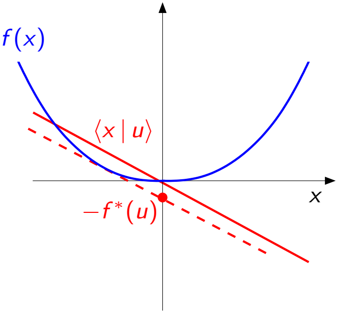
$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big).$$



Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \,]-\infty,+\infty]$.
The **conjugate** of f is the function $f^*\colon \mathcal{H} \to [-\infty,+\infty]$ such that

$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x\in\mathcal{H}} \big(\langle x \mid u\rangle - f(x)\big).$$



Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \,]-\infty, +\infty]$.
The **conjugate** of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x\in\mathcal{H}} \big(\langle x \mid u \rangle - f(x) \big).$$

Example :

- $(\forall x \in \mathbb{R}^N)\ f(x) = \frac{1}{q}\|x\|_q^q$ with $q \in]1, +\infty[$

  $\Rightarrow (\forall u \in \mathbb{R}^N)\ f^*(u) = \frac{1}{q^*}\|u\|_{q^*}^{q^*}$ with $\frac{1}{q} + \frac{1}{q^*} = 1$

Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \,]-\infty, +\infty]$.
The conjugate of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x \in \mathcal{H}} \big(\langle x \mid u \rangle - f(x) \big).$$

- f^* is l.s.c. and convex.

Conjugate: definition

Let $\mathcal{H}$ be a Hilbert space and $f\colon \mathcal{H} \to \,]-\infty, +\infty]$.
The **conjugate** of f is the function $f^*\colon \mathcal{H} \to [-\infty, +\infty]$ such that

$$(\forall u \in \mathcal{H}) \qquad f^*(u) = \sup_{x \in \mathcal{H}} \big(\langle x \mid u \rangle - f(x) \big).$$

Moreau-Fenchel theorem

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

$$f \text{ is l.s.c. and convex} \Leftrightarrow f^{**} = f.$$

- Consequence: If $f \in \Gamma_0(\mathcal{H})$ then $f^* \in \Gamma_0(\mathcal{H})$.

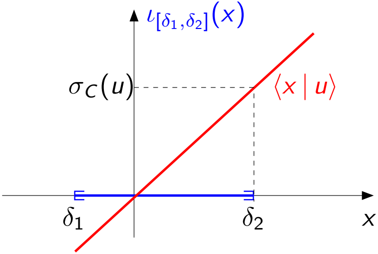
Conjugate versus Fourier transform

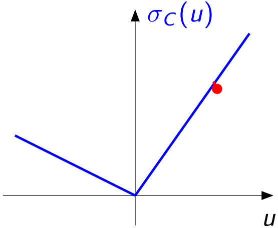
	conjugate		Fourier transform	
Property	$h(x)$	$h^*(u)$	$h(x)$	$\widehat{h}(\nu)$
invariant function	$\frac{1}{2}\|x\|^2$	$\frac{1}{2}\|u\|^2$	$\exp(-\pi\|x\|^2)$	$\exp(-\pi\|\nu\|^2)$
translation $c \in \mathcal{H}$	$f(x-c)$	$f^*(u) + \langle u \mid c\rangle$	$f(x-c)$	$\exp(-\jmath 2\pi \langle \nu \mid c\rangle)\widehat{f}(\nu)$
dual translation $c \in \mathcal{H}$	$f(x) + \langle x \mid c\rangle$	$f^*(u-c)$	$\exp(\jmath 2\pi \langle x \mid c\rangle) f(x-c)$	$\widehat{f}(\nu - c)$
scalar multiplication $\alpha \in \,]0,+\infty[$	$\alpha f(x)$	$\alpha f^*\left(\frac{u}{\alpha}\right)$	$\alpha f(x)$	$\alpha \widehat{f}(\nu)$
scaling $\alpha \in \mathbb{R}^*$	$f\left(\frac{x}{\alpha}\right)$	$f^*(\alpha u)$	$f\left(\frac{x}{\alpha}\right)$	$\|\alpha\|\widehat{f}(\alpha\nu)$
isomorphism $L \in \mathcal{B}(\mathcal{G},\mathcal{H})$	$f(Lx)$	$f^*\left((L^{-1})^\top u\right)$	$f(Lx)$	$\frac{1}{\|\det(L)\|}\widehat{f}\left((L^{-1})^\top \nu\right)$
reflection	$f(-x)$	$f^*(-u)$	$f(-x)$	$\widehat{f}(-\nu)$
separability	$\sum_{n=1}^{N} \varphi_n(x^{(n)})$ $x = (x^{(n)})_{1<n<N}$	$\sum_{n=1}^{N} \varphi_n^*(u^{(n)})$ $u = (u^{(n)})_{1<n<N}$	$\prod_{n=1}^{N} \varphi_n(x^{(n)})$ $x = (x^{(n)})_{1<n<N}$	$\prod_{n=1}^{N} \widehat{\varphi}_n(\nu^{(n)})$ $\nu = (\nu^{(n)})_{1<n<N}$
isotropy	$\psi(\|x\|)$	$\psi^*(\|u\|)$	$\psi(\|x\|)$	$\widetilde{\psi}(\|\nu\|)$
identity element of convolution	$\iota_{\{0\}}(x)$	0	$\delta(x)$	1
identity element of addition/product	0	$\iota_{\{0\}}(u)$	1	$\delta(\nu)$

Conjugate: example

Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$.
σ_C is the support function of C if

$$(\forall u \in \mathcal{H}) \qquad \sigma_C(u) = \sup_{x \in C} \langle x \mid u \rangle$$
$$= \iota_C^*(u).$$

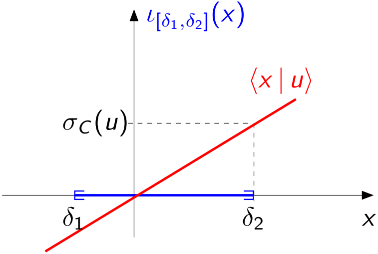


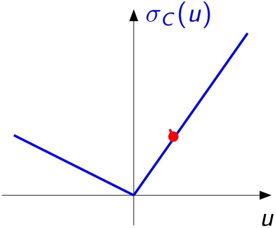


Conjugate: example

Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$.
σ_C is the support function of C if

$$(\forall u \in \mathcal{H}) \qquad \sigma_C(u) = \sup_{x \in C} \langle x \mid u \rangle$$
$$= \iota_C^*(u).$$

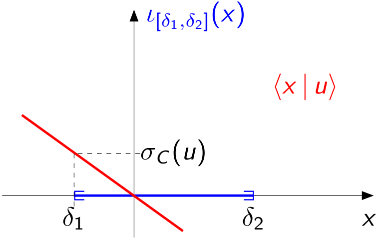


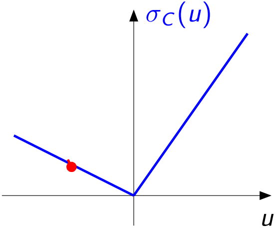


Conjugate: example

Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$.
σ_C is the support function of C if

$$(\forall u \in \mathcal{H}) \qquad \sigma_C(u) = \sup_{x \in C} \langle x \mid u \rangle$$
$$= \iota_C^*(u).$$





Conjugate: example

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

$$(\forall x \in \mathcal{H})(\forall \alpha \in \,]0,+\infty[) \qquad f(\alpha x) = \alpha f(x).$$

f is positively homogeneous and belongs to $\Gamma_0(\mathcal{H})$

$$\Updownarrow$$

$f = \sigma_C$ where C is a nonempty closed convex subset of $\mathcal{H}$.

Conjugate: example

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

$$(\forall x \in \mathcal{H})(\forall \alpha \in]0,+\infty[) \qquad f(\alpha x) = \alpha f(x).$$

f is positively homogeneous and belongs to $\Gamma_0(\mathcal{H})$

$\Updownarrow$

$f = \sigma_C$ where C is a nonempty closed convex subset of $\mathcal{H}$.

- Example 1: Let $f\colon \mathbb{R} \to]-\infty,+\infty] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases}$
with $-\infty \leq \delta_1 < \delta_2 \leq +\infty$. Then, $f = \sigma_C$ where C is the closed real interval such that $\inf C = \delta_1$ and $\sup C = \delta_2$.

Conjugate: example

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

$$(\forall x \in \mathcal{H})(\forall \alpha \in \,]0,+\infty[) \qquad f(\alpha x) = \alpha f(x).$$

f is positively homogeneous and belongs to $\Gamma_0(\mathcal{H})$

$$\Updownarrow$$

$f = \sigma_C$ where C is a nonempty closed convex subset of $\mathcal{H}$.

- Example 2: Let f be a ℓ^q norm of $\mathbb{R}^N$ with $q \in [1,+\infty]$.
  We have $f = \sigma_C$ where

$$C = \left\{ y \in \mathbb{R}^N \;\middle|\; \|y\|_{q^*} \le 1 \right\} \qquad \text{with } \frac{1}{q} + \frac{1}{q^*} = 1.$$

  Particular case : ℓ^1 norm of $\mathbb{R}^N \Rightarrow C = [-1,1]^N$.

Conjugate : subdifferential

Fenchel-Young inequality : if f proper then

$$(\forall (x,u) \in \mathcal{H}^2) \qquad f(x) + f^*(u) \geq \langle x \mid u \rangle \, .$$

Conjugate : subdifferential

Fenchel-Young inequality : if f proper then

$$(\forall (x,u) \in \mathcal{H}^2) \qquad f(x) + f^*(u) \geq \langle x \mid u \rangle .$$

Proof :

Because f proper, $f^*(u) = \sup_{y \in \mathcal{H}} \langle u \mid y \rangle - f(y) \neq -\infty$ and
$f^*(u) \geq \langle x \mid u \rangle - f(x)$.

Conjugate : subdifferential

Fenchel-Young inequality : if f proper then

$$(\forall (x,u) \in \mathcal{H}^2) \qquad f(x) + f^*(u) \geq \langle x \mid u \rangle \,.$$

If f proper then

$$(\forall (x,u) \in \mathcal{H}^2) \qquad u \in \partial f(x) \quad \Leftrightarrow \quad f(x) + f^*(u) = \langle x \mid u \rangle \,.$$

Conjugate : subdifferential

Fenchel-Young inequality : if f proper then

$$(\forall (x,u) \in \mathcal{H}^2) \qquad f(x) + f^*(u) \geq \langle x \mid u \rangle \,.$$

If f proper then

$$(\forall (x,u) \in \mathcal{H}^2) \qquad u \in \partial f(x) \quad \Leftrightarrow \quad f(x) + f^*(u) = \langle x \mid u \rangle \,.$$

Proof :

$$\begin{aligned} f(x) + f^*(u) = \langle x \mid u \rangle \quad &\Leftrightarrow \quad (\forall y \in \mathcal{H})\ \langle u \mid y \rangle - f(y) \leq \langle x \mid u \rangle - f(x) \\ &\Leftrightarrow \quad (\forall y \in \mathcal{H})\ f(y) \geq f(x) + \langle u \mid y - x \rangle \end{aligned}$$

Conjugate : subdifferential

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ proper and $x \in \mathcal{H}$.
If $u \in \partial f(x)$ then $x \in \partial f^*(u)$.

Conjugate : subdifferential

Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ proper and $x \in \mathcal{H}$.
If $u \in \partial f(x)$ then $x \in \partial f^*(u)$.

Proof :

$$
\begin{aligned}
u \in \partial f(x) \quad & \Leftrightarrow \quad (\forall y \in \mathcal{H})\ \ f(y) \geq f(x) + \langle u \mid y - x\rangle \\
& \Leftrightarrow \quad (\forall y \in \mathcal{H})\ \ f(x) - \langle u \mid x\rangle + \langle u \mid y\rangle - f(y) \leq 0 \\
& \Rightarrow \quad f(x) - \langle u \mid x\rangle + f^*(u) \leq 0 \\
& \Leftrightarrow \quad (\forall v \in \mathcal{H})\ \ \langle v \mid x\rangle - f(x) \geq f^*(u) + \langle x \mid v - u\rangle \\
\text{Fenchel-Young} \quad & \Rightarrow \quad (\forall v \in \mathcal{H})\ \ f^*(v) \geq f^*(u) + \langle x \mid v - u\rangle \\
& \Leftrightarrow \quad x \in \partial f^*(u)
\end{aligned}
$$

Conjugate : subdifferential

Let $f : \mathcal{H} \to \,]-\infty, +\infty]$ proper and $x \in \mathcal{H}$.
If $u \in \partial f(x)$ then $x \in \partial f^*(u)$.

Let $f \in \Gamma_0(\mathcal{H})$ and $x \in \mathcal{H}$.
$u \in \partial f(x) \Leftrightarrow x \in \partial f^*(u)$

Conjugate : subdifferential

Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ proper and $x \in \mathcal{H}$.
If $u \in \partial f(x)$ then $x \in \partial f^*(u)$.

Let $f \in \Gamma_0(\mathcal{H})$ and $x \in \mathcal{H}$.
$u \in \partial f(x) \Leftrightarrow x \in \partial f^*(u)$

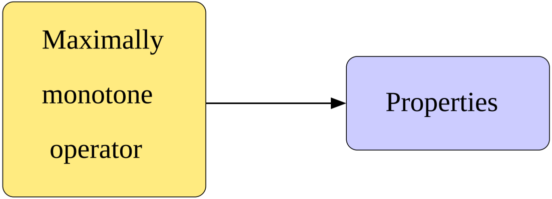
Proof :
We have $u \in \partial f(x) \Rightarrow x \in \partial f^*(u)$.
Moreover, $f \in \Gamma_0(\mathcal{H}) \Rightarrow f^* \in \Gamma_0(\mathcal{H})$ and $f^{**} = f$.
Thus $x \in \partial f^*(u) \Rightarrow u \in \partial f^{**}(x) = \partial f(x)$.

Conjugate : subdifferential

Let $f : \mathcal{H} \to]-\infty, +\infty]$ proper and $x \in \mathcal{H}$.
If $u \in \partial f(x)$ then $x \in \partial f^*(u)$.

Let $f \in \Gamma_0(\mathcal{H})$ and $x \in \mathcal{H}$.
$u \in \partial f(x) \Leftrightarrow x \in \partial f^*(u)$

Let $f \in \Gamma_0(\mathcal{H})$.
$(\partial f)^{-1} = \partial f^*$


Maximally
monotone
operator
Properties

Maximally monotone operator: sum

Let A and B be two maximally monotone operators.
$A + B$ is monotone but may not be maximally monotone.

Maximally monotone operator: sum

Let $\mathcal{H}$ be a Hilbert space.
Let A and B be two maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$ such that one of the following assumptions is satisfied:

- $\operatorname{dom} B = \mathcal{H}$
- $\operatorname{dom} A \cap \operatorname{int}(\operatorname{dom} B) \neq \varnothing$
- $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$

then $A + B$ is maximally monotone.

Maximally monotone operator: sum

Let $\mathcal{H}$ be a Hilbert space.
Let A and B be two maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$ such that one of the following assumptions is satisfied:

- $\operatorname{dom} B = \mathcal{H}$
- $\operatorname{dom} A \cap \operatorname{int}(\operatorname{dom} B) \neq \varnothing$
- $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$

then $A + B$ is maximally monotone.

Consequence: Let $\alpha \in [0, +\infty[$. If A is maximally monotone, then $A + \alpha \mathrm{Id}$ is maximally monotone.

Maximally monotone operator: linear transform

Let $\mathcal{H}$ and $\mathcal{G}$ two Hilbert spaces.
Let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be a maximally monotone operator and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that one of the following assumptions is satisfied:

- L surjective
- $0 \in \operatorname{int}(\operatorname{dom} B - \operatorname{ran} L)$

then L^*BL is maximally monotone.

Maximally monotone operator: linear transform

Let $\mathcal{H}$ and $\mathcal{G}$ two Hilbert spaces.
Let $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be a maximally monotone operator and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that one of the following assumptions is satisfied:

- L surjective
- $0 \in \operatorname{int}(\operatorname{dom} B - \operatorname{ran} L)$

then L^*BL is maximally monotone.

Consequence: Let $\mu \in \left]0, +\infty\right[$.
If A is maximally monotone and $LL^* = \mu \mathrm{Id}$, then L^*AL is maximally monotone.

Proof: $LL^* = \mu \mathrm{Id} \Rightarrow \operatorname{ran} L = \mathcal{H}$.

Part 2: Nonexpansive operators

1. Background on nonexpansive operators
   - Definition
   - Properties
   - Examples
   - Resolvent

2. Proximal operator
   - Definition
   - Properties
   - Examples

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
A is nonexpansive if $(\forall(x,y) \in C^2) \quad \|Ax - Ay\| \leq \|x - y\|$.

Nonexpansive operator: definition

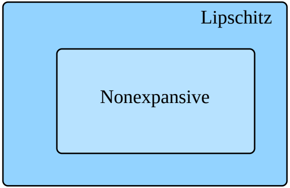
Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\nu \in \,]0, +\infty[$
$\nu^{-1}A$ is nonexpansive if $\big(\forall (x,y) \in C^2\big) \quad \|Ax - Ay\| \leq \nu \|x - y\|$.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\nu \in \,]0, +\infty[$

$\nu^{-1}A$ is nonexpansive if $\big(\forall (x,y) \in C^2\big) \quad \|Ax - Ay\| \leq \nu \|x - y\|$.

$\nu^{-1}A$ is nonexpansive $\Leftrightarrow$ A is ν-Lipschitzian.



Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$
A is firmly nonexpansive if

$$(\forall (x,u) \in \operatorname{gra} A)(\forall (y,v) \in \operatorname{gra} A) \quad \|u - v\|^2 \leq \langle u - v \mid x - y \rangle.$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
A is firmly nonexpansive if

$$(\forall x \in C)(\forall y \in C) \quad \|Ax - Ay\|^2 \leq \langle Ax - Ay \mid x - y \rangle.$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
A is **firmly nonexpansive** if

$$\big(\forall (x,y) \in C^2\big) \quad \|Ax - Ay\|^2 + \|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x - y\|^2 .$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
A is firmly nonexpansive if

$$(\forall (x,y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x - y\|^2.$$

- A is firmly nonexpansive $\Leftrightarrow$ $\mathrm{Id} - A$ is firmly nonexpansive.
- A is firmly nonexpansive $\Leftrightarrow$ $2A - \mathrm{Id}$ is nonexpansive.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
A is **firmly nonexpansive** if

$$(\forall (x,y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x - y\|^2.$$

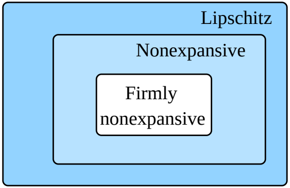
- A is firmly nonexpansive $\Leftrightarrow$ $\mathrm{Id} - A$ is firmly nonexpansive.
- A is firmly nonexpansive $\Leftrightarrow$ $\underbrace{2A - \mathrm{Id}}_{\text{Reflection of A}}$ is nonexpansive.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
A is **firmly nonexpansive** if

$$(\forall (x,y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x - y\|^2.$$

A is firmly nonexpansive $\Rightarrow$ A is nonexpansive.



Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\beta \in\ ]0,+\infty[$.
A is β-cocoercive if βA is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle .$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\beta \in\ ]0,+\infty[$.
A is β-cocoercive if βA is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle.$$

- Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $L \in \mathcal{B}(\mathcal{G},\mathcal{H})$ nonzero, and $A\colon \mathcal{H} \to \mathcal{H}$. A is β-cocoercive $\Rightarrow$ L^*AL is $\|L\|^{-2}\beta$-cocoercive.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\beta \in]0,+\infty[$.
A is β-cocoercive if βA is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle.$$

- Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $L \in \mathcal{B}(\mathcal{G},\mathcal{H})$ nonzero, and $A\colon \mathcal{H} \to \mathcal{H}$. A is β-cocoercive $\Rightarrow$ L^*AL is $\|L\|^{-2}\beta$-cocoercive.

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

  $\langle L^*ALx - L^*ALy \mid x - y\rangle = \langle ALx - ALy \mid Lx - Ly\rangle \geq \beta\|ALx - ALy\|^2$

  Moreover, $\|L^*ALx - L^*ALy\|^2 \leq \|L\|^2\|ALx - ALy\|^2$.
  Thus $\langle L^*ALx - L^*ALy \mid x - y\rangle \geq \beta\|L^*ALx - L^*ALy\|^2/\|L\|^2$.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\beta \in \,]0, +\infty[$.
A is **β-cocoercive** if βA is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle .$$

- Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ nonzero, and $A\colon \mathcal{H} \to \mathcal{H}$. A is β-cocoercive $\Rightarrow$ L^*AL is $\|L\|^{-2}\beta$-cocoercive.
- A is β-cocoercive $\Rightarrow$ A is β^{-1}-Lipschitzian.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\beta \in \,]0,+\infty[$.
A is β-cocoercive if βA is firmly nonexpansive, i.e.,

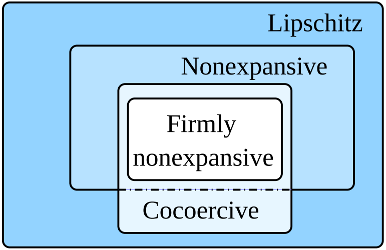
$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle.$$

- Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $L \in \mathcal{B}(\mathcal{G},\mathcal{H})$ nonzero, and $A\colon \mathcal{H} \to \mathcal{H}$. A is β-cocoercive $\Rightarrow$ L^*AL is $\|L\|^{-2}\beta$-cocoercive.
- A is β-cocoercive $\Rightarrow$ A is β^{-1}-Lipschitzian.
- $A\colon \mathcal{H} \to \mathcal{H}$ is β-cocoercive $\Rightarrow$ A is maximally monotone.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$ and $\beta \in \,]0,+\infty[$.
A is β-cocoercive if βA is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle .$$



Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A : C \to \mathcal{H}$ and let $\alpha \in]0, 1[$.
A is α-averaged if there exists a nonexpansive operator $R \colon C \to \mathcal{H}$ such that

$$A = (1 - \alpha)\mathrm{Id} + \alpha R .$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A : C \to \mathcal{H}$ and let $\alpha \in]0,1[$.
A is α-averaged if

$$\big(\forall (x,y) \in C^2\big) \quad \|Ax - Ay\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-A)x - (\mathrm{Id}-A)y\|^2 \leq \|x-y\|^2.$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A: C \to \mathcal{H}$ and let $\alpha \in]0,1[$.
A is α-averaged if

$$(\forall (x,y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id}-A)x - (\mathrm{Id}-A)y\|^2 \leq \|x-y\|^2.$$

- A is α-averaged $\Rightarrow$ A is nonexpansive.
- A is $\frac{1}{2}$-averaged $\Leftrightarrow$ A is firmly nonexpansive.
- A is α-averaged $\Rightarrow$ A is α'-averaged for every $\alpha' \in [\alpha, 1[$.
- Let $\lambda \in]0, 1/\alpha[$. A is α-averaged $\Rightarrow$ $(1-\lambda)\mathrm{Id} + \lambda A$ is $\lambda\alpha$-averaged.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A : C \to \mathcal{H}$ and let $\alpha \in]0,1[$.
A is α-averaged if

$$(\forall (x,y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x-y\|^2.$$

- Let $(\omega_i)_{1 \leq i \leq n} \in]0,1]^n$ be such that $\sum_{i=1}^n \omega_i = 1$ and let $(\alpha_i)_{1 \leq i \leq n} \in]0,1[^n$. If, for every $i \in \{1, \dots, n\}$, $A_i : C \to \mathcal{H}$ is α_i-averaged, then $\sum_{i=1}^n \omega_i A_i$ is α-averaged with $\alpha = \max_{1 \leq i \leq n} \alpha_i$.
- Let $(\alpha_i)_{1 \leq i \leq n} \in]0,1[^n$. If, for every $i \in \{1, \dots, n\}$, $A_i : C \to C$ is α_i-averaged, then $A_1 \cdots A_n$ is α-averaged with $\alpha = \dfrac{n}{n - 1 + \frac{1}{\max_{1 \leq i \leq n} \alpha_i}}$.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A: C \to \mathcal{H}$ and let $\alpha \in]0,1[$.
A is α-averaged if

$$(\forall (x,y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x-y\|^2.$$

$A: \mathcal{H} \to \mathcal{H}$ is α-averaged with $\alpha \in]0,1/2] \Rightarrow A$ is maximally monotone.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A : C \to \mathcal{H}$ and let $\alpha \in]0,1[$.
A is α-averaged if

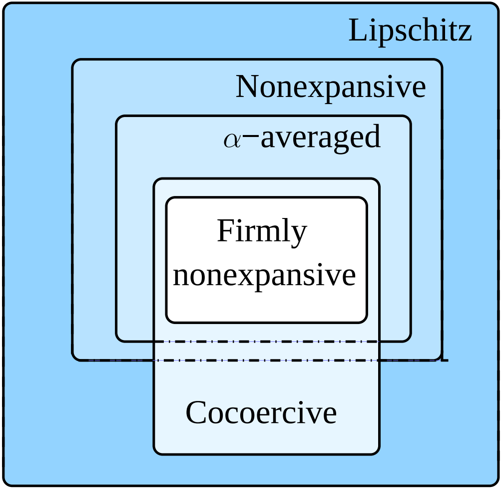
$$\big(\forall (x,y) \in C^2\big) \quad \|Ax - Ay\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \le \|x-y\|^2.$$

$A : \mathcal{H} \to \mathcal{H}$ is α-averaged with $\alpha \in]0,1/2] \Rightarrow A$ is maximally monotone.

Proof : A continuous. Moreover, for all $(x,y) \in \mathcal{H}^2$,

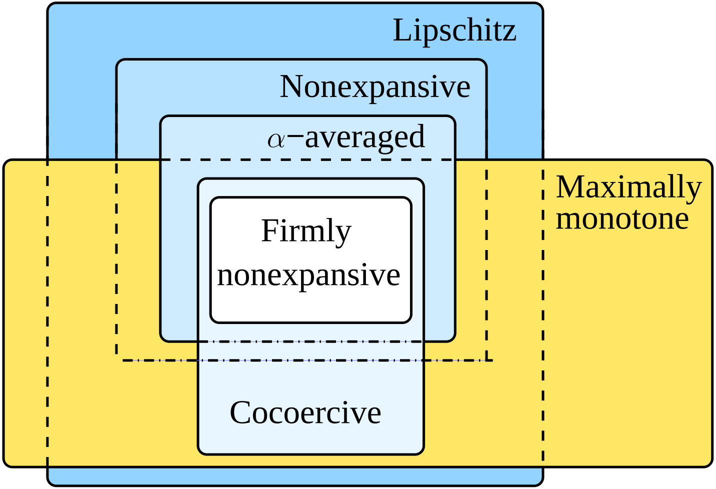
$$0 \le \|Ax - Ay\|^2 + (1-2\alpha)\|x-y\|^2 \le 2(1-\alpha)\,\langle x - y \mid Ax - Ay\rangle\,.$$

Nonexpansive operator: recap



Nonexpansive operator: recap

(if the domain C is equal to $\mathcal{H}$)



Nonexpansive operator: properties

Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty subset of $\mathcal{H}$.
Let $A\colon C \to \mathcal{H}$.
Let $\beta \in\ ]0,+\infty[$ and $\gamma \in]0,2\beta[$.
If A is β-cocoercive, then $\mathrm{Id} - \gamma A$ is $\gamma/(2\beta)$-averaged.

Proof :
A β-cocoercive $\Leftrightarrow$ βA firmly nonexpansive.
There exists a nonexpansive operator $R\colon C \to \mathcal{H}$ such that $\beta A = (\mathrm{Id} + R)/2$.
Thus

$$\mathrm{Id} - \gamma A = \Big(1 - \frac{\gamma}{2\beta}\Big)\mathrm{Id} + \frac{\gamma}{2\beta}(-R).$$

$(-R)$ being nonexpansive, $\mathrm{Id} - \gamma A$ is $\gamma/(2\beta)$-averaged.

Nonexpansive operators

Nonexpansive operators → What is their use ?

Nonexpansive operator: example

Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0, +\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Nonexpansive operator: example

Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0,+\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Descent lemma

Let $\mathcal{H}$ be a Hilbert space, $f\colon \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0,+\infty[$.
If f is differentiable and ∇f ν-Lipschitzian, then

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

Let $\mathcal{H}$ be a Hilbert space, $f\colon \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0,+\infty[$.
If f is differentiable and ∇f ν-Lipschitzian then, for all $(x,y) \in \mathcal{H}^2$,

$$f^*(\nabla f(y)) \ge f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2.$$

Nonexpansive operator: example

Descent lemma

Let $\mathcal{H}$ be a Hilbert space, $f\colon \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0,+\infty[$.
If f is differentiable and ∇f ν-Lipschitzian, then

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

Nonexpansive operator: example

Descent lemma

Let $\mathcal{H}$ be a Hilbert space, $f\colon \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0, +\infty[$.
If f is differentiable and ∇f ν-Lipschitzian, then

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

Proof :
For all $(x, y) \in \mathcal{H}^2$ and $t \in \mathbb{R}$, let $\varphi(t) = f(x + t(y - x))$.
φ is differentiable and $\varphi'(t) = \langle y - x \mid \nabla f(x + t(y - x)) \rangle$. We have then

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt$$

$$\Leftrightarrow \; f(y) - f(x) - \langle y - x \mid \nabla f(x) \rangle$$

$$= \int_0^1 \langle y - x \mid \nabla f(x + t(y - x)) - \nabla f(x) \rangle \, dt.$$

Nonexpansive operator: example

Descent lemma

Let $\mathcal{H}$ be a Hilbert space, $f: \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0, +\infty[$.
If f is differentiable and ∇f ν-Lipschitzian, then

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

Proof :

$$f(y) - f(x) - \langle y - x \mid \nabla f(x) \rangle$$
$$= \int_0^1 \langle y - x \mid \nabla f(x + t(y-x)) - \nabla f(x) \rangle \, dt.$$

From Cauchy-Schwarz inequality,

$$\langle y - x \mid \nabla f(x + t(y-x)) - \nabla f(x) \rangle$$
$$\leq \|y - x\| \|\nabla f(x + t(y-x)) - \nabla f(x)\| \leq t\nu \|y - x\|^2.$$

Nonexpansive operator: example

Let $\mathcal{H}$ be a Hilbert space, $f\colon \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0,+\infty[$.
If f is differentiable and ∇f ν-Lipschitzian then, for all $(x,y) \in \mathcal{H}^2$,

$$f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2.$$

Nonexpansive operator: example

Let $\mathcal{H}$ be a Hilbert space, $f: \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0, +\infty[$.
If f is differentiable and ∇f ν-Lipschitzian then, for all $(x, y) \in \mathcal{H}^2$,

$$f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x) \rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2.$$

Proof :
From descent lemma, for all $(x, y, z) \in \mathcal{H}^3$,

$$\begin{aligned} f^*(\nabla f(y)) &\geq \langle z \mid \nabla f(y) \rangle - f(z) \\ &\geq \langle z \mid \nabla f(y) - \nabla f(x) \rangle + \langle x \mid \nabla f(x) \rangle - f(x) - \frac{\nu}{2}\|z - x\|^2. \end{aligned}$$

From Cauchy-Schwarz inequality,

$$\langle x \mid \nabla f(x) \rangle - f(x) = f^*(\nabla f(x)).$$

Thus

$$f^*(\nabla f(y)) \geq \langle z \mid \nabla f(y) - \nabla f(x) \rangle + f^*(\nabla f(x)) - \frac{\nu}{2}\|z - x\|^2$$

Nonexpansive operator: example

Let $\mathcal{H}$ be a Hilbert space, $f\colon \mathcal{H} \to \mathbb{R}$ and $\nu \in \,]0,+\infty[$.
If f is differentiable and ∇f ν-Lipschitzian then, for all $(x,y) \in \mathcal{H}^2$,

$$f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2.$$

Proof :
Consequently,

$$\begin{aligned} f^*(\nabla f(y)) \geq& \langle z \mid \nabla f(y) - \nabla f(x)\rangle + f^*(\nabla f(x)) - \frac{\nu}{2}\|z - x\|^2\\ =& f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle\\ &+ \langle z - x \mid \nabla f(y) - \nabla f(x)\rangle - \frac{\nu}{2}\|z - x\|^2. \end{aligned}$$

This yields to

$$\begin{aligned} f^*(\nabla f(y)) \geq& f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle\\ &+ \big(\nu\|\cdot\|^2/2\big)^*\big(\nabla f(y) - \nabla f(x)\big)\\ \geq& f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2. \end{aligned}$$

Nonexpansive operator: example

Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0, +\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Nonexpansive operator: example

Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0,+\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Preuve :
Pour tout $(x,y) \in \mathcal{H}^2$,

$$f^*(\nabla f(y)) \geq f^*(\nabla f(x)) + \langle x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{2\nu}\|\nabla f(y) - \nabla f(x)\|^2$$

et symétriquement

$$f^*(\nabla f(x)) \geq f^*(\nabla f(y)) + \langle y \mid \nabla f(x) - \nabla f(y)\rangle + \frac{1}{2\nu}\|\nabla f(x) - \nabla f(y)\|^2.$$

En sommant,

$$-\langle y - x \mid \nabla f(y) - \nabla f(x)\rangle + \frac{1}{\nu}\|\nabla f(x) - \nabla f(y)\|^2 \leq 0.$$

Nonexpansive operator: example

Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0, +\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$, $\nu \in \,]0, +\infty[$ and $\gamma \in]0, 2\nu^{-1}[$.
f differentiable and ∇f ν-Lipschitzian $\Rightarrow$ $\underbrace{\mathrm{Id} - \gamma\nabla f}_{\text{gradient descent operator}}$ is $\gamma\nu/2$-averaged.

Nonexpansive operator: example

Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0, +\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$, $\nu \in \,]0, +\infty[$ and $\gamma \in]0, 2\nu^{-1}[$.
f differentiable and ∇f ν-Lipschitzian $\Rightarrow$ $\mathrm{Id} - \gamma \nabla f$ is $\gamma\nu/2$-averaged.

Nonexpansive operator: example

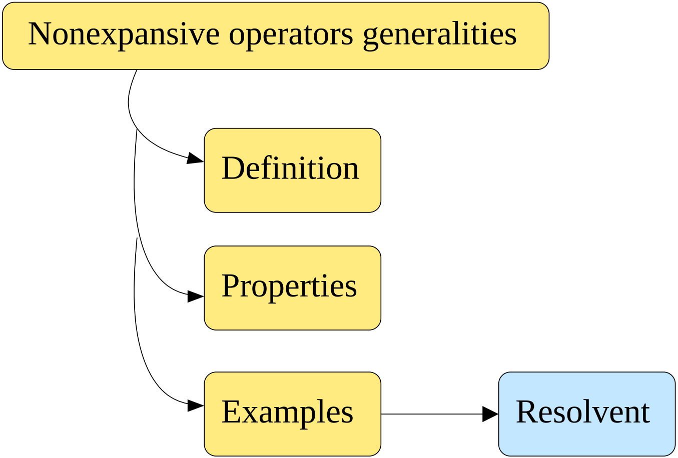
Baillon-Haddad theorem

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\nu \in \,]0,+\infty[$.
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ ∇f ν^{-1}-cocoercive.

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$, $\nu \in \,]0,+\infty[$ and $\gamma \in]0, 2\nu^{-1}[$.
f differentiable and ∇f ν-Lipschitzian $\Rightarrow$ $\mathrm{Id} - \gamma\nabla f$ is $\gamma\nu/2$-averaged.

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$, and $\nu \in \,]0,+\infty[$
f differentiable and ∇f ν-Lipschitzian $\Leftrightarrow$ f^* is ν^{-1}-strongly convex.

Remark : f^* is ν^{-1}-strongly convex if $f^* - \nu^{-1}\|\cdot\|^2/2$ is convex.


Nonexpansive operators generalities
Definition
Properties
Examples
Resolvent

Resolvent: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

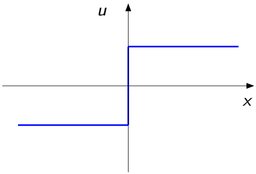
$$J_A = (\mathrm{Id} + A)^{-1}.$$

Resolvent: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

$$J_A = (\mathrm{Id} + A)^{-1}.$$

- Example :



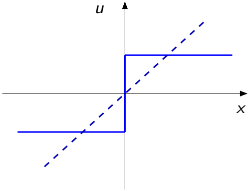
A $\quad$ $A + \mathrm{Id}$? $\quad$ J_A ?

Resolvent: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

$$J_A = (\mathrm{Id} + A)^{-1}.$$

- Example :



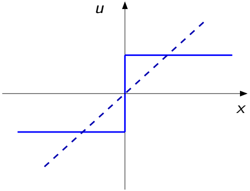
A and Id $\qquad$ $A + \mathrm{Id}$? $\qquad$ J_A ?

Resolvent: definition

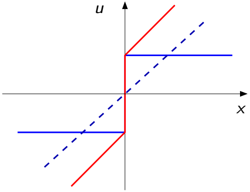
Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

$$J_A = (\mathrm{Id} + A)^{-1}.$$

- Example :



A and Id



$A + \mathrm{Id}$

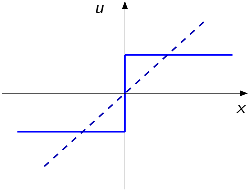
J_A ?

Resolvent: definition

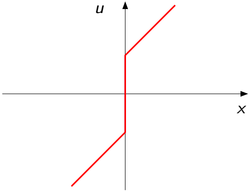
Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

$$J_A = (\mathrm{Id} + A)^{-1}.$$

- Example :



A and Id



$A + \mathrm{Id}$

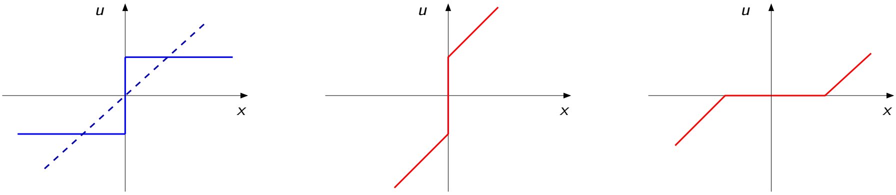
J_A ?

Resolvent: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

$$J_A = (\mathrm{Id} + A)^{-1}.$$

- Example :



A and Id — $A + \mathrm{Id}$ — J_A

Resolvent: definition

The range of an operator $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ is

$$\operatorname{ran} B = \left\{ u \in \mathcal{H} \;\middle|\; \exists x \in \mathcal{H}, u \in Bx \right\}.$$

Minty theorem

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator.

$$\operatorname{ran}(\mathrm{Id} + A) = \mathcal{H} \quad \Leftrightarrow \quad A \text{ is maximally monotone.}$$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Proof : A is monotone

$$\Leftrightarrow \big(\forall (x,u) \in \mathrm{gra} A\big)\big(\forall (y,v) \in \mathrm{gra} A\big) \quad \langle x-y \mid u-v\rangle \geq 0$$
$$\Leftrightarrow \big(\forall (x,u) \in \mathrm{gra} A\big)\big(\forall (y,v) \in \mathrm{gra} A\big) \quad \langle x-y \mid x-y+u-v\rangle \geq \|x-y\|^2$$
$$\Leftrightarrow \big(\forall (x,u') \in \mathrm{gra}(\mathrm{Id}+A)\big)\big(\forall (y,v') \in \mathrm{gra}(\mathrm{Id}+A)\big) \quad \big\langle x-y \mid u'-v'\big\rangle \geq \|x-y\|^2$$
$$\Leftrightarrow \big(\forall (u',x) \in \mathrm{gra} J_A\big)\big(\forall (v',y) \in \mathrm{gra} J_A\big) \quad \big\langle u'-v' \mid x-y\big\rangle \geq \|x-y\|^2$$
$$\Leftrightarrow J_A \text{ firmly nonexpansive}$$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Remark : $J_A : \operatorname{ran}(\mathrm{Id} + A) \to \mathcal{H}$.

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is maximally monotone $\Leftrightarrow$ $J_A \colon \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive.

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is maximally monotone $\Leftrightarrow$ $J_A\colon \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive.

Proof: A monotone $\Leftrightarrow$ $J_A\colon \mathrm{ran}\,(\mathrm{Id} + A) \to \mathcal{H}$ firmly nonexpansive
$+$ Minty theorem.

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is maximally monotone $\Leftrightarrow$ $J_A : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive.

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ maximally monotone and $\gamma \in \,]0, +\infty[$. For every $x \in \mathcal{H}$, there exists a unique $p \in \mathcal{H}$ such that $x - p \in \gamma A p$ and thus $p = J_{\gamma A} x$.

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is monotone $\Leftrightarrow$ J_A is firmly nonexpansive.

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
A is maximally monotone $\Leftrightarrow$ $J_A \colon \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive.

Let $\mathcal{H}$ be a Hilbert space. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ maximally monotone and $\gamma \in \,]0, +\infty[$. For every $x \in \mathcal{H}$, there exists a unique $p \in \mathcal{H}$ such that $x - p \in \gamma A p$ and thus $p = J_{\gamma A} x$.

Proof: $x \in (\mathrm{Id} + \gamma A)(p) \Leftrightarrow p \in (\mathrm{Id} + \gamma A)^{-1} x \Leftrightarrow p = J_{\gamma A} x$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone and $\gamma \in]0, +\infty[$.

- $J_{\gamma A}$ and $\mathrm{Id} - J_{\gamma A}$ are firmly nonexpansive.
- The reflected resolvent $R_{\gamma A} = 2J_{\gamma A} - \mathrm{Id}$ is nonexpansive.
- ${}^{\gamma}A$ is γ-cocoercive.

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ and $\gamma \in]0, +\infty[$.
The Yosida approximation of A of index γ is

$$ {}^{\gamma}A = \frac{1}{\gamma}(\mathrm{Id} - J_{\gamma A}). $$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.

- Let $z \in \mathcal{H}$ and $B = A(\cdot - z)$. Then $J_B = z + J_A(\cdot - z)$.
- Let $z \in \mathcal{H}$ and $B = z + A$. Then $J_B = J_A(\cdot - z)$.
- Let $\alpha \in [0, +\infty[$ and $B = A + \alpha \mathrm{Id}$. Then $J_B = J_{\frac{A}{1+\alpha}}\left(\frac{\cdot}{1+\alpha}\right)$

Proof:

For all $x \in \mathcal{H}$,

$$
\begin{aligned}
p = J_{A+\alpha \mathrm{Id}} x \quad & \Leftrightarrow \quad x - p \in (A + \alpha \mathrm{Id})(p) \\
& \Leftrightarrow \quad (1+\alpha)^{-1} x - p \in (1+\alpha)^{-1} A p \\
& \Leftrightarrow \quad p = J_{(1+\alpha)^{-1} A}\big((1+\alpha)^{-1} x\big).
\end{aligned}
$$

Resolvent: properties

For every $i \in \{1, \dots, n\}$, let $\mathcal{H}_i$ be a Hilbert space and $A_i \colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ be a maximally monotone operator.

$$J_{A_1 \times \cdots \times A_n} = J_{A_1} \times \cdots \times J_{A_n} \colon \mathcal{H}_1 \times \cdots \times \mathcal{H}_n \to \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$$
$$(x_1, \dots, x_n) \mapsto (J_{A_1} x_1, \dots, J_{A_n} x_n).$$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $\gamma \in]0, +\infty[$.

$$J_{\gamma A^{-1}} = \mathrm{Id} - \gamma J_{\gamma^{-1} A}(\gamma^{-1} \cdot)$$

Proof: For all $x \in \mathcal{H}$,

$$\begin{aligned}
p = J_{\gamma A^{-1}} x \quad & \Leftrightarrow \quad x \in (\mathrm{Id} + \gamma A^{-1})(p) \\
& \Leftrightarrow \quad \gamma^{-1}(x - p) \in A^{-1} p \\
& \Leftrightarrow \quad p \in A\big(\gamma^{-1}(x - p)\big) \\
& \Leftrightarrow \quad \gamma^{-1} p \in \gamma^{-1} A\big(\gamma^{-1}(x - p)\big) \\
& \Leftrightarrow \quad \gamma^{-1} x \in (\mathrm{Id} + \gamma^{-1} A)\big(\gamma^{-1}(x - p)\big) \\
& \Leftrightarrow \quad \gamma^{-1}(x - p) = J_{\gamma^{-1} A}\big(\gamma^{-1} x\big) \\
& \Leftrightarrow \quad p = x - \gamma J_{\gamma^{-1} A}\big(\gamma^{-1} x\big).
\end{aligned}$$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $\gamma \in \,]0, +\infty[$.

$$J_{\gamma A^{-1}} = \mathrm{Id} - \gamma J_{\gamma^{-1} A}(\gamma^{-1}\cdot)$$

Remark: $J_A + J_{A^{-1}} = \mathrm{Id}$.

Resolvent: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $LL^* = \mu \mathrm{Id}$ where $\mu \in \,]0, +\infty[$. Then

$$J_{L^*AL} = \mathrm{Id} - L^* \circ {}^{\mu}A \circ L \; .$$

Resolvent: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $LL^* = \mu \mathrm{Id}$ where $\mu \in \,]0, +\infty[$. Then

$$J_{L^*AL} = \mathrm{Id} - L^* \circ {}^{\mu}A \circ L \,.$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ be a unitary operator. Then

$$J_{L^*AL} = L^* J_A L.$$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $B = \rho A(\rho\cdot)$ where $\rho \in \mathbb{R}^*$. Then

$$J_B = \rho^{-1} J_{\rho^2 A}(\rho\cdot).$$

Proof: Set $L = \rho\mathrm{Id}$ and apply formula

$$J_{L^* A L} = \mathrm{Id} - L^* \circ {}^{\mu}A \circ L.$$

Resolvent: properties

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $B = \rho A(\rho \cdot)$ where $\rho \in \mathbb{R}^*$. Then

$$J_B = \rho^{-1} J_{\rho^2 A}(\rho \cdot).$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and $B = -A(-\cdot)$. Then

$$J_B = -J_A(-\cdot).$$

Resolvent →

Resolvent →

→ Proximity operator

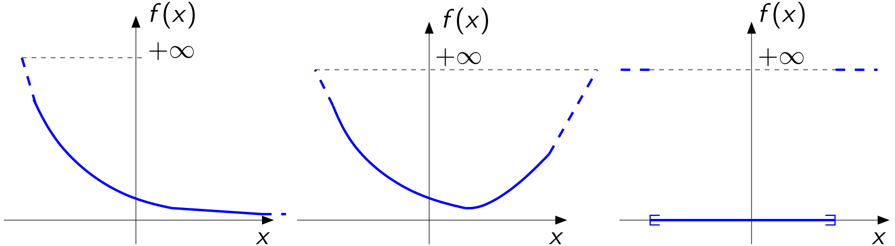
Convex analysis

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

Convex analysis

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

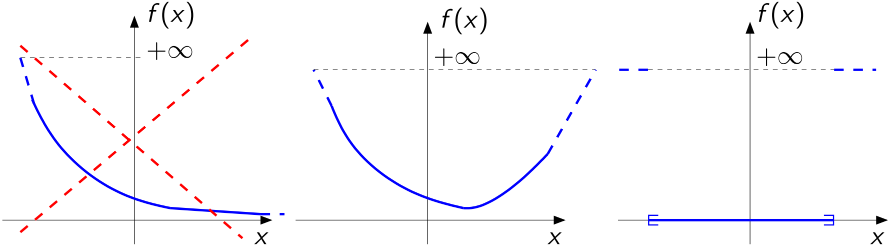
Coercive functions ?



Convex analysis

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

Coercive functions ?



Convex analysis

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

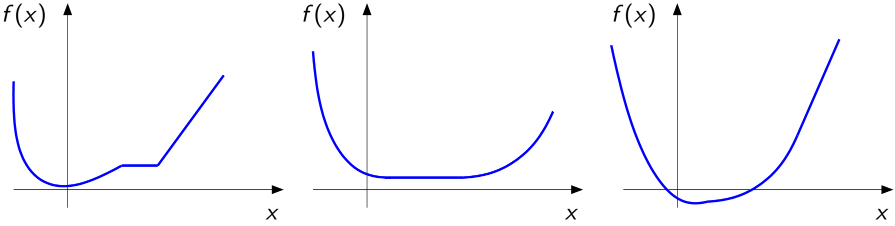
$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{dom} f)(\forall \alpha \in]0,1[)$$
$$x \neq y \quad \Rightarrow \quad f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y).$$

Convex analysis

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

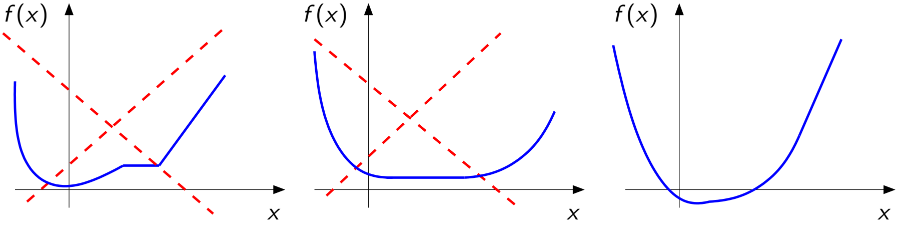


Convex analysis

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 ?



Convex analysis

Let $\mathcal{H}$ be a Hilbert space and C be a closed convex of $\mathcal{H}$.
Let $f \in \Gamma_0(\mathcal{H})$ such that $\mathrm{dom}\, f \cap C \neq \varnothing$.
If f is coercive or C is bounded, then there exists $p \in C$ such that

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

Moreover, if f is strictly convex, this minimizer p is unique.

Proximity operator: definition

Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0,+\infty[$.
For every $x \in \mathcal{H}$, there exists a unique $p \in \mathcal{H}$ such that

$$f(p) + \frac{1}{2\gamma}\|p - x\|^2 = \inf_{y\in\mathcal{H}} f(y) + \frac{1}{2\gamma}\|y - x\|^2 .$$

Proximity operator: definition

Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0,+\infty[$.
For every $x \in \mathcal{H}$, there exists a unique $p \in \mathcal{H}$ such that

$$f(p) + \frac{1}{2\gamma}\|p - x\|^2 = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma}\|y - x\|^2 .$$

Proof: $f \in \Gamma_0(\mathcal{H}) \Leftrightarrow f^* \in \Gamma_0(\mathcal{H})$. There exists $u \in \mathcal{H}$ such that $f^*(u) \in \mathbb{R}$. From Fenchel-Young inequality, we have

$$(\forall y \in \mathcal{H}) \qquad f(y) \geq \langle u \mid y \rangle - f^*(u).$$

Then $f(y) + (2\gamma)^{-1}\|y - x\|^2 \to +\infty$ when $\|y\| \to +\infty$.
Moreover $(2\gamma)^{-1}\| \cdot - x\|^2$ being strictly convex, $f + (2\gamma)^{-1}\| \cdot - x\|^2$ is a strictly convex coercive function.

Proximity operator: definition

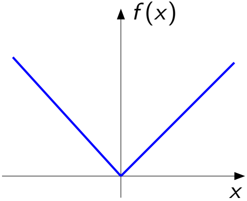
Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$.

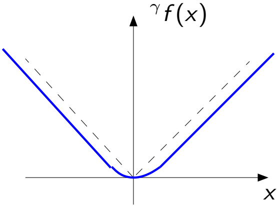
- The **Moreau envelope** of f of parameter $\gamma \in \,]0, +\infty[$ is

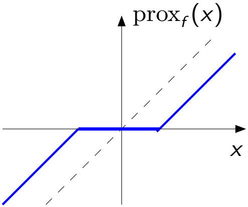
$$^{\gamma}f : \mathcal{H} \to \mathbb{R} : x \mapsto \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma}\|y - x\|^2.$$

- The **proximity operator** of f is

$$\mathrm{prox}_f : \mathcal{H} \to \mathcal{H} : x \mapsto \operatorname*{argmin}_{y \in \mathcal{H}} \; f(y) + \frac{1}{2}\|y - x\|^2.$$







Proximity operator: definition

Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$.

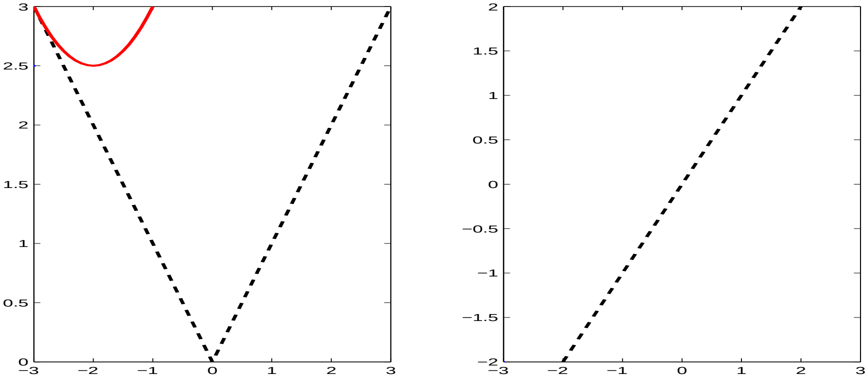
- The Moreau envelope of f of parameter $\gamma \in]0, +\infty[$ is

$$^{\gamma}f\colon \mathcal{H} \to \mathbb{R}\colon x \mapsto \inf_{y\in\mathcal{H}} f(y) + \frac{1}{2\gamma}\|y-x\|^2.$$

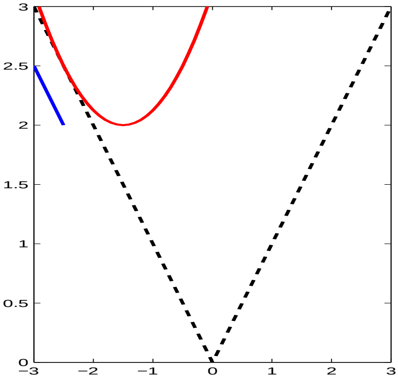
- The proximity operator of f is

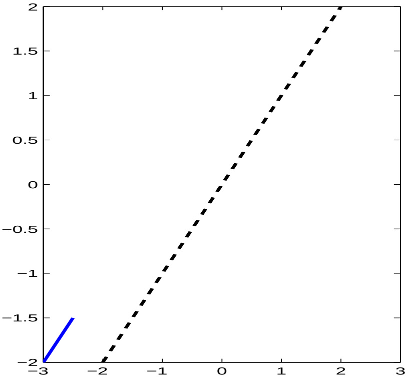
$$\mathrm{prox}_f\colon \mathcal{H} \to \mathcal{H}\colon x \mapsto \underset{y\in\mathcal{H}}{\mathrm{argmin}}\ f(y) + \frac{1}{2}\|y-x\|^2.$$

Proximity operator: definition

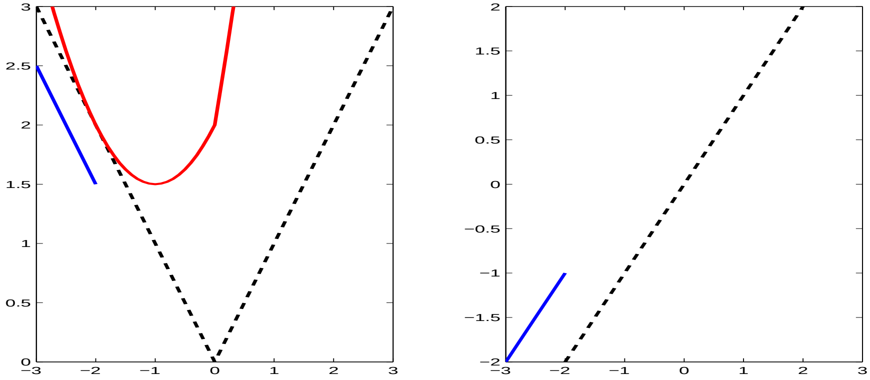


Proximity operator: definition

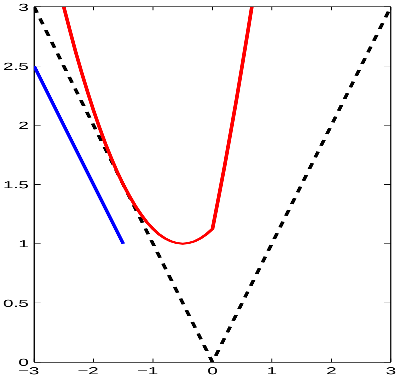


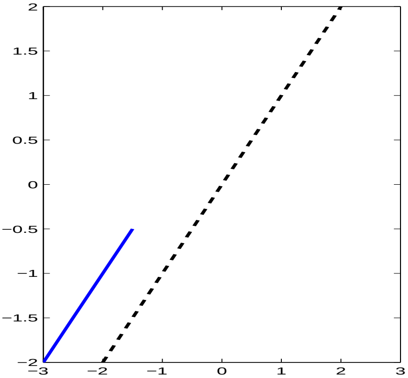


Proximity operator: definition

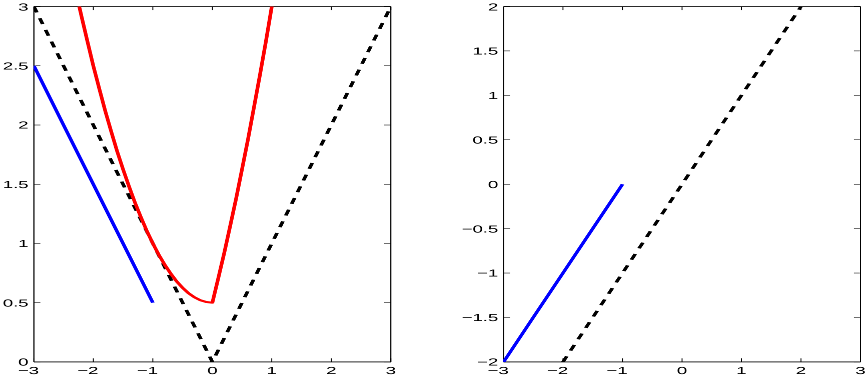


Proximity operator: definition

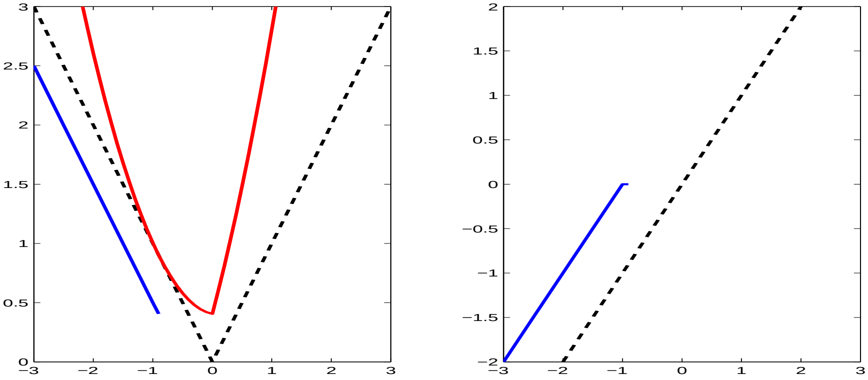




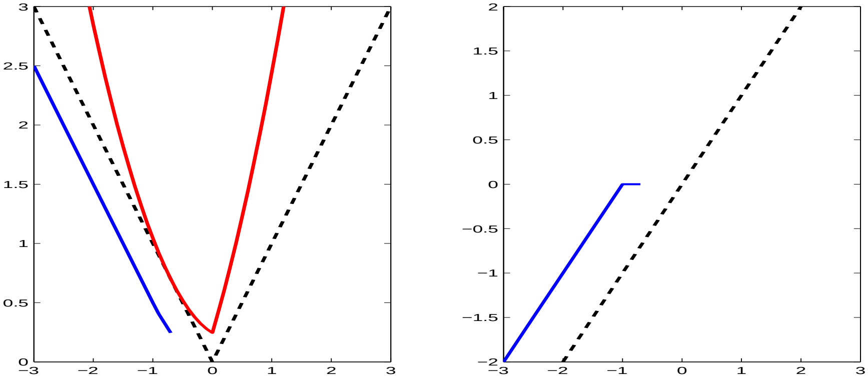
Proximity operator: definition



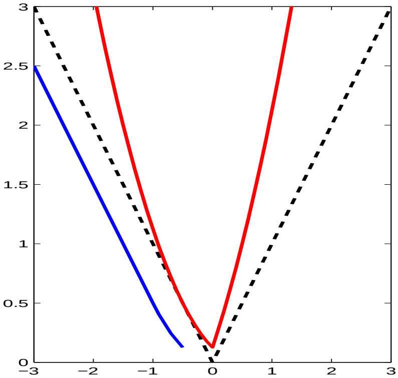
Proximity operator: definition

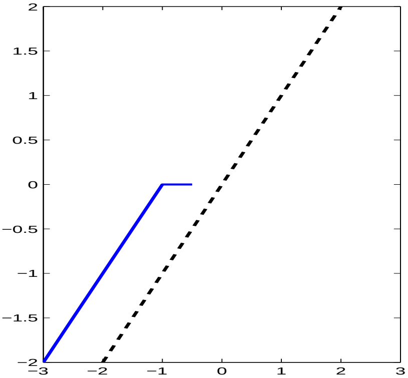


Proximity operator: definition

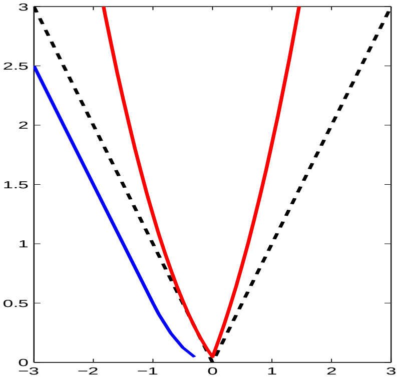


Proximity operator: definition

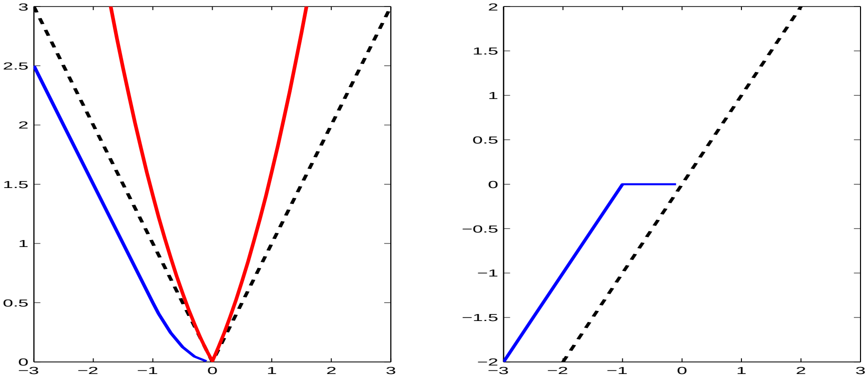




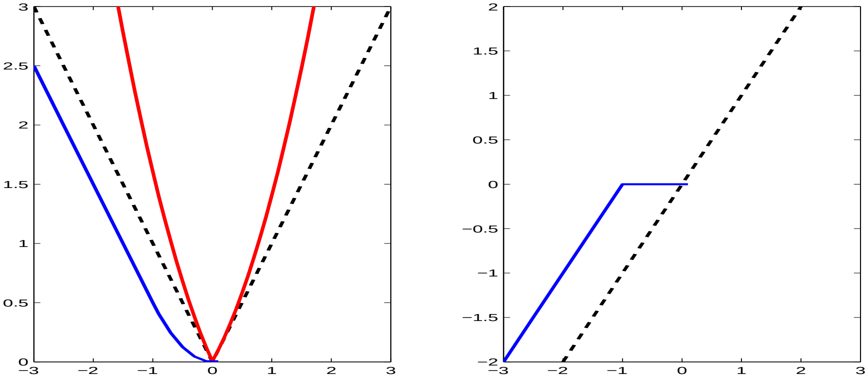
Proximity operator: definition



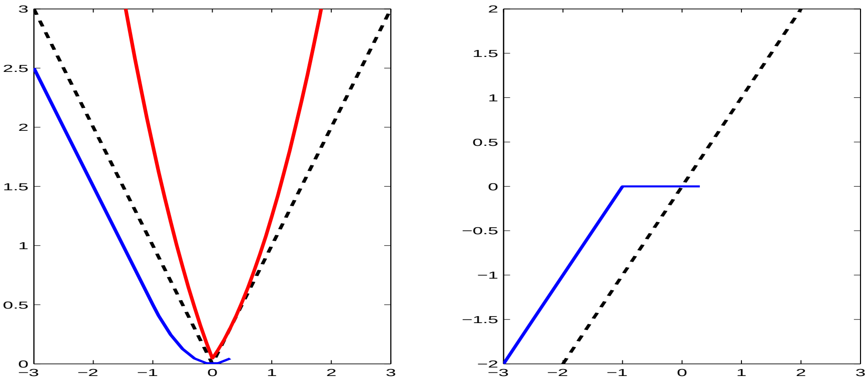
Proximity operator: definition



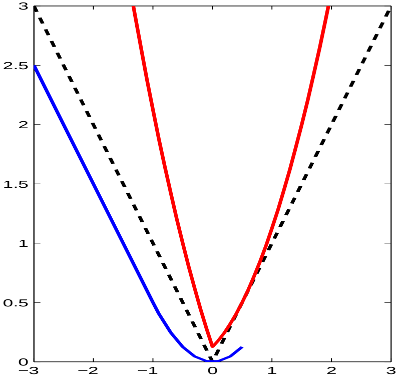
Proximity operator: definition

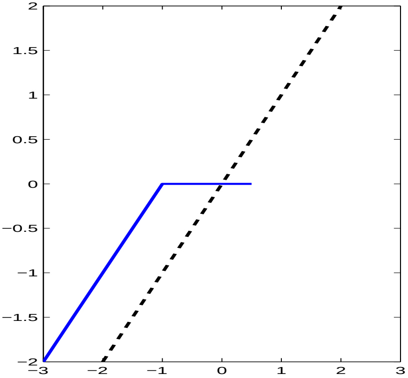


Proximity operator: definition

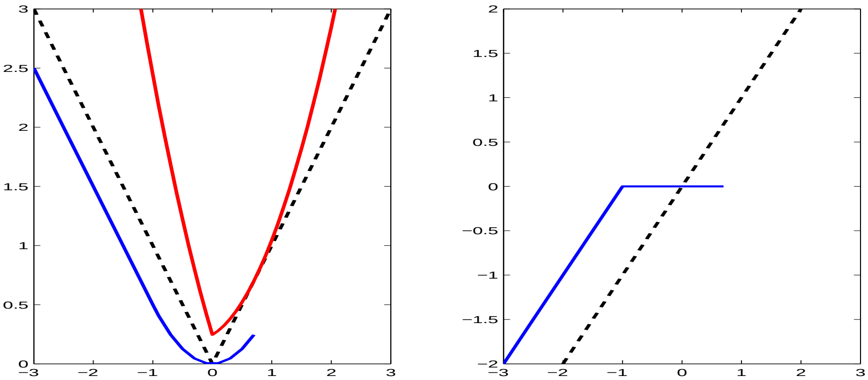


Proximity operator: definition

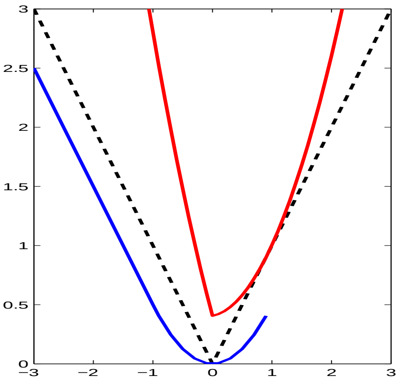


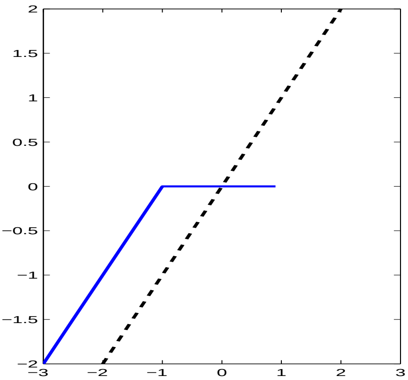


Proximity operator: definition

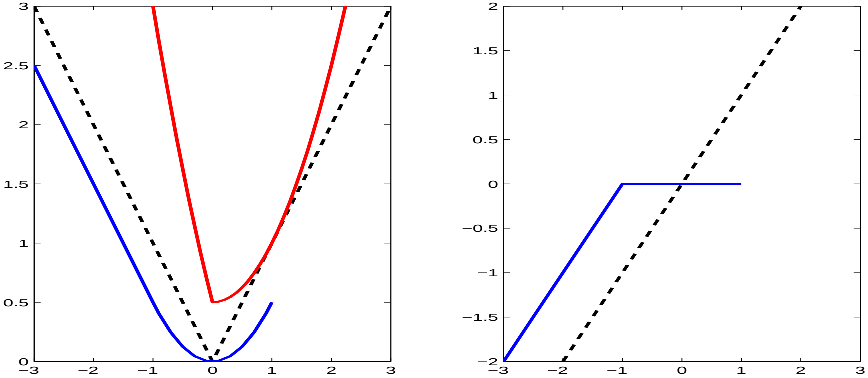


Proximity operator: definition

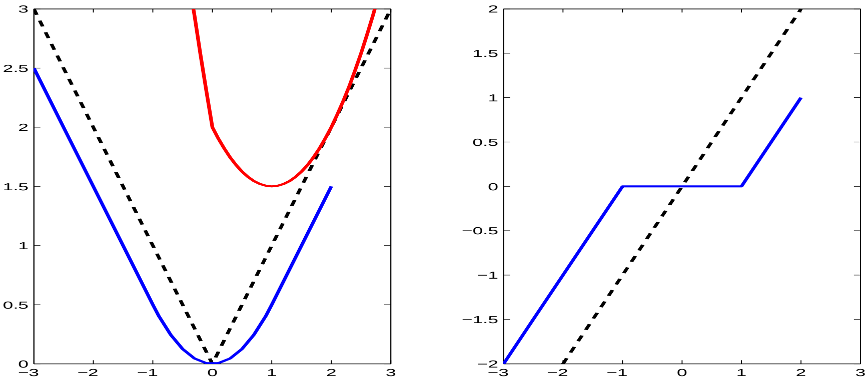




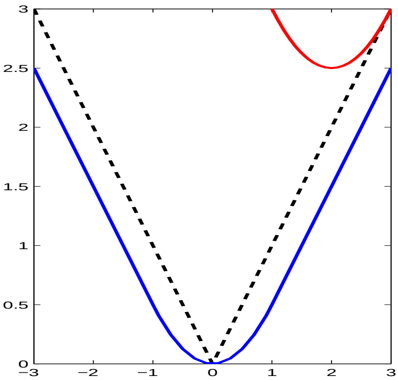
Proximity operator: definition

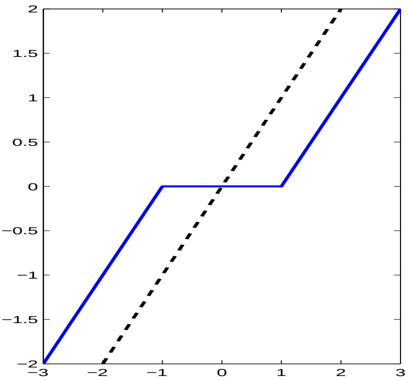


Proximity operator: definition



Proximity operator: definition





Proximity operator: definition

Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.
If $\operatorname{dom} f \cap \operatorname{int}(\operatorname{dom} g) \neq \varnothing$ then $\partial(f+g) = \partial f + \partial g$.

Let $\mathcal{H}$ be a Hilbert space and $f \in \Gamma_0(\mathcal{H})$.

$$\operatorname{prox}_f = J_{\partial f}.$$

Proximity operator: definition

Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.
If $\mathrm{dom}\, f \cap \mathrm{int}\,(\mathrm{dom}\, g) \neq \varnothing$ then $\partial(f+g) = \partial f + \partial g$.

Let $\mathcal{H}$ be a Hilbert space and $f \in \Gamma_0(\mathcal{H})$.

$$\mathrm{prox}_f = J_{\partial f}.$$

Proof: By using Fermat's rule, for every $x \in \mathcal{H}$,

$$\begin{aligned}
p = \arg\min \;\; f + (2\gamma)^{-1}\|\cdot - x\|^2 &\Leftrightarrow \quad 0 \in \partial\Big(f + \frac{1}{2}\|\cdot - x\|^2\Big)(p) \\
&\Leftrightarrow \quad 0 \in \partial f(p) + p - x \\
&\Leftrightarrow \quad x \in (\mathrm{Id} + \partial f)(p) \\
&\Leftrightarrow \quad p = (\mathrm{Id} + \partial f)^{-1}(x).
\end{aligned}$$

Proximity operator: definition

Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.
If $\operatorname{dom} f \cap \operatorname{int}(\operatorname{dom} g) \neq \varnothing$ then $\partial(f+g) = \partial f + \partial g$.

Let $\mathcal{H}$ be a Hilbert space and $f \in \Gamma_0(\mathcal{H})$.

$$\operatorname{prox}_f = J_{\partial f}.$$

Remark: As $\operatorname{dom}(\operatorname{prox}_f) = \mathcal{H}$, this provides a proof that ∂f is maximally monotone !

Proximity operator: properties

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $(x, p) \in \mathcal{H}^2$.

$$p = \mathrm{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \;\; \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Proximity operator: properties

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $(x, p) \in \mathcal{H}^2$.

$$p = \mathrm{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \ \ \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Proof: $(\Rightarrow)$ Let $p_\alpha = \alpha y + (1-\alpha)p$ where $y \in \mathcal{H}$ and $\alpha \in]0,1]$. We have

$$f(p) + \frac{1}{2}\|p - x\|^2 \leq f(p_\alpha) + \frac{1}{2}\|p_\alpha - x\|^2$$
$$\leq \alpha f(y) + (1-\alpha)f(p) + \frac{1}{2}\|p - x + \alpha(y-p)\|^2.$$

Consequently
$f(p) \leq \alpha f(y) + (1-\alpha)f(p) + \alpha \langle y - p \mid p - x \rangle + \frac{\alpha^2}{2}\|y - p\|^2$
$\Leftrightarrow \langle y - p \mid x - p \rangle + f(p) \leq f(y) + \frac{\alpha}{2}\|y - p\|^2.$
The results comes from $\alpha \to 0$.

Proximity operator: properties

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $(x, p) \in \mathcal{H}^2$.

$$p = \mathrm{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \;\; \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Preuve: $(\Leftarrow)$

On a, pour tout $y \in \mathcal{H}$,

$$\begin{aligned} f(p) + \frac{1}{2}\|p - x\|^2 &\leq f(y) + \langle y - p \mid p - x \rangle + \frac{1}{2}\|p - x\|^2 \\ &= f(y) + \frac{1}{2}\|y - p + p - x\|^2 - \frac{1}{2}\|y - p\|^2 \\ &\leq f(y) + \frac{1}{2}\|y - x\|^2. \end{aligned}$$

Proximity operator: properties

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $(x, p) \in \mathcal{H}^2$.

$$p = \mathrm{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \;\; \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0, +\infty[$.
${}^{\gamma}f$ is differentiable and $\nabla\, {}^{\gamma}f$ is γ^{-1}-Lipschitzian

$$(\forall x \in \mathcal{H}) \quad \nabla \underbrace{{}^{\gamma}f}_{\text{Moreau envelope}} = \gamma^{-1}(\mathrm{Id} - \mathrm{prox}_{\gamma f}) = \underbrace{{}^{\gamma}\partial f}_{\text{Yosida approximation}}.$$

<u>Proof</u>: Previous property + ... calculations.

Proximity operator: properties

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $(x, p) \in \mathcal{H}^2$.

$$p = \mathrm{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \ \ \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0, +\infty[$.
${}^{\gamma}f$ is differentiable and $\nabla {}^{\gamma}f$ is γ^{-1}-Lipschitzian

$$(\forall x \in \mathcal{H}) \quad \nabla \underbrace{{}^{\gamma}f}_{\text{Moreau envelope}} = \gamma^{-1}(\mathrm{Id} - \mathrm{prox}_{\gamma f}) = \underbrace{{}^{\gamma}\partial f}_{\text{Yosida approximation}} .$$

Interpretation: ${}^{\gamma}f$ is a smooth approximation of f.

Proximity operator: properties

Let $\mathcal{H}$ be a Hilbert space, $x \in \mathcal{H}$ and $f \in \Gamma_0(\mathcal{H})$.

Properties	$g(x)$	$\mathrm{prox}_g x$
Translation	$f(x-z), z \in \mathcal{H}$	$z + \mathrm{prox}_f(x-z)$
Quadratic perturbation	$f(x) + \alpha \parallel x \parallel^2 /2 + \langle z \mid x \rangle + \gamma$ $z \in \mathcal{H}, \alpha > 0, \gamma \in \mathbb{R}$	$\mathrm{prox}_{\frac{f}{\alpha+1}}(\frac{x-z}{\alpha+1})$
Scale change	$f(\rho x), \rho \in \mathbb{R}^*$	$\frac{1}{\rho}\mathrm{prox}_{\rho^2 f}(\rho x)$
Reflection	$f(-x)$	$-\mathrm{prox}_f(-x)$
Moreau envelope	${}^\gamma f(x) = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma}\|x-y\|^2$ $\gamma > 0$	$\frac{1}{1+\gamma}\left(\gamma x + \mathrm{prox}_{(1+\gamma)f}(x)\right)$

Proximity operator: properties

For every $i \in \{1, \ldots, n\}$, let $\mathcal{H}_i$ be a Hilbert space and $f_i \in \Gamma_0(\mathcal{H}_i)$.
For all $(x_1, \ldots, x_n) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_n$,
if

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i).$$

then

$$\mathrm{prox}_f(x_1, \ldots, x_n) = \left(\mathrm{prox}_{f_i}(x_i)\right)_{1 \le i \le n}.$$

Proximity operator: properties

Let $\mathcal{H}$ be a separable Hilbert space.
Let $(b_i)_{i\in I}$ be an orthonormal basis of $\mathcal{H}$.
For every $i \in I$, let $\varphi_i \in \Gamma_0(\mathbb{R})$ such that $\varphi_i \geq 0$. For every $x \in \mathcal{H}$,
if

$$f(x) = \sum_{i\in I} \varphi_i(\langle x \mid b_i \rangle)$$

then

$$\operatorname{prox}_f(x) = \sum_{i\in I} \operatorname{prox}_{\varphi_i}(\langle x \mid b_i \rangle) b_i.$$

Remark: The assumption $(\forall i \in I)\ \varphi_i \geq 0$ can be relaxed if $\mathcal{H}$ is finite dimensional.

Proximity operator: properties

Let $\mathcal{H}$ be a separable Hilbert space.
Let $(b_i)_{i\in I}$ be an orthonormal basis of $\mathcal{H}$.
For every $i \in I$, let $\varphi_i \in \Gamma_0(\mathbb{R})$ such that $\varphi_i \geq 0$. For every $x \in \mathcal{H}$, if

$$f(x) = \sum_{i\in I} \varphi_i(\langle x \mid b_i\rangle)$$

then

$$\mathrm{prox}_f(x) = \textstyle\sum_{i\in I} \mathrm{prox}_{\varphi_i}(\langle x \mid b_i\rangle) b_i.$$

Example: $\mathcal{H} = \mathbb{R}^N$, $(b_i)_{1\leq i\leq N}$ canonical basis of $\mathbb{R}^N$, $f = \lambda\|\cdot\|_1$ with $\lambda \in [0, +\infty[$.

$$(\forall x = (x^{(i)})_{1\leq i\leq N} \in \mathbb{R}^N) \qquad \mathrm{prox}_{\lambda\|\cdot\|_1}(x) = \big(\mathrm{prox}_{\lambda|\cdot|}(x^{(i)})\big)_{1\leq i\leq N}$$

Proximity operator: properties

Moreau decomposition formula

Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0,+\infty[$.

$$(\forall x \in \mathcal{H}) \qquad \mathrm{prox}_{\gamma f^*} x = x - \gamma \mathrm{prox}_{\gamma^{-1} f}(\gamma^{-1} x).$$

Proximity operator: properties

Moreau decomposition formula

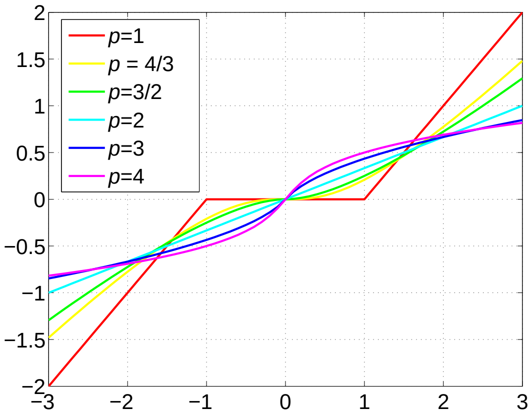
Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0, +\infty[$.

$$(\forall x \in \mathcal{H}) \qquad \operatorname{prox}_{\gamma f^*} x = x - \gamma \operatorname{prox}_{\gamma^{-1} f}(\gamma^{-1} x).$$

Example: If $\mathcal{H} = \mathbb{R}^N$, $f = \frac{1}{q}\|\cdot\|_q^q$ with $q \in]1, +\infty[$, then $f^* = \frac{1}{q^*}\|\cdot\|_{q^*}^{q^*}$ with $1/q + 1/q^* = 1$, and

$$(\forall x \in \mathbb{R}^N) \qquad \operatorname{prox}_{\frac{\gamma}{q^*}\|\cdot\|_{q^*}^{q^*}} x = x - \gamma \operatorname{prox}_{\frac{1}{\gamma q}\|\cdot\|_q^q}(\gamma^{-1} x).$$

Proximity operator: properties



Proximity operator: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $\mathrm{ran}\, L = \mathcal{H}$. Then

$$\partial(f \circ L) = L^* \, \partial f \, L.$$

Proximity operator: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $\operatorname{ran} L = \mathcal{H}$. Then

$$\partial(f \circ L) = L^* \, \partial f \, L.$$

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $LL^* = \mu \mathrm{Id}$ where $\mu \in \,]0, +\infty[$. Then

$$\mathrm{prox}_{f \circ L} = \mathrm{Id} - \mu^{-1} L^* \circ (\mathrm{Id} - \mathrm{prox}_{\mu f}) \circ L.$$

Proximity operator: properties

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $\operatorname{ran} L = \mathcal{H}$. Then

$$\partial(f \circ L) = L^* \, \partial f \, L.$$

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $LL^* = \mu \mathrm{Id}$ where $\mu \in \,]0, +\infty[$. Then

$$\mathrm{prox}_{f \circ L} = \mathrm{Id} - \mu^{-1} L^* \circ (\mathrm{Id} - \mathrm{prox}_{\mu f}) \circ L.$$

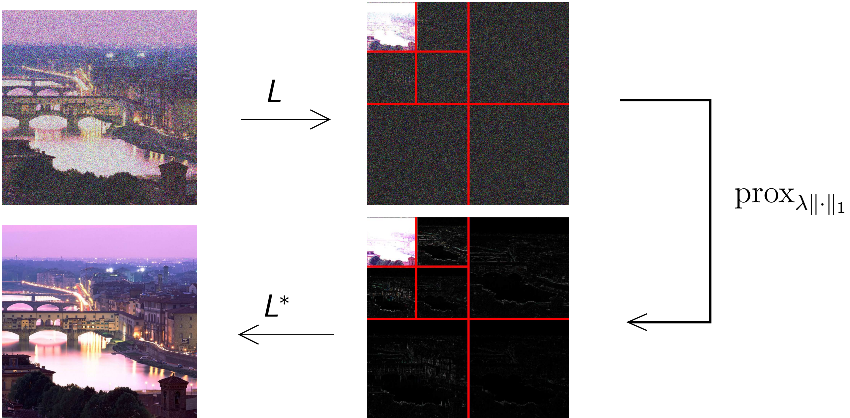
Remark :

Useful property for data fidelity terms involving a neg-log-likelihood f and a synthesis tight frame operator L.

Proximity operator: properties

Particular case : $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ unitary, $\mathrm{prox}_{f \circ L} = L^* \mathrm{prox}_f L$.

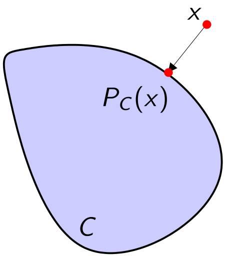
- Illustration: denoising using an ℓ_1 penalty on the coefficients resulting from an orthogonal wavelet transform L.



Proximity operator: examples

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

$$(\forall x \in \mathcal{H}) \qquad \mathrm{prox}_{\iota_C}(x) = \underset{y \in C}{\mathrm{argmin}} \frac{1}{2}\|y - x\|^2 = P_C(x).$$



Proximity operator: examples

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

$$(\forall x \in \mathcal{H}) \qquad \mathrm{prox}_{\iota_C}(x) = \underset{y \in C}{\mathrm{argmin}} \frac{1}{2}\|y - x\|^2 = P_C(x).$$

Remark :

- $p = P_C(x) \Leftrightarrow x - p \in \partial\iota_C(p) = N_C(p)$
  $\Leftrightarrow (\forall y \in C)\ \langle y - p \mid x - p\rangle \leq 0$.
  Particular case: if C is a vector space: $p = P_C(x) \Leftrightarrow x - p \in C^\perp$.
- ${}^\gamma\iota_C = (2\gamma)^{-1} d_C^2$ where d_C distance to the convex set C is defined by $d_C \colon x \mapsto \inf_{y\in C} \|y - x\| = \|x - P_C x\|$. We have then $\nabla d_C^2 = \nabla(\,{}^{\frac{1}{2}}\iota_C) = 2(\mathrm{Id} - P_C)$.

Proximity operator: examples

Quadratic function :
Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$, $\gamma \in \,]0, +\infty[$ and $z \in \mathcal{G}$.

$$f = \gamma \left\| L \cdot - z \right\|^2 / 2 \quad \Rightarrow \quad \mathrm{prox}_f = (\mathrm{Id} + \gamma L^* L)^{-1}(\cdot + \gamma L^* z).$$

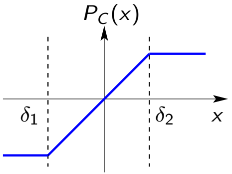
Proximity operator: examples

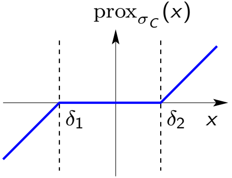
Support function :
Let $\mathcal{H}$ be a Hilbert space and $C \subset \mathcal{H}$ be nonempty closed convex.

$$(\forall x \in \mathcal{H}) \qquad \mathrm{prox}_{\sigma_C} = \mathrm{Id} - P_C.$$

Soft-thresholding : $\mathcal{H} = \mathbb{R}$, $\delta_1 = \inf C$ and $\delta_2 = \sup C$. For every $x \in \mathbb{R}$,

$$\sigma_C(x) = \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases} \Rightarrow \ \mathrm{prox}_{\sigma_C}(x) = \mathrm{soft}_C(x) = \begin{cases} x - \delta_1 & \text{if } x < \delta_1 \\ 0 & \text{if } x \in C \\ x - \delta_2 & \text{if } x > \delta_2. \end{cases}$$





Part 3: Search for a zero

1. Zeros of a (maximally) monotone operator

2. Fixed points

3. Convergence
   - Definition
   - Fejér monotonicity
   - Demiclosedness principle

4. Algorithms
   - Krasnosel'skii Mann
   - Douglas-Rachford
   - PPXA
   - *Forward-Backward*
   - *Forward-Backward-Forward*

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator.
The set of zeros of A, denoted $\operatorname{zer} A$, is

$$\operatorname{zer} A = \{x \in \mathcal{H} \,|\, 0 \in Ax\}.$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
$\operatorname{zer} A$ is closed and convex.

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator.
The set of zeros of A, denoted $\operatorname{zer} A$, is

$$\operatorname{zer} A = \{x \in \mathcal{H} \mid 0 \in Ax\}.$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
If one of the following assumptions is satisfied

- A is surjective
- $\operatorname{dom} A$ is bounded,

then $\operatorname{zer} A \neq \varnothing$.

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space.
$A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is **strictly monotone** if

$$\big(\forall (x_1, u_1) \in \operatorname{gra} A\big)\big(\forall (x_2, u_2) \in \operatorname{gra} A\big) \quad x_1 \neq x_2 \Rightarrow \langle u_1 - u_2 \mid x_1 - x_2 \rangle > 0.$$

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space.
$A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is strictly monotone if

$$\big(\forall (x_1, u_1) \in \mathrm{gra} A\big)\big(\forall (x_2, u_2) \in \mathrm{gra} A\big) \quad x_1 \neq x_2 \Rightarrow \langle u_1 - u_2 \mid x_1 - x_2 \rangle > 0.$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a strictly monotone operator.
$\mathrm{zer}\, A$ is at most a singleton .

Proof:

If $(x_1, x_2) \in (\mathrm{zer}\, A)^2$ and $x_1 \neq x_2$ then $0 = \langle x_1 - x_2 \mid 0 - 0 \rangle \leq 0$!

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space and $\beta \in \,]0,+\infty[$.
$A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is **β-strongly monotone** if

$$\big(\forall(x_1,u_1) \in \operatorname{gra} A\big)\big(\forall(x_2,u_2) \in \operatorname{gra} A\big) \quad \langle u_1 - u_2 \mid x_1 - x_2\rangle \geq \beta\|x_1 - x_2\|^2.$$

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space and $\beta \in \,]0, +\infty[$.
$A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is **β-strongly monotone** if

$$\big(\forall (x_1, u_1) \in \operatorname{gra} A\big)\big(\forall (x_2, u_2) \in \operatorname{gra} A\big) \quad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq \beta \|x_1 - x_2\|^2.$$

Let $\mathcal{H}$ be a Hilbert space and $\beta \in \,]0, +\infty[$.
$A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is β-strongly monotone $\Leftrightarrow$ $A^{-1}\colon \mathcal{H} \to \mathcal{H}$ is β-cocoercive.

Let $\mathcal{H}$ be a Hilbert space and $\beta \in \,]0, +\infty[$.
$A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is β-strongly monotone $\Rightarrow$ A is strictly monotone.

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space and $\beta \in\ ]0,+\infty[$.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
A β-strongly monotone $\Rightarrow$ $\operatorname{zer} A$ is a singleton.

Monotone operator: zeros

Let $\mathcal{H}$ be a Hilbert space and $\beta \in]0, +\infty[$.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
A β-strongly monotone $\Rightarrow$ $\operatorname{zer} A$ is a singleton.

Proof: For all $(x_1, u_1) \in \operatorname{gra} A$ et $(x_2, u_2) \in \operatorname{gra} A$,

$$\begin{aligned}
\|x_1 - x_2\| \|u_1\| &\geq \langle x_1 - x_2 \mid u_1 \rangle \\
&= \langle x_1 - x_2 \mid u_1 - u_2 \rangle + \langle x_1 - x_2 \mid u_2 \rangle \\
&\geq \frac{\beta}{2} \|x_1 - x_2\|^2 + \langle x_1 - x_2 \mid u_2 \rangle
\end{aligned}$$

Setting x_2 and u_2, if $\|x_1\| \to +\infty$, we can deduce that $\|u_1\| \to +\infty$.
This proves that $\lim_{\|x\| \to +\infty} \inf \|Ax\| = +\infty$ and thus $\operatorname{zer} A \neq \varnothing$.
Moreover, A being strictly monotone, there is only one element in $\operatorname{zer} A$.

Monotone operator: zeros and minimizer

Let $\mathcal{H}$ be a Hilbert space.
Let $f\colon \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
f is strictly convex $\Rightarrow \partial f$ is strictly monotone.

Monotone operator: zeros and minimizer

Let $\mathcal{H}$ be a Hilbert space.
Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ be a proper function.
f is strictly convex $\Rightarrow$ ∂f is strictly monotone.

<u>Proof</u>: Let $(x_1,u_1) \in \operatorname{gra}\partial f$ and $(x_2,u_2) \in \operatorname{gra}\partial f$ such that $x_1 \neq x_2$. We have, for all $\alpha \in]0,1[$,

$$\begin{aligned} f(x_1) + \langle u_1 \mid \alpha(x_2 - x_1)\rangle &\leq f(x_1 + \alpha(x_2 - x_1)) \\ &< (1-\alpha)f(x_1) + \alpha f(x_2) \\ \Rightarrow \langle u_1 \mid x_2 - x_1\rangle < f(x_2) - f(x_1). \end{aligned}$$

Symmetrically, $\langle u_2 \mid x_1 - x_2\rangle < f(x_1) - f(x_2)$ and, par summation,

$$\langle u_1 - u_2 \mid x_2 - x_1\rangle < 0$$

Monotone operator: zeros and minimizer

Let $\mathcal{H}$ be a Hilbert space.
Let $f\colon \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
f is strictly convex $\Rightarrow \partial f$ is strictly monotone.

Let $\mathcal{H}$ be a Hilbert space.
Let $f\colon \mathcal{H} \to]-\infty, +\infty]$ be a proper function.
f is strictly convex $\Rightarrow f$ has at most one minimizer.

Monotone operator: zeros and minimizer

Let $\mathcal{H}$ be a Hilbert space and $\beta \in \,]0, +\infty[$.
Let $f \colon \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
f is β-strongly convex $\Rightarrow \partial f$ is β-strongly monotone.

Monotone operator: zeros and minimizer

Let $\mathcal{H}$ be a Hilbert space and $\beta \in \,]0, +\infty[$.
Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a proper function.
f is β-strongly convex $\Rightarrow$ ∂f is β-strongly monotone.

Proof: f β-strongly convex if $f = h + \beta\|\cdot\|^2/2$ where h is convex. Let $(x_1, u_1) \in \operatorname{gra}\partial f$ and $(x_2, u_2) \in \operatorname{gra}\partial f$. We have

$$u_1 \in \partial f(x_1) \Leftrightarrow u_1 - \beta x_1 \in \partial h(x_1).$$

Moreover,

$$\begin{aligned} f(x_2) = h(x_2) + \frac{\beta}{2}\|x_2\|^2 &\geq h(x_1) + \langle u_1 - \beta x_1 \mid x_2 - x_1 \rangle + \frac{\beta}{2}\|x_2\|^2 \\ &= f(x_1) + \langle u_1 \mid x_2 - x_1 \rangle + \frac{\beta}{2}\|x_1 - x_2\|^2. \end{aligned}$$

Symmetrically, $f(x_1) \geq f(x_2) + \langle u_2 \mid x_1 - x_2 \rangle + \frac{\beta}{2}\|x_2 - x_1\|^2$.
Consequently, by summation,

$$0 \geq \langle u_1 - u_2 \mid x_2 - x_1 \rangle + \beta\|x_1 - x_2\|^2.$$

Monotone operator: zeros and minimizer

Let $\mathcal{H}$ be a Hilbert space and $\beta \in\ ]0,+\infty[$.
Let $f\colon \mathcal{H} \to\ ]-\infty,+\infty]$ be a proper function.
f is β-strongly convex $\Rightarrow$ ∂f is β-strongly monotone.

Let $\mathcal{H}$ be a Hilbert space and $\beta \in\ ]0,+\infty[$.
Let $f \in \Gamma_0(\mathcal{H})$.
f is β-strongly convex $\Rightarrow$ f has a unique minimizer.

Fixed point algorithms: zeros and fixed points

Let $\mathcal{H}$ be a Hilbert space. Let $B : \mathcal{H} \to 2^{\mathcal{H}}$.
The set of fixed points of B is denoted by

$$\mathrm{Fix} B = \{x \in \mathcal{H} \,|\, x \in Bx\}$$

Fixed point algorithms: zeros and fixed points

Let $\mathcal{H}$ be a Hilbert space. Let $B : \mathcal{H} \to 2^{\mathcal{H}}$.
The set of fixed points of B is denoted by

$$\mathrm{Fix} B = \{x \in \mathcal{H} \,|\, x \in Bx\}$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator and let $\gamma \in \,]0, +\infty[$.
Then

$$\mathrm{Fix} J_{\gamma A} = \mathrm{zer}\, A$$

Fixed point algorithms: zeros and fixed points

Let $\mathcal{H}$ be a Hilbert space. Let $B : \mathcal{H} \to 2^{\mathcal{H}}$.
The set of fixed points of B is denoted by

$$\mathrm{Fix} B = \{x \in \mathcal{H} \,|\, x \in Bx\}$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator and let $\gamma \in \,]0, +\infty[$.
Then

$$\mathrm{Fix} J_{\gamma A} = \mathrm{zer}\, A$$

<u>Proof</u> : $(\forall x \in \mathcal{H}) \qquad 0 \in Ax \quad \Leftrightarrow \quad \dots$

$$\Leftrightarrow \quad \dots$$

$$\Leftrightarrow \quad \dots$$

$$\Leftrightarrow \quad x = J_{\gamma A} x$$

Fixed point algorithms: zeros and fixed points

Let $\mathcal{H}$ be a Hilbert space. Let $B : \mathcal{H} \to 2^{\mathcal{H}}$.
The set of fixed points of B is denoted by

$$\mathrm{Fix} B = \{x \in \mathcal{H} \,|\, x \in Bx\}$$

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone operator and let $\gamma \in \,]0, +\infty[$.
Then

$$\mathrm{Fix} J_{\gamma A} = \mathrm{zer}\, A$$

Proof :

$$\begin{aligned} (\forall x \in \mathcal{H}) \qquad 0 \in Ax \quad &\Leftrightarrow \quad 0 \in \gamma Ax \\ &\Leftrightarrow \quad x \in (\mathrm{Id} + \gamma A)x \\ &\Leftrightarrow \quad x \in (\mathrm{Id} + \gamma A)^{-1}x \\ &\Leftrightarrow \quad x = J_{\gamma A}x \end{aligned}$$

Fixed point algorithm: convergence

Let $\mathcal{H}$ be a Hilbert space.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and $\widehat{x} \in \mathcal{H}$.

- $(x_n)_{n\in\mathbb{N}}$ converges strongly to $\widehat{x}$ if

$$\lim_{n\to+\infty} \|x_n - \widehat{x}\| = 0.$$

  It is denoted by $x_n \to \widehat{x}$.

- $(x_n)_{n\in\mathbb{N}}$ converges weakly to $\widehat{x}$ if

$$(\forall y \in \mathcal{H}) \qquad \lim_{n\to+\infty} \langle y \mid x_n - \widehat{x}\rangle = 0.$$

  It is denoted by $x_n \rightharpoonup \widehat{x}$.

Remark: In a finite dimensional Hilbert space, strong and weak convergences are equivalent.

Fixed point algorithm: convergence

Let $\mathcal{H}$ be a Hilbert space.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of $\mathcal{H}$.
$(x_n)_{n\in\mathbb{N}}$ converges weakly if and only if

- $(x_n)_{n\in\mathbb{N}}$ is bounded

  and
- $(x_n)_{n\in\mathbb{N}}$ posseses at most one sequential cluster point in the weak topology.

- $\widehat{x}$ is a sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ in the weak topology if there exists a sub-sequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ that converges weakly to $\widehat{x}$.

Fixed point algorithm: convergence

Let $\mathcal{H}$ be a Hilbert space.
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of $\mathcal{H}$.
$(x_n)_{n\in\mathbb{N}}$ converges weakly if and only if

- $(x_n)_{n\in\mathbb{N}}$ is bounded

  and

- $(x_n)_{n\in\mathbb{N}}$ posseses at most one sequential cluster point in the weak topology.

Illustration:

x_0	x_1	x_2	x_3	x_4	x_5	...
1	-1	1	-1	1	-1	...

$\rightarrow$ $(x_n)_{n\in\mathbb{N}}$ is bounded but it has 2 sequential cluster points: -1 and 1.

$\rightarrow$ $(x_n)_{n\in\mathbb{N}}$ does not converge.

Fixed point algorithm: Fejér-monotone sequence

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

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

Fixed point algorithm: Fejér-monotone sequence

Fejér-monotone convergence

Let $\mathcal{H}$ be a Hilbert space and let D be a nonempty subset of $\mathcal{H}$.
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 sequential cluster point of $(x_n)_{n\in\mathbb{N}}$ in the weak topology lies in D.

Fixed point algorithm: Fejér-monotone sequence

Proof:

If $(\|x_n - x\|)_{n\in\mathbb{N}}$ converges then $(\|x_n - x\|)_{n\in\mathbb{N}}$ and thus $(x_n)_{n\in\mathbb{N}}$ are bounded.

We assume that $(x_{n_k})_{k\in\mathbb{N}}$ and $(x_{n_\ell})_{\ell\in\mathbb{N}}$ are such that $x_{n_k} \rightharpoonup \widehat{x}$ and $x_{n_\ell} \rightharpoonup \widehat{x}'$ where $(\widehat{x}, \widehat{x}') \in D^2$. For all $n \in \mathbb{N}$,

$$2\left\langle x_n \mid \widehat{x}' - \widehat{x} \right\rangle = \|x_n - \widehat{x}\|^2 - \|x_n - \widehat{x}'\|^2 - \|\widehat{x}\|^2 + \|\widehat{x}'\|^2.$$

Because $(\|x_n - \widehat{x}\|)_{n\in\mathbb{N}}$ and $(\|x_n - \widehat{x}'\|)_{n\in\mathbb{N}}$ converge, there exists $\alpha \in \mathbb{R}$ such that $\langle x_n \mid \widehat{x}' - \widehat{x}\rangle \to \alpha$ and thus $\langle x_{n_k} \mid \widehat{x}' - \widehat{x}\rangle \to \langle \widehat{x} \mid \widehat{x}' - \widehat{x}\rangle = \alpha$. Similarly, $\langle \widehat{x}' \mid \widehat{x}' - \widehat{x}\rangle = \alpha$. Consequently $\|\widehat{x}' - \widehat{x}\|^2 = 0 \Rightarrow \widehat{x} = \widehat{x}'$.

Fixed point algorithm: Fejér-monotone sequence

Demiclosedness principle

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$.
Let $T\colon C \to \mathcal{H}$ be a nonexpansive operator.
If $(x_n)_{n\in\mathbb{N}}$ is a sequence in C that converges weakly to $\widehat{x}$ and if $Tx_n - x_n \to 0$ then

$$\widehat{x} \in \operatorname{Fix} T.$$

Fixed point algorithm: Fejér-monotone sequence

Demiclosedness principle

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$. Let $T\colon C \to \mathcal{H}$ be a nonexpansive operator. If $(x_n)_{n\in\mathbb{N}}$ is a sequence in C that converges weakly to $\widehat{x}$ and if $Tx_n - x_n \to 0$ then

$$\widehat{x} \in \operatorname{Fix} T.$$

Proof:
We have $\widehat{x} - P_C\widehat{x} \in N_C\big(P_C\widehat{x}\big)$.
Because $(\forall n \in \mathbb{N})\ x_n \in C$, we have

$$\langle x_n - P_C\widehat{x} \mid \widehat{x} - P_C\widehat{x}\rangle \leq 0.$$

Using $x_n \rightharpoonup \widehat{x}$, we deduce that $\|\widehat{x} - P_C\widehat{x}\|^2 = 0$, so $\widehat{x} = P_C(\widehat{x}) \in C$.
Therefore, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in C that converges weakly to $\widehat{x} \in C$.

Fixed point algorithm: Fejér-monotone sequence

Demiclosedness principle

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$. Let $T\colon C \to \mathcal{H}$ be a nonexpansive operator. If $(x_n)_{n\in\mathbb{N}}$ is a sequence in C that converges weakly to $\widehat{x}$ and if $Tx_n - x_n \to 0$ then

$$\widehat{x} \in \operatorname{Fix} T.$$

Proof:

$x_n \rightharpoonup \widehat{x} \Rightarrow \widehat{x} \in C$ and $T\widehat{x}$ is defined. For all $n \in \mathbb{N}$,

$$\|x_n - T\widehat{x}\|^2 = \|x_n - \widehat{x}\|^2 + \|\widehat{x} - T\widehat{x}\|^2 + 2\left\langle x_n - \widehat{x} \mid \widehat{x} - T\widehat{x}\right\rangle$$

$$\|x_n - T\widehat{x}\|^2 = \|x_n - Tx_n\|^2 + \|Tx_n - T\widehat{x}\|^2 + 2\left\langle x_n - Tx_n \mid Tx_n - T\widehat{x}\right\rangle$$

$$\Rightarrow \|\widehat{x} - T\widehat{x}\|^2 = \|x_n - Tx_n\|^2 + \|Tx_n - T\widehat{x}\|^2 - \|x_n - \widehat{x}\|^2 + 2\left\langle x_n - Tx_n \mid Tx_n - T\widehat{x}\right\rangle - 2\left\langle x_n - \widehat{x} \mid \widehat{x} - T\widehat{x}\right\rangle$$

Fixed point algorithm: Fejér-monotone sequence

Demiclosedness principle

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$. Let $T: C \to \mathcal{H}$ be a nonexpansive operator.
If $(x_n)_{n\in\mathbb{N}}$ is a sequence in C that converges weakly to $\widehat{x}$ and if $Tx_n - x_n \to 0$ then

$$\widehat{x} \in \operatorname{Fix} T.$$

Proof:

$$\begin{aligned}\|\widehat{x} - T\widehat{x}\|^2 =& \|x_n - Tx_n\|^2 + \|Tx_n - T\widehat{x}\|^2 - \|x_n - \widehat{x}\|^2 \\ &+ 2\langle x_n - Tx_n \mid Tx_n - T\widehat{x}\rangle - 2\langle x_n - \widehat{x} \mid \widehat{x} - T\widehat{x}\rangle .\end{aligned}$$

T being non expansive and, using Cauchy-Schwarz inequality,

$$\begin{aligned}\|\widehat{x} - T\widehat{x}\|^2 &\le \|x_n - Tx_n\|^2 + 2\|x_n - Tx_n\|\|Tx_n - T\widehat{x}\| - 2\langle x_n - \widehat{x} \mid \widehat{x} - T\widehat{x}\rangle \\ &\le \|x_n - Tx_n\|^2 + 2\|x_n - Tx_n\|\|x_n - \widehat{x}\| - 2\langle x_n - \widehat{x} \mid \widehat{x} - T\widehat{x}\rangle .\end{aligned}$$

$x_n \rightharpoonup \widehat{x} \Rightarrow (x_n)_{n\in\mathbb{N}}$ bounded. Taking the limit, the result is proved .

Fixed point algorithm: Fejér-monotone sequences

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty subset of $\mathcal{H}$.
Let $T\colon C \to C$ be a nonexpansive operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

Fixed point algorithm: Fejér-monotone sequences

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty subset of $\mathcal{H}$.
Let $T\colon C \to C$ be a nonexpansive operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

If $x_n - Tx_n \to 0$,

Fixed point algorithm: Fejér-monotone sequences

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty subset of $\mathcal{H}$.
Let $T\colon C \to C$ be a nonexpansive operator such that $\operatorname{Fix} T \neq \varnothing$.
Let $x_0 \in C$,

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

If $x_n - Tx_n \to 0$, then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{Fix} T$.

Proof : $(x_n)_{n\in\mathbb{N}}$ Fejér-monotone with respect to $\operatorname{Fix} T$ + demiclosedness principle.

Fixed point algorithm: proximal point algorithm

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

Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{A} x_n,$$

Fixed point algorithm: proximal point algorithm

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

Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator

Let $\gamma \in \,]0, +\infty[$

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma A} x_n,$$

Fixed point algorithm: proximal point algorithm

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $\operatorname{zer} A \neq \varnothing$.
Let $\gamma \in \left]0, +\infty\right[$ et $x_0 \in \mathcal{H}$.
If

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma A} x_n,$$

then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer} A$.

Proof: $J_{\gamma A}\colon \mathcal{H} \to \mathcal{H}$ is nonexpansive and $x_n - J_{\gamma A} x_n \to 0$.

Fixed point algorithm: proximal point algorithm

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator such that $\operatorname{zer} A \neq \varnothing$.
Let $(\gamma_n)_{n\in\mathbb{N}}$ a sequence in $]0,+\infty[$ such that $\sum_{n=0}^{+\infty} \gamma_n^2 = +\infty$ and $x_0 \in \mathcal{H}$.
If

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma_n A} x_n,$$

then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer} A$.

Proof: ... more challenging.

Optimization method: proximal point algorithm

Let $\mathcal{H}$ be a Hilbert space.
Let $f \in \Gamma_0(\mathcal{H})$ such that $\operatorname{Argmin} f \neq \varnothing$.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}_{\gamma_n f} x_n$$

Optimization method: proximal point algorithm

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

Let $f \in \Gamma_0(\mathcal{H})$ such that $\operatorname{Argmin} f \neq \varnothing$.

Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $]0,+\infty[$ such that $\sum_{n=0}^{+\infty} \gamma_n^2 = +\infty$ and $x_0 \in \mathcal{H}$.

If

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{prox}_{\gamma_n f} x_n$$

then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a (global) minimizer of f.

Fixed point algorithm: Fejér-monotone sequence

Krasnosel'skii-Mann algorithm

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$.
Let $T\colon C \to C$ be a nonexpansive operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Fixed point algorithm: Fejér-monotone sequence

Krasnosel'skii-Mann algorithm

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$.
Let $T \colon C \to C$ be a nonexpansive operator
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Fixed point algorithm: Fejér-monotone sequence

Krasnosel'skii-Mann algorithm

Let $\mathcal{H}$ be a Hilbert space and C be a nonempty closed convex subset of $\mathcal{H}$.
Let $T\colon C \to C$ be a nonexpansive operator such that $\operatorname{Fix} T \neq \varnothing$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(1-\lambda_n) = +\infty$.
Let $x_0 \in C$ and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Then, $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{Fix} T$

Proof:

- $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to $\operatorname{Fix} T$.
- $(Tx_n - x_n)_{n\in\mathbb{N}}$ converges strongly to 0.

Fixed point algorithm: Douglas-Rachford

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

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be two maximally monotone operators.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n (z_n - y_n). \end{cases}$$

Fixed point algorithm: Douglas-Rachford

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

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be two maximally monotone operators.

Let $\gamma \in \,]0,+\infty[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,2]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

Fixed point algorithm: Douglas-Rachford

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

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be two maximally monotone operators.

Let $\gamma \in \,]0, +\infty[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0, 2]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$.

We assume that $\operatorname{zer}(A + B) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

The following properties are satisfied:

- $x_n \to \widehat{x}$
- $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ converge to $J_{\gamma B}\widehat{x} \in \operatorname{zer}(A + B)$.

Fixed point algorithm: Douglas-Rachford

Proof: We set $T = R_{\gamma A} R_{\gamma B}$.

1. T is nonexpansive.

2. $\operatorname{zer}(A+B) = J_{\gamma B}(\operatorname{Fix} T)$.

Fixed point algorithm: Douglas-Rachford

Proof: We set $T = R_{\gamma A} R_{\gamma B}$.

1. T is nonexpansive.

2. $\operatorname{zer}(A+B) = J_{\gamma B}(\operatorname{Fix} T)$.

   Proof: Let $x \in \mathcal{H}$.

$$
\begin{aligned}
0 \in \gamma(Ax + Bx) &\Leftrightarrow (\exists y \in \mathcal{H}) \;\; x - y \in \gamma Ax \text{ et } y - x \in \gamma Bx \\
&\Leftrightarrow (\exists y \in \mathcal{H}) \;\; 2x - y \in (\mathrm{Id} + \gamma A)x \text{ et } x = J_{\gamma B} y \\
&\Leftrightarrow (\exists y \in \mathcal{H}) \;\; x = J_{\gamma A}(R_{\gamma B} y) \text{ et } x = J_{\gamma B} y \\
&\Leftrightarrow (\exists y \in \mathcal{H}) \;\; R_{\gamma A}(R_{\gamma B} y) = 2x - R_{\gamma B} y = y \\
&\qquad\qquad \text{et } x = J_{\gamma B} y \\
&\Leftrightarrow (\exists y \in \mathrm{Fix} R_{\gamma A} R_{\gamma B}) \;\; x = J_{\gamma B} y.
\end{aligned}
$$

Fixed point algorithm: Douglas-Rachford

Proof: We set $T = R_{\gamma A} R_{\gamma B}$.

1. T is nonexpansive.

2. $\operatorname{zer}(A+B) = J_{\gamma B}(\operatorname{Fix} T)$.

3. $\varnothing \neq \operatorname{zer}(A+B) = J_{\gamma B}(\operatorname{Fix} T) \Rightarrow \operatorname{Fix} T \neq \varnothing$.

4. We have

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \big(J_{\gamma A}(2 J_{\gamma B} x_n - x_n) - J_{\gamma B} x_n \big)$$
$$= x_n + \frac{\lambda_n}{2} \big(T x_n - x_n \big).$$

$\Rightarrow$ Krasnosel'skii-Mann algorithm with relaxation steps $(\lambda_n/2)_{n\in\mathbb{N}}$.

5. Invoke continuity of $J_{\gamma B}$.

Fixed point algorithm: Douglas-Rachford

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ be two maximally monotone operators.
Let $\gamma \in \,]0,+\infty[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,2]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$.
We assume that $\operatorname{zer}(A+B) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

The following properties are satisfied:

- $x_n \rightharpoonup \widehat{x}$
- $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ converge weakly to $J_{\gamma B}\widehat{x} \in \operatorname{zer}(A+B)$.

Optimization algorithm: Douglas-Rachford

Let $\mathcal{H}$ be a Hilbert space.
Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \mathrm{prox}_{\gamma g} x_n \\ z_n = \mathrm{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

Optimization algorithm: Douglas-Rachford

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

Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.

Let $\gamma \in]0, +\infty[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0, 2]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$.

We assume that $\mathrm{zer}\,(\partial f + \partial g) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \mathrm{prox}_{\gamma g} x_n \\ z_n = \mathrm{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

The following properties are satisfied:

- $x_n \rightharpoonup \widehat{x}$
- $z_n - y_n \to 0$, $y_n \rightharpoonup \widehat{y}$, $z_n \rightharpoonup \widehat{y}$ where $\widehat{y} = \mathrm{prox}_{\gamma g}\widehat{x} \in \mathrm{Argmin}(f + g)$.

Optimization algorithm: Douglas-Rachford

Image restoration : $z = A\overline{x} + n$

- $\overline{x} \in \mathbb{R}^N$: original image (unknown),
- $A \in \mathbb{R}^{N \times N}$: blur operator,
- $n \in \mathbb{R}^N$: additive noise (white zero-mean Gaussian),
- $z \in \mathbb{R}^N$: observation = blur + noise

Degraded image z

$\Rightarrow$ Find an image $\widehat{x}$ close to $\overline{x}$ using z

Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta\|Wx\|_1 \quad \text{with} \quad \eta \in\ ]0,+\infty[\quad \text{and} \quad W \in \mathbb{R}^{N\times N}$$

- Douglas-Rachford algorithm with $g = \|A\cdot - z\|_2^2$ and $f = \eta\|W\cdot\|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\mathrm{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in \,]0, +\infty[\quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

- Douglas-Rachford algorithm with $g = \|A \cdot - z\|_2^2$ and $f = \eta \|W \cdot \|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \mathrm{prox}_{\gamma g} x_n \quad \rightarrow \quad \text{Closed form} \\ z_n = \mathrm{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n (z_n - y_n). \end{cases}$$

Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in]0, +\infty[\quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

- Douglas-Rachford algorithm with $g = \|A \cdot - z\|_2^2$ and $f = \eta \| W \cdot \|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n (z_n - y_n). \end{cases} \quad \rightarrow \quad \text{Closed form}$$

$$\boxed{\operatorname{prox}_{\gamma\eta|\cdot|_1} = \text{soft}_{[-\gamma\eta, \gamma\eta]}}$$

Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta\|Wx\|_1 \quad \text{with} \quad \eta \in]0,+\infty[\quad \text{and} \quad W \in \mathbb{R}^{N\times N}$$

- Douglas-Rachford algorithm with $g = \|A\cdot - z\|_2^2$ and $f = \eta\|W\cdot\|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

$\rightarrow$ Closed form if $W^{-1} = W^*$ (e.g. wavelet transform)

$$\boxed{\operatorname{prox}_{f\circ W} = W^* \operatorname{prox}_f(W\cdot)}$$

Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\mathrm{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in]0, +\infty[\quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

- Douglas-Rachford algorithm

Degraded image z

Restored image $\widehat{x}$ [DR – DWT]

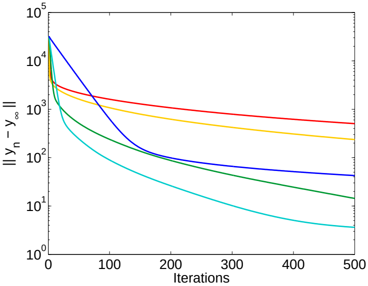
Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta\|Wx\|_1 \quad \text{with} \quad \eta \in \,]0,+\infty[\quad \text{and} \quad W \in \mathbb{R}^{N\times N}$$

- Douglas-Rachford algorithm

$$\begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$



$$\gamma = \{50, 10^2, 5.10^2, 10^3, 5.10^3\}$$

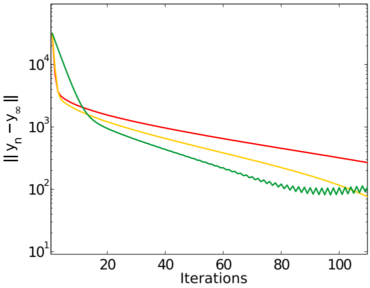
Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta\|Wx\|_1 \quad \text{with} \quad \eta \in \,]0,+\infty[\quad \text{and} \quad W \in \mathbb{R}^{N\times N}$$

- Douglas-Rachford algorithm

$$\begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$



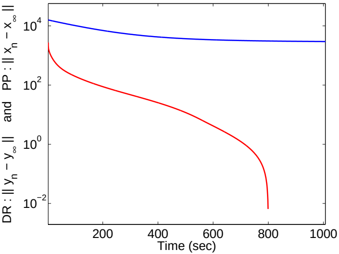
$$\lambda_n \equiv \{1, 1.5, 2.1\}$$

Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in]0, +\infty[\quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

- Douglas-Rachford algorithm



→ DR (red)

→ Proximal point algo. (blue)

Optimization algorithm: Douglas-Rachford

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $g \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $\operatorname{ran} L$ is closed and L^*L is an isomorphism.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ c_n = (L^*L)^{-1} L^* y_n \\ x_{n+1} = x_n + \lambda_n \big(L(2c_n - v_n) - y_n \big) \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

Optimization algorithm: Douglas-Rachford

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $g \in \Gamma_0(\mathcal{H})$ and $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ such that $\operatorname{ran} L$ is closed and L^*L is an isomorphism.
Let $\gamma \in \,]0, +\infty[$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0, 2]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$.
We assume that $\operatorname{zer}(L^* \circ \partial g \circ L) \neq \varnothing$. Let $x_0 \in \mathcal{H}$, $v_0 = (L^*L)^{-1}L^*x_0$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ c_n = (L^*L)^{-1}L^*y_n \\ x_{n+1} = x_n + \lambda_n\big(L(2c_n - v_n) - y_n\big) \\ v_{n+1} = v_n + \lambda_n(c_n - v_n). \end{cases}$$

We have then:
$v_n \rightharpoonup \widehat{v}$ where $\widehat{v} \in \operatorname{Argmin}(g \circ L)$.

Optimization algorithm: Douglas-Rachford

Sketch of proof:

$$\underset{v\in\mathcal{G}}{\text{minimize}}\ g(Lv) \quad\Leftrightarrow\quad \underset{x\in\mathcal{H}}{\text{minimize}}\ \iota_E(x)+g(x)$$

where $E=\mathrm{ran}\,L$.
We apply Douglas-Rachford algorithm with $f=\iota_E$ and we set

$$(\forall n\in\mathbb{N})\qquad P_E y_n = Lc_n \ \text{ et } \ P_E x_n = Lv_n$$

where $c_n=\underset{c\in\mathcal{H}}{\text{argmin}}\ \|y_n-Lc\|^2=(L^*L)^{-1}L^*y_n$.

Optimization algorithm: Douglas-Rachford

Particular case of Douglas-Rachford algorithm: PPXA+
$\mathcal{H} = \mathcal{H}_1 \times \cdots \mathcal{H}_m$ where $\mathcal{H}_1, \ldots, \mathcal{H}_m$ Hilbert spaces

$$
(\forall n \in \mathbb{N}) \qquad
\begin{cases}
y_{n,i} = \mathrm{prox}_{\gamma g_i} x_{n,i}, \quad i \in \{1, \ldots, m\} \\
c_n = \left(\sum_{i=1}^m L_i^* L_i\right)^{-1} \sum_{i=1}^m L_i^* y_{n,i} \\
x_{n+1,i} = x_{n,i} + \lambda_n\big(L_i (2c_n - v_n) - y_{n,i}\big), \quad i \in \{1, \ldots, m\} \\
v_{n+1} = v_n + \lambda_n(c_n - v_n).
\end{cases}
$$

Optimization algorithm: Douglas-Rachford

Particular case of Douglas-Rachford algorithm: PPXA+

$\mathcal{H} = \mathcal{H}_1 \times \cdots \mathcal{H}_m$ where $\mathcal{H}_1, \ldots, \mathcal{H}_m$ Hilbert spaces

$\big(\forall x = (x_1, \ldots, x_m) \in \mathcal{H}\big)\ g(x) = \sum_{i=1}^m g_i(x_i)$

where $(\forall i \in \{1, \ldots, m\}\ g_i \in \Gamma_0(\mathcal{H}_i)$

$L\colon v \mapsto (L_1 v, \ldots, L_m v)$ where $(\forall i \in \{1, \ldots, m\})\ L_i \in \mathcal{B}(\mathcal{G}, \mathcal{H}_i)$.

Let $(x_{0,i})_{1 \le i \le m} \in \mathcal{H}$, $v_0 = (\sum_{i=1}^m L_i^* L_i)^{-1} \sum_{i=1}^m L_i^* x_{0,i}$ and

$$(\forall n \in \mathbb{N}) \quad \left\{ \begin{aligned} & y_{n,i} = \mathrm{prox}_{\gamma g_i} x_{n,i}, \quad i \in \{1, \ldots, m\} \\ & c_n = (\textstyle\sum_{i=1}^m L_i^* L_i)^{-1} \sum_{i=1}^m L_i^* y_{n,i} \\ & x_{n+1,i} = x_{n,i} + \lambda_n \big(L_i (2c_n - v_n) - y_{n,i} \big), \quad i \in \{1, \ldots, m\} \\ & v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{aligned} \right.$$

We have then $v_n \rightharpoonup \widehat{v} \in \mathrm{Argmin} \sum_{i=1}^m g_i \circ L_i$.

Optimization algorithm: Douglas-Rachford

Particular case of Douglas-Rachford algorithm: PPXA

$\mathcal{H} = \mathcal{H}_1 \times \cdots \mathcal{H}_m$ where $\mathcal{H}_1, \ldots, \mathcal{H}_m$ Hilbert spaces

$(\forall x = (x_1, \ldots, x_m) \in \mathcal{H})\ g(x) = \sum_{i=1}^m g_i(x_i)$

where $(\forall i \in \{1, \ldots, m\}\ g_i \in \Gamma_0(\mathcal{H}_i)$

$L\colon v \mapsto (L_1 v, \ldots, L_m v)$ where $(\forall i \in \{1, \ldots, m\})\ L_i \in \mathcal{B}(\mathcal{G}, \mathcal{H}_i)$.

If $\mathcal{H}_1 = \ldots = \mathcal{H}_m$ and $L_1 = \ldots = L_m = \mathrm{Id} \Rightarrow$ Consensus trick

Let $(x_{0,i})_{1 \le i \le m} \in \mathcal{H}$, $v_0 = \frac{1}{m} \sum_{i=1}^m x_{0,i}$ et

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \mathrm{prox}_{\gamma g_i} x_{n,i}, \quad i \in \{1, \ldots, m\} \\ c_n = \frac{1}{m} \sum_{i=1}^m y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n \big(2c_n - v_n - y_{n,i}\big), \quad i \in \{1, \ldots, m\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

We have then $v_n \rightharpoonup \widehat{v} \in \mathrm{Argmin} \sum_{i=1}^m g_i$.

Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta\|[H^* \quad V^*]^* x\|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in \,]0,+\infty[\\ H, V \in \mathbb{R}^{N\times N} \\ C = [0,255]^N \end{cases}$$

- PPXA+ with $g_1 = \|A\cdot - z\|_2^2$ and $L_1 = \mathrm{Id}$
  $g_2 = \eta\|\cdot\|_{1,2}$ and $L_2 = [H^* \quad V^*]^*$
  $g_3 = \iota_C$ and $L_3 = \mathrm{Id}$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \operatorname{prox}_{\gamma g_i} x_{n,i},\, i \in \{1,2,3\} \\ c_n = \left(\sum_{i=1}^3 L_i^* L_i\right)^{-1} \sum_{i=1}^3 L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n\big(L_i(2c_n - v_n) - y_{n,i}\big),\ i \in \{1,2,3\} \\ v_{n+1} = v_n + \lambda_n(c_n - v_n). \end{cases}$$

Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[H^* \quad V^*]^* x\|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in \,]0, +\infty[\\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

- PPXA+ with $g_1 = \|A \cdot - z\|_2^2$ and $L_1 = \mathrm{Id}$
  $g_2 = \eta \| \cdot \|_{1,2}$ and $L_2 = [H^* \quad V^*]^*$
  $g_3 = \iota_C$ and $L_3 = \mathrm{Id}$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \operatorname{prox}_{\gamma g_i} x_{n,i}, \, i \in \{1, 2, 3\} & \rightarrow \quad \text{Closed form} \\ c_n = (\sum_{i=1}^3 L_i^* L_i)^{-1} \sum_{i=1}^3 L_i^* y_{n,i} & \rightarrow \quad \text{Closed form} \\ x_{n+1,i} = x_{n,i} + \lambda_n \big(L_i(2c_n - v_n) - y_{n,i} \big), \; i \in \{1, 2, 3\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[H^* \quad V^*]^* x\|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in]0, +\infty[\\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

- PPXA+

Degraded image z

Restored image $\widehat{x}$ [PPXA – TV]

Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[H^* \quad V^*]^* x\|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in \,]0, +\infty[\\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

- PPXA+

Degraded image z

Restored image $\widehat{x}$ [DR – DWT]

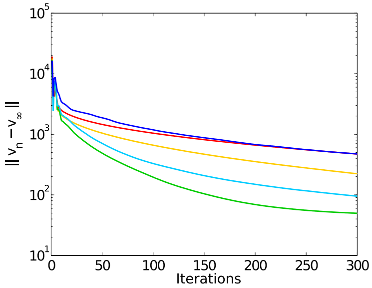
Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[H^* \quad V^*]^* x\|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in]0,+\infty[\\ H, V \in \mathbb{R}^{N\times N} \\ C = [0,255]^N \end{cases}$$

- PPXA+

$$\begin{cases} y_{n,i} = \operatorname{prox}_{\gamma g_i} x_{n,i}, \\ c_n = \left(\sum_{i=1}^3 L_i^* L_i\right)^{-1} \sum_{i=1}^2 L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n \big(L_i(2c_n - v_n) - y_{n,i}\big), \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$



$\gamma = \{5.10^2, 10^3, 5.10^3, 10^4, 5.10^4\}$

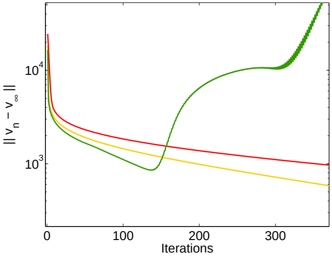
Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\widehat{x} \in \underset{x\in\mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[H^* \quad V^*]^* x\|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in\,]0, +\infty[\\ H, V \in \mathbb{R}^{N\times N} \\ C = [0, 255]^N \end{cases}$$

- PPXA+

$$\begin{cases} y_{n,i} = \operatorname{prox}_{\gamma g_i} x_{n,i}, \\ c_n = \left(\sum_{i=1}^3 L_i^* L_i\right)^{-1} \sum_{i=1}^2 L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n \big(L_i(2c_n - v_n) - y_{n,i}\big), \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$



$\lambda_n \equiv \{1, 1.8, 2.1\}$

Fixed point algorithm: α-averaged operator

Let $\mathcal{H}$ be a Hilbert space.
Let $T\colon \mathcal{H} \to \mathcal{H}$ be a α-averaged operator with $\alpha \in]0,1[$

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n (Tx_n - x_n).$$

Fixed point algorithm: α-averaged operator

Let $\mathcal{H}$ be a Hilbert space.
Let $T : \mathcal{H} \to \mathcal{H}$ be a α-averaged operator with $\alpha \in]0, 1[$

Let $(\lambda_n)_{n\in\mathbb{N}}$ a sequence in $[0, 1/\alpha]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(1 - \alpha\lambda_n) = +\infty$.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Fixed point algorithm: α-averaged operator

Let $\mathcal{H}$ be a Hilbert space.
Let $T : \mathcal{H} \to \mathcal{H}$ be a α-averaged operator with $\alpha \in]0, 1[$ such that $\operatorname{Fix} T \neq \varnothing$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ a sequence in $[0, 1/\alpha]$ such that $\sum_{n\in\mathbb{N}} \lambda_n(1 - \alpha\lambda_n) = +\infty$.
Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Then, $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{Fix} T$.

Proof : T α-average, there exists a contraction R such that $T = (1 - \alpha)\mathrm{Id} + \alpha R$. $\operatorname{Fix} R = \operatorname{Fix} T$. Let $(\forall n \in \mathbb{N})$ $\mu_n = \alpha\lambda_n \in [0, 1]$. The iterations becomes

$$\begin{aligned} x_{n+1} &= x_n + \lambda_n(Tx_n - x_n) \\ &= x_n + \mu_n(Rx_n - x_n). \end{aligned}$$

+ Use properties of Krasnosel'skii-Mann algorithm.

Fixed point algorithm: Forward-Backward

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

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.

Let $B\colon \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator with $\beta \in \,]0, +\infty[$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma B x_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma A} y_n - x_n). \end{cases}$$

Fixed point algorithm: Forward-Backward

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
Let $B\colon \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator with $\beta \in \,]0, +\infty[$.
Let $\gamma \in]0, 2\beta[$ and $\delta = \min\{1, \beta/\gamma\} + 1/2$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0, \delta[$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma B x_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma A} y_n - x_n). \end{cases}$$

Fixed point algorithm: Forward-Backward

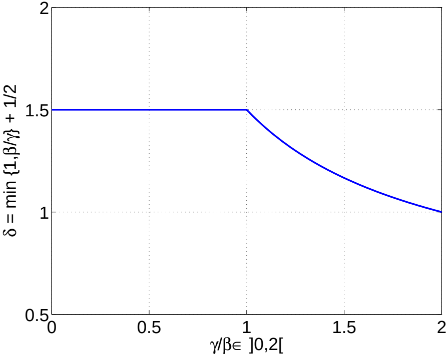
Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.
Let $B\colon \mathcal{H} \to \mathcal{H}$ be a β-cocoercive operator with $\beta \in \,]0,+\infty[$.
Let $\gamma \in]0,2\beta[$ and $\delta = \min\{1,\beta/\gamma\} + 1/2$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[0,\delta[$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$.
We assume that $\operatorname{zer}(A+B) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma B x_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma A} y_n - x_n). \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer}(A+B)$.

Fixed point algorithm: Forward-Backward

Let $\gamma \in]0, 2\beta[$ and $\delta = \min\{1, \beta/\gamma\} + 1/2$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ a sequence in $[0, \delta[$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$.



Fixed point algorithm: Forward-Backward

Proof: Let $T = J_{\gamma A}(\mathrm{Id} - \gamma B)$.

1. $\mathrm{Fix}\, T = \mathrm{zer}\,(A + B) \neq \varnothing$:

$$(\forall x \in \mathcal{H}) \quad x \in \mathrm{Fix}\, T \Leftrightarrow x - \gamma Bx \in (\mathrm{Id} + \gamma A)x \Leftrightarrow 0 \in Ax + Bx.$$

2. The iterations can be written as

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n (Tx_n - x_n)$$

3. T is α-averaged : B β-cocoercive and $\gamma \in]0, 2\beta[\Rightarrow \mathrm{Id} - \gamma B$ is $\gamma/(2\beta)$-averaged and $J_{\gamma A}$ is $1/2$-averaged.
   Then, T is α-averaged with

$$\alpha = \frac{2}{1 + \frac{1}{\max\{\frac{1}{2}, \frac{\gamma}{2\beta}\}}} \quad \Leftrightarrow \quad \alpha^{-1} = \delta.$$

Fixed point algorithm: Forward-Backward

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

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.

Let $B\colon \mathcal{H} \to \mathcal{H}$ be β-cocoercive where $\beta \in \,]0, +\infty[$.

Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence in $[\underline{\gamma}, \overline{\gamma}]$ where $0 < \underline{\gamma} \leq \overline{\gamma} < 2\beta$.

Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\underline{\lambda}, 1]$ where $\underline{\lambda} \in]0, 1]$.

We assume that $\operatorname{zer}(A + B) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma_n B x_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma_n A} y_n - x_n). \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ converges weakly to a point in $\operatorname{zer}(A + B)$.

Remark: When $B = 0$ and $\lambda_n \equiv 1$, the algorithm reduces to the proximal point algorithm.

Optimization algorithm: Forward-Backward

Let $\mathcal{H}$ be a Hilbert space.
Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$ such that ∇g is $\frac{1}{\beta}$-Lipschitz where $\beta \in]0,+\infty[$.
Let $(\gamma_n)_{n\in\mathbb{N}}$ in $[\underline{\gamma},\overline{\gamma}]$ where $0 < \underline{\gamma} \leq \overline{\gamma} < 2\beta$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\underline{\lambda},1]$ where $\underline{\lambda} \in]0,1]$.
We assume that $\mathrm{Argmin}(f+g) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n(\mathrm{prox}_{\gamma_n f} y_n - x_n). \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ converges weakly to a minimizer of $f+g$.

Optimization algorithm: Forward-Backward

MM (Majoration-Minimization) interpretation

For all $n \in \mathbb{N}$, let $p_n = \operatorname{prox}_{\gamma_n f}\big(x_n - \gamma_n \nabla g(x_n)\big)$. We have:

$$f(x_{n+1}) = f\big((1-\lambda_n)x_n + \lambda_n p_n\big) \leq (1-\lambda_n) f(x_n) + \lambda_n f(p_n).$$

We use the descent lemma,

$$g(x_{n+1}) \leq g(x_n) + \langle \nabla g(x_n) \mid x_{n+1} - x_n \rangle + \frac{1}{2\beta}\|x_{n+1} - x_n\|^2,$$

the proximity operator definition

$$x_n - \gamma_n \nabla g(x_n) - p_n \in \partial f(p_n)$$

and, the subdifferential definition

$$\begin{aligned} & \gamma_n f(p_n) + \langle x_n - \gamma_n \nabla g(x_n) - p_n \mid x_n - p_n \rangle \leq \gamma_n f(x_n) \\ \Leftrightarrow \quad & \lambda_n f(p_n) + \langle \nabla g(x_n) \mid x_{n+1} - x_n \rangle + \gamma_n^{-1}\lambda_n^{-1}\|x_{n+1} - x_n\|^2 \leq \lambda_n f(x_n) \end{aligned}$$

since $x_{n+1} - x_n = \lambda_n(p_n - x_n)$.

Optimization algorithm: Forward-Backward

MM (Majoration-Minimization) interpretation

$$\Rightarrow f(x_{n+1}) + g(x_{n+1}) + \left(\frac{1}{\gamma_n \lambda_n} - \frac{1}{2\beta}\right)\|x_{n+1} - x_n\|^2 \leq f(x_n) + g(x_n).$$

Thus, if $\gamma_n^{-1}\lambda_n^{-1} - \frac{1}{2}\beta^{-1} \geq 0 \Rightarrow \gamma_n\lambda_n \leq 2\beta$, $\big(f(x_n) + g(x_n)\big)_{n\in\mathbb{N}}$ is a decreasing sequence.

Because $(\forall n \in \mathbb{N})$ $p_n = \mathrm{prox}_{\gamma_n f} y_n \in \mathrm{dom}\, f$, if $x_0 \in \mathrm{dom}\, f$, then $(\forall n \in \mathbb{N})$ $x_{n+1} = (1 - \lambda_n)x_n + \lambda_n p_n \in \mathrm{dom}\, f$.

Thus, $\big(f(x_n) + g(x_n)\big)_{n\in\mathbb{N}}$ is a real decreasing sequence.

Optimization algorithm: Forward-Backward

Beck-Teboule proximal gradient algorithm (Nesterov acceleration)

Let $x_0 = z_0 \in \mathcal{H}$, $t_0 = 1$, and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = z_n - \beta \nabla g(z_n) \\ x_{n+1} = \operatorname{prox}_{\beta f} y_n \\ t_{n+1} = \frac{1+\sqrt{4t_n^2+1}}{2} \\ \lambda_n = 1 + \frac{t_n - 1}{t_{n+1}} \\ z_{n+1} = z_n + \lambda_n (x_{n+1} - x_n). \end{cases}$$

Convergence of $(f(x_n))_{n\in\mathbb{N}}$ at optimal $O(1/n^2)$ rate, but convergence of $(x_n)_{n\in\mathbb{N}}$ not secured theoretically.

Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u \in \mathbb{R}^N}{\operatorname{Argmin}} \|AF^* u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in \,]0, +\infty[\text{ and } F \in \mathbb{R}^{N \times M}$$

- FB with $g = \|AF^* \cdot - z\|_2^2$, $f = \eta \| \cdot \|_1$ and $\widehat{x} = F^* \widehat{u}$.

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (\operatorname{prox}_{\gamma_n f} y_n - x_n) \end{cases}$$

Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u \in \mathbb{R}^N}{\mathrm{Argmin}} \|AF^* u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in\,]0, +\infty[\text{ and } F \in \mathbb{R}^{N \times M}$$

- FB with $g = \|AF^* \cdot - z\|_2^2$, $f = \eta \|\cdot\|_1$ and $\widehat{x} = F^* \widehat{u}$.

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) & \rightarrow \quad \text{Closed form: } 2FA^*(AF^* \cdot - z) \\ x_{n+1} = x_n + \lambda_n (\mathrm{prox}_{\gamma_n f} y_n - x_n) & \rightarrow \quad \text{Closed form} \end{cases}$$

Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u\in\mathbb{R}^N}{\operatorname{Argmin}} \|AF^*u - z\|_2^2 + \eta\|u\|_1 \text{ with } \eta \in \,]0,+\infty[\text{ and } F \in \mathbb{R}^{N\times M}$$

- FB

Degraded image z

Restored image [FB – DTT]

Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u \in \mathbb{R}^N}{\mathrm{Argmin}} \|AF^* u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in \,]0, +\infty[\text{ and } F \in \mathbb{R}^{N \times M}$$

- FB

Degraded image z

Restored image [PPXA – TV]

Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u \in \mathbb{R}^N}{\mathrm{Argmin}} \|AF^* u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in \,]0, +\infty[\text{ and } F \in \mathbb{R}^{N \times M}$$

- FB

Degraded image z

Restored image [DR – DWT]

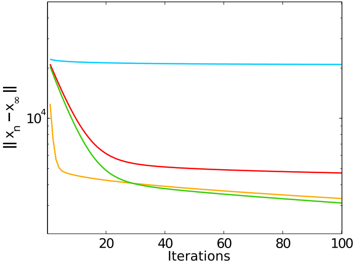
Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u\in\mathbb{R}^N}{\operatorname{Argmin}} \|AF^*u - z\|_2^2 + \eta\|u\|_1 \text{ with } \eta \in \,]0,+\infty[\text{ and } F \in \mathbb{R}^{N\times M}$$

- FB

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n(\operatorname{prox}_{\gamma_n f} y_n - x_n) \end{cases}$$



$$2\|AF^*\|^2\gamma_n \equiv \{0.1, 1.5, 1.9, 2\}$$

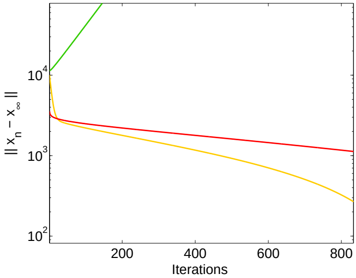
Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\widehat{u} \in \underset{u\in\mathbb{R}^N}{\operatorname{Argmin}} \|AF^* u - z\|_2^2 + \eta\|u\|_1 \text{ with } \eta \in\ ]0,+\infty[\text{ and } F \in \mathbb{R}^{N\times M}$$

- FB

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (\operatorname{prox}_{\gamma_n f} y_n - x_n) \end{cases}$$



$$\lambda_n \equiv \{0.5, 1, 1.1\}$$

Optimization algorithm: projected gradient

Let $\mathcal{H}$ be a Hilbert space.
Let C be a nonempty closed convex subset of $\mathcal{H}$.
Let $g \in \Gamma_0(\mathcal{H})$ such that ∇g is $\frac{1}{\beta}$-Lipschitz where $\beta \in \,]0,+\infty[$.
Let $(\gamma_n)_{n\in\mathbb{N}}$ a sequence in $[\underline{\gamma},\overline{\gamma}]$ where $0 < \underline{\gamma} \leq \overline{\gamma} < 2\beta$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $[\underline{\lambda},1]$ where $\underline{\lambda} \in]0,1]$.
We assume that $\operatorname{Argmin}_{x\in C} g(x) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (P_C y_n - x_n). \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ converges weakly to a minimizer of g on C.

Optimization algorithm: projected gradient

Image restoration : Variational approach – regularized

$$\widehat{u} \in \underset{u \in \mathbb{R}^N}{\operatorname{Argmin}} \|AF^* u - z\|_2^2 + \eta \|u\|_1 \quad \text{with} \quad \begin{cases} \eta \in \,]0, +\infty[\\ F \in \mathbb{R}^{N \times M} \end{cases}$$

Image restoration : Variational approach – constrained

$$\widehat{u} \in \underset{u \in \mathbb{R}^N, \|u\|_1 \leq \epsilon}{\operatorname{Argmin}} \|AF^* u - z\|_2^2 \quad \text{with} \quad \begin{cases} \epsilon \in \,]0, +\infty[\\ F \in \mathbb{R}^{N \times M} \end{cases}$$

Fixed point algorithm: *Forward-Backward-Forward*

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

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator.

Let $B\colon \mathcal{H} \to \mathcal{H}$ be a monotone operator, $\frac{1}{\beta}$-Lipschitz where $\beta \in \left]0, +\infty\right[$.

Let $\varepsilon \in]0, \beta/(1+\beta)[$ and $(\gamma_n)_{n\in\mathbb{N}}$ in $[\varepsilon, (1-\epsilon)\beta]$.
We assume that $\operatorname{zer}(A+B) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma_n B x_n \\ p_n = J_{\gamma_n A} y_n \\ q_n = p_n - \gamma_n B p_n \\ x_{n+1} = x_n - y_n + q_n. \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ and $(p_n)_{n\in\mathbb{N}}$ converge weakly to $\widehat{x} \in \operatorname{zer}(A+B)$.

Fixed point algorithm: *Forward-Backward-Forward*

Let $\mathcal{H}$ be a Hilbert space.
Let $f \in \Gamma_0(\mathcal{H})$.
Let $g \in \Gamma_0(\mathcal{H})$ such that ∇g is $\frac{1}{\beta}$-Lipschitz where $\beta \in]0, +\infty[$.
Let $\varepsilon \in]0, \beta/(1+\beta)[$ and $(\gamma_n)_{n\in\mathbb{N}}$ in $[\varepsilon, (1-\epsilon)\beta]$.
We assume that $\operatorname{Argmin}(f+g) \neq \varnothing$. Let $x_0 \in \mathcal{H}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ p_n = \operatorname{prox}_{\gamma_n f} y_n \\ q_n = p_n - \gamma_n \nabla g(p_n) \\ x_{n+1} = x_n - y_n + q_n. \end{cases}$$

$(x_n)_{n\in\mathbb{N}}$ and $(p_n)_{n\in\mathbb{N}}$ converge weakly to $\widehat{x} \in \operatorname{Argmin}(f+g)$.

Part 4: Duality

1. General duality concepts
   - Primal and dual problems
   - Duality theorems
   - Inf-convolution

2. Augmented Lagrangian algorithms
   - ADMM
   - SDMM

3. Primal-dual algorithms
   - FB-based PD algorithm
   - M+LFBF algorithm

Duality

Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{H} \to 2^{\mathcal{H}}$.
The parallel sum of A and B is

$$A \,\square\, B = (A^{-1} + B^{-1})^{-1}.$$

- If A and B are monotone, then $A \,\square\, B$ is monotone.
- $A \,\square\, N_{\{0\}} = A$ where $N_{\{0\}} = \partial \iota_{\{0\}}$ and $\operatorname{gra} N_{\{0\}} = \{0\} \times \mathcal{H}$.
- $A \,\square\, B = B \,\square\, A$.

Duality

Primal problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, $C\colon \mathcal{H} \to 2^{\mathcal{H}}$, $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ and $D\colon \mathcal{G} \to 2^{\mathcal{G}}$ be monotone operators. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Find $\widehat{x} \in \mathcal{H}$ such that

$$0 \in A\widehat{x} + C\widehat{x} + L^*(B \,\Box\, D)L\widehat{x}.$$

Dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$, $C\colon \mathcal{H} \to 2^{\mathcal{H}}$, $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ and $D\colon \mathcal{G} \to 2^{\mathcal{G}}$ be monotone operators. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. Find $\widehat{v} \in \mathcal{G}$ such that

$$0 \in -L(A^{-1} \,\Box\, C^{-1})(-L^*\widehat{v}) + B^{-1}\widehat{v} + D^{-1}\widehat{v}.$$

Duality theorem

Let $\widehat{x} \in \mathcal{H}$ and $\widehat{v} \in \mathcal{G}$. $(\widehat{x}, \widehat{v})$ is a **Kuhn-Tucker point** if

$$\begin{cases} -L^*\widehat{v} \in (A + C)\widehat{x} \\ \quad L\widehat{x} \in (B^{-1} + D^{-1})\widehat{v}. \end{cases}$$

If $\widehat{x} \in \mathcal{H}$ is a solution to the primal problem, then there exists a solution $\widehat{v}$ to the dual problem such that $(\widehat{x}, \widehat{v})$ is a Kuhn-Tucker point.

Duality theorem

Let $\widehat{x} \in \mathcal{H}$ and $\widehat{v} \in \mathcal{G}$. $(\widehat{x}, \widehat{v})$ is a **Kuhn-Tucker point** if

$$\begin{cases} -L^*\widehat{v} \in (A + C)\widehat{x} \\ \quad L\widehat{x} \in (B^{-1} + D^{-1})\widehat{v}. \end{cases}$$

If $\widehat{v} \in \mathcal{G}$ is a solution to the dual problem, then there exists a solution $\widehat{x}$ to the primal problem such that $(\widehat{x}, \widehat{v})$ is a Kuhn-Tucker point.

Duality theorem

Let $\widehat{x} \in \mathcal{H}$ and $\widehat{v} \in \mathcal{G}$. $(\widehat{x}, \widehat{v})$ is a **Kuhn-Tucker point** if

$$\begin{cases} -L^*\widehat{v} \in (A + C)\widehat{x} \\ L\widehat{x} \in (B^{-1} + D^{-1})\widehat{v}. \end{cases}$$

If $(\widehat{x}, \widehat{v})$ is a Kuhn-Tucker point, then $\widehat{x}$ is a solution to the primal problem and $\widehat{v}$ is a solution to the dual problem.

Inf-convolution

Let $\mathcal{H}$ be a Hilbert space.
Let $f : \mathcal{H} \to]-\infty, +\infty]$ and $g : \mathcal{H} \to]-\infty, +\infty]$.
The inf-convolution of f and g is

$$f \,\square\, g : \mathcal{H} \to [-\infty, +\infty] : x \mapsto \inf_{y \in \mathcal{H}} f(y) + g(x - y)$$

Inf-convolution

Let $\mathcal{H}$ be a Hilbert space.
Let $f \colon \mathcal{H} \to]-\infty,+\infty]$ and $g \colon \mathcal{H} \to]-\infty,+\infty]$.
The inf-convolution of f and g is

$$f \,\Box\, g \colon \mathcal{H} \to [-\infty,+\infty] \colon x \mapsto \inf_{\substack{(u,v)\in\mathcal{H}^2 \\ u+v=x}} f(u)+g(v)$$

Inf-convolution

Let $\mathcal{H}$ be a Hilbert space.
Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ and $g\colon \mathcal{H} \to \,]-\infty, +\infty]$.
The inf-convolution of f and g is

$$f \,\Box\, g\colon \mathcal{H} \to [-\infty, +\infty]\colon x \mapsto \inf_{\substack{(u,v)\in\mathcal{H}^2 \\ u+v=x}} f(u) + g(v)$$

- $f \,\Box\, \iota_{\{0\}} = f$
- $f \,\Box\, g = g \,\Box\, f$
- $\mathrm{dom}\,(f \,\Box\, g) = \mathrm{dom}\, f + \mathrm{dom}\, g$
- ${}^{\gamma}f = f \,\Box\, \frac{1}{2\gamma}\|\cdot\|^2$, $\gamma \in \,]0, +\infty[$.

Inf-convolution

Let $\mathcal{H}$ be a Hilbert space.
Let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ and $g\colon \mathcal{H} \to \,]-\infty,+\infty]$.
The **inf-convolution** of f and g is

$$f \,\Box\, g\colon \mathcal{H} \to [-\infty,+\infty]\colon x \mapsto \inf_{y\in\mathcal{H}} f(y) + g(x-y)$$

If $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ and $g\colon \mathcal{H} \to \,]-\infty,+\infty]$ are convex,
then $f \,\Box\, g$ is convex.

Inf-convolution

	conjugate		Fourier transform ($\mathcal{H}$ finite dimensional)	
Property	$h(x)$	$h^*(u)$	$h(x)$	$\widehat{h}(\nu)$
inf-convolution /convolution	$(f \,\square\, g)(x)$ $= \inf_{y\in\mathcal{H}} f(y) + g(x-y)$	$f^*(u) + g^*(u)$	$(f \star g)(x)$ $= \int_{\mathcal{H}} f(y)g(x-y)dy$	$\widehat{f}(\nu)\widehat{g}(\nu)$
sum/product	$f(x) + g(x)$ $f \in \Gamma_0(\mathcal{H})$ $g \in \Gamma_0(\mathcal{H})$ $\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \varnothing$	$(f^* \,\square\, g^*)(u)$	$f(x)g(x)$	$(\widehat{f} \star \widehat{g})(\nu)$

Inf-convolution

Let $\mathcal{H}$ be a Hilbert space.
Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.
If $\mathrm{dom}\, f^* \cap \mathrm{int}\,(\mathrm{dom}\, g^*) \neq \varnothing$, then

$$\partial(f \underbrace{\,\square\,}_{\text{inf-convolution}} g) = \partial f \underbrace{\,\square\,}_{\text{parallel sum}} \partial g.$$

Inf-convolution

Let $\mathcal{H}$ be a Hilbert space.
Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{H})$.
If $\operatorname{dom} f^* \cap \operatorname{int}(\operatorname{dom} g^*) \neq \varnothing$, then

$$\partial(f \underbrace{\square}_{\text{inf-convolution}} g) = \partial f \underbrace{\square}_{\text{parallel sum}} \partial g.$$

Proof:

$$\begin{aligned}
\partial(f \square g) &= \partial(f^* + g^*)^* \\
&= \big(\partial(f^* + g^*)\big)^{-1} \\
&= \big(\partial f^* + \partial g^*\big)^{-1} \\
&= \big((\partial f)^{-1} + (\partial g)^{-1}\big)^{-1} \\
&= \partial f \square \partial g.
\end{aligned}$$

Fenchel-Rockafellar duality

Primal problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $f : \mathcal{H} \to]-\infty, +\infty]$, $g : \mathcal{G} \to]-\infty, +\infty]$. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
We want to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + g(Lx).$$

Dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.
Let $f : \mathcal{H} \to]-\infty, +\infty]$, $g : \mathcal{G} \to]-\infty, +\infty]$. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.
We want to

$$\underset{v \in \mathcal{G}}{\text{minimize}} \;\; f^*(-L^* v) + g^*(v).$$

Fenchel-Rockafellar duality

Weak duality

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces.

Let f be a proper fonction from $\mathcal{H}$ to $]-\infty,+\infty]$, g be a proper function from $\mathcal{G}$ to $]-\infty,+\infty]$ and $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$. Let

$$\mu = \inf_{x\in\mathcal{H}} f(x) + g(Lx) \quad \text{and} \quad \mu^* = \inf_{v\in\mathcal{G}} f^*(-L^*v) + g^*(v).$$

We have $\mu \geq -\mu^*$. If $\mu \in \mathbb{R}$, $\mu + \mu^*$ is called the duality gap.

Fenchel-Rockafellar duality

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$ and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. If $\mathrm{int}\,(\mathrm{dom}\, g) \cap L(\mathrm{dom}\, f) \neq \varnothing$ and $\mathrm{dom}\, g \cap \mathrm{int}\,\big(L(\mathrm{dom}\, f)\big) \neq \varnothing$, then

$$\partial f + L^* \partial g L = \partial(f + g \circ L)$$

Fenchel-Rockafellar duality

Duality theorem (1)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$ and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

$$\mathrm{zer}\,(\partial f + L \partial g L^*) \neq \varnothing \quad \Leftrightarrow \quad \mathrm{zer}\,\big((-L)\partial f^*(-L^*) + \partial g^*\big) \neq \varnothing$$

Fenchel-Rockafellar duality

Duality theorem (2)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$ and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

- If there exists $\widehat{x} \in \mathcal{H}$ such that $0 \in \partial f(\widehat{x}) + L^* \partial g(L\widehat{x})$ then $\widehat{x}$ is a solution to the primal problem. Moreover, $-L^*\widehat{v} \in \partial f(\widehat{x})$ and $L\widehat{x} \in \partial g^*(\widehat{v})$ where $\widehat{v}$ is a solution to the dual problem.
- If there exists $(\widehat{x}, \widehat{v}) \in \mathcal{H} \times \mathcal{G}$ such that $-L^*\widehat{v} \in \partial f(\widehat{x})$ and $L\widehat{x} \in \partial g^*(\widehat{v})$ then $\widehat{x}$ (resp. $\widehat{v}$) is a solution to the primal (resp. dual) problem.

<u>Particular case</u>: If $f = \varphi + \frac{1}{2}\|\cdot - z\|^2$ where $\varphi \in \Gamma_0(\mathcal{H})$ and $z \in \mathcal{H}$, then

$$-L^*\widehat{v} \in \partial f(\widehat{x}) \Leftrightarrow -L^*\widehat{v} \in \partial \varphi(\widehat{x}) + \widehat{x} - z.$$

We have then $\widehat{x} = \operatorname{prox}_{\varphi}(-L^*\widehat{v} + z)$.

Fenchel-Rockafellar duality

Duality theorem (3): strong duality

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces, $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$ and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$. If $\mathrm{int}\,(\mathrm{dom}\, g) \cap L(\mathrm{dom}\, f) \neq \varnothing$ or $\mathrm{dom}\, g \cap \mathrm{int}\,\big(L(\mathrm{dom}\, f)\big) \neq \varnothing$, then

$$\mu = \inf_{x \in \mathcal{H}} f(x) + g(Lx) = -\min_{v \in \mathcal{G}} f^*(-L^* v) + g^*(v) = -\mu^* .$$

Fenchel-Rockafellar duality

Extension of the duality theorem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert space, $f \in \Gamma_0(\mathcal{H})$, $h \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$, $\ell \in \Gamma_0(\mathcal{G})$ such that $\operatorname{dom} h = \mathcal{H}$ and $\operatorname{dom} \ell^* = \mathcal{G}$, $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

If there exists $\widehat{x} \in \mathcal{H}$ such that $0 \in \partial f(\widehat{x}) + L^*(\partial g \,\square\, \partial \ell)(L\widehat{x}) + \partial h(\widehat{x})$ then

- $\widehat{x}$ is a solution to the primal problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f(x) + (g \,\square\, \ell)(Lx) + h(x)$$

- $-L^*\widehat{v} \in \partial f(\widehat{x}) + \partial h(\widehat{x})$ and $L\widehat{x} \in \partial g^*(\widehat{v}) + \partial \ell^*(\widehat{v})$ where $\widehat{v}$ is a solution to the dual problem

$$\underset{v \in \mathcal{G}}{\text{minimize}} \;\; (f^* \,\square\, h^*)(-L^* v) + g^*(v) + \ell^*(v).$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = \mathrm{prox}_{\gamma g^*} u_n \\ w_n = \mathrm{prox}_{\gamma f^* \circ (-L^*)}(2v_n - u_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0,+\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n \in \gamma\partial\big(f^* \circ (-L^*)\big)w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n \in -\gamma L \circ \partial f^* \circ (-L^*) w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n \in -\gamma L \circ \underbrace{\partial f^* \circ (-L^*) w_n}_{x_n \in} \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \left]0, +\infty\right[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n = -\gamma L x_n \\ x_n \in \partial f^*(-L^* w_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \operatorname{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ 2v_n - u_n - w_n = -\gamma L x_n \\ -L^* w_n \in \partial f(x_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} v_n = u_n - \gamma \mathrm{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ L^*(u_n - 2v_n - \gamma L x_n) \in \partial f(x_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

using $y_n = \gamma^{-1}(u_n - v_n)$ and $z_n = \gamma^{-1} v_n$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \operatorname{prox}_{\gamma^{-1}g}(\gamma^{-1} u_n) \\ L^*(y_n - z_n - Lx_n) \in \frac{1}{\gamma}\partial f(x_n) \\ u_{n+1} = \gamma z_n + \gamma L x_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_n = \mathrm{prox}_{\gamma^{-1} g}(\gamma^{-1} u_n) \\ x_n = \underset{x \in \mathcal{H}}{\mathrm{argmin}} \;\; \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ u_{n+1} = \gamma z_n + \gamma L x_n \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\equiv$ Douglas-Rachford for the dual problem

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.
Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.
The Douglas-Rachford iterations to minimize $f^* \circ (-L^*) + g^*$ are

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \; \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}} (z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}}\;\; f(x)+g(Lx) \quad\Leftrightarrow\quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}}\;\; f(x)+g(y)$$

Lagrange function:

$$\big(\forall(x,y,v)\in\mathcal{H}\times\mathcal{G}^2\big)\;\; \mathcal{L}(x,y,z)=f(x)+g(y)+\langle v\mid Lx-y\rangle$$

where $v\in\mathcal{G}$ denotes the Lagrange multiplier.

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}}\; f(x)+g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}}\; f(x)+g(y)$$

Augmented Lagrange function: let $\gamma \in \,]0,+\infty[$, we define

$$(\forall(x,y,z)\in\mathcal{H}\times\mathcal{G}^2)\quad \widetilde{\mathcal{L}}(x,y,z)=f(x)+g(y)+\gamma\langle z \mid Lx-y\rangle+\frac{\gamma}{2}\|Lx-y\|^2$$

$$\Rightarrow \quad \begin{cases}\partial_x\widetilde{\mathcal{L}}(x,y,z)=\partial f(x)+\gamma L^*z+\gamma L^*(Lx-y)\\ \partial_y\widetilde{\mathcal{L}}(x,y,z)=\partial g(y)-\gamma z+\gamma(y-Lx)\\ \partial_z\widetilde{\mathcal{L}}(x,y,z)=Lx-y.\end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}} \; f(x)+g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}} \; f(x)+g(y)$$

Augmented Lagrange function:

Thus $(0,0,0) \in \partial_x\mathcal{L}(\widehat{x},\widehat{y},\widehat{z}) \times \partial_y\mathcal{L}(\widehat{x},\widehat{y},\widehat{z}) \times \partial_z\mathcal{L}(\widehat{x},\widehat{y},\widehat{z})$

$$\Leftrightarrow \quad \begin{cases} -\gamma L^*\widehat{z} \in \partial f(\widehat{x}) \\ \gamma\widehat{z} \in \partial g(L\widehat{x}) \quad \Leftrightarrow \quad L\widehat{x} \in \partial g^*(\gamma\widehat{z}) \\ \widehat{y} = L\widehat{x}. \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\Rightarrow$ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}}\ f(x)+g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}}\ f(x)+g(y)$$

Augmented Lagrange function: To sum up, if

$$\left(\forall(x,y,z)\in\mathcal{H}\times\mathcal{G}^2\right)\ \widetilde{\mathcal{L}}(x,y,z)=f(x)+g(y)+\gamma\,\langle z\mid Lx-y\rangle+\frac{\gamma}{2}\|Lx-y\|^2$$

then $(\widehat{x},\widehat{y},\widehat{z})$ critical point of $\widetilde{\mathcal{L}}\Rightarrow\widehat{y}=L\widehat{x}$, $\widehat{x}$ solution to the primal problem and $\gamma\widehat{z}$ solution to the dual problem.
Moreover, $(\widehat{x},\widehat{y},\widehat{z})$ is a critical point of $\widetilde{\mathcal{L}}\Leftrightarrow$ $(\widehat{x},\widehat{y},\widehat{z})$ saddle point of $\widetilde{\mathcal{L}}$, i.e. $\widetilde{\mathcal{L}}(\widehat{x},\widehat{y},z)\leq\widetilde{\mathcal{L}}(\widehat{x},\widehat{y},\widehat{z})\leq\widetilde{\mathcal{L}}(x,y,\widehat{z})$.

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

$\Rightarrow$ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}}\; f(x)+g(Lx) \quad\Leftrightarrow\quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}}\; f(x)+g(y)$$

Augmented Lagrange function: Algorithm to find $(\widehat{x},\widehat{y},\widehat{z})$:

$$(\forall n\in\mathbb{N})\qquad \begin{cases} x_n = \underset{x\in\mathcal{H}}{\operatorname{argmin}}\; \widetilde{\mathcal{L}}(x,y_n,z_n)\\ y_{n+1} = \underset{y\in\mathcal{G}}{\operatorname{argmin}}\; \widetilde{\mathcal{L}}(x_n,y,z_n)\\ z_{n+1} \text{ such that } \widetilde{\mathcal{L}}(x_n,y_{n+1},z_{n+1}) \geq \widetilde{\mathcal{L}}(x_n,y_{n+1},z_n). \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}}\; f(x)+g(Lx) \quad\Leftrightarrow\quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}}\; f(x)+g(y)$$

Augmented Lagrange function: Algorithm to find $(\widehat{x},\widehat{y},\widehat{z})$:

$$(\forall n\in\mathbb{N})\qquad \begin{cases} x_n = \underset{x\in\mathcal{H}}{\operatorname{argmin}}\; f(x)+\gamma\langle z_n \mid Lx-y_n\rangle + \frac{\gamma}{2}\|Lx-y_n\|^2\\ y_{n+1} = \underset{y\in\mathcal{G}}{\operatorname{argmin}}\; g(y)+\gamma\langle z_n \mid Lx_n-y\rangle + \frac{\gamma}{2}\|Lx_n-y\|^2\\ z_{n+1} = z_n + \frac{1}{\gamma}\nabla_z\widetilde{\mathcal{L}}(x_n,y_{n+1},z_n). \end{cases}$$

Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x\in\mathcal{H}}{\text{minimize}} \; f(x)+g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x\in\mathcal{H},y\in\mathcal{G}\\ Lx=y}}{\text{minimize}} \; f(x)+g(y)$$

Augmented Lagrange function: Algorithm to find $(\widehat{x},\widehat{y},\widehat{z})$:

$$(\forall n\in\mathbb{N}) \qquad \begin{cases} x_n = \underset{x\in\mathcal{H}}{\text{argmin}} \; \frac{1}{2}\|Lx-y_n+z_n\|^2+\frac{1}{\gamma}f(x)\\ y_{n+1} = \text{prox}_{\frac{g}{\gamma}}(z_n+Lx_n)\\ z_{n+1} = z_n+Lx_n-y_{n+1}. \end{cases}$$

Augmented Lagrange method

ADMM algorithm (*Alternating-direction method of multipliers*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ et $g \in \Gamma_0(\mathcal{G})$.

Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \; \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = L x_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}} (z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

Augmented Lagrange method

ADMM algorithm (*Alternating-direction method of multipliers*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $f \in \Gamma_0(\mathcal{H})$ et $g \in \Gamma_0(\mathcal{G})$.

Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ such that L^*L is an isomorphism and let $\gamma \in \,]0, +\infty[$.

We assume that $\operatorname{int}(\operatorname{dom} g) \cap L(\operatorname{dom} f) \neq \varnothing$ or $\operatorname{dom} g \cap \operatorname{int}\big(L(\operatorname{dom} f)\big) \neq \varnothing$ and that $\operatorname{Argmin}(f + g \circ L) \neq \varnothing$. Let

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \ \frac{1}{2}\|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = L x_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}}(z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

We have:

- $x_n \rightharpoonup \widehat{x}$ where $\widehat{x} \in \operatorname{Argmin}(f + g \circ L)$
- $\gamma z_n \rightharpoonup \widehat{v}$ where $\widehat{v} \in \operatorname{Argmin}(f^* \circ (-L^*) + g^*)$.

Augmented Lagrangian method

SDMM algorithm (*Simultaneous-direction method of multipliers*)

Let $\mathcal{H}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_m$ be Hilbert spaces.

Let $(\forall i \in \{1, \ldots, m\})$ $g_i \in \Gamma_0(\mathcal{G}_i)$ and $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$. Let $\gamma \in]0, +\infty[$.

We assume that $\sum_{i=1}^m L_i^* L_i$ is an isomorphism

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \left(\sum_{i=1}^m L_i^* L_i\right)^{-1} \sum_{i=1}^m L_i^*(y_{n,i} - z_{n,i}) \\ s_{n,i} = L_i x_n, \quad i \in \{1, \ldots, m\} \\ y_{n+1,i} = \operatorname{prox}_{\frac{g_i}{\gamma}}\left(z_{n,i} + s_{n,i}\right), \quad i \in \{1, \ldots, m\} \\ z_{n+1,i} = z_{n,i} + s_{n,i} - y_{n+1,i}, \quad i \in \{1, \ldots, m\}. \end{cases}$$

Augmented Lagrangian method

SDMM algorithm (*Simultaneous-direction method of multipliers*)

Let $\mathcal{H}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_m$ be Hilbert spaces.

Let $(\forall i \in \{1, \ldots, m\})$ $g_i \in \Gamma_0(\mathcal{G}_i)$ and $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$. Let $\gamma \in]0, +\infty[$.

We assume that $\sum_{i=1}^m L_i^* L_i$ is an isomorphism and there exists $\widetilde{x} \in \mathcal{H}$ such that $(\forall i \in \{1, \ldots, m\})$ $L_i \widetilde{x} \in \operatorname{int}(\operatorname{dom} g_i)$. Let

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \left(\sum_{i=1}^m L_i^* L_i\right)^{-1} \sum_{i=1}^m L_i^*(y_{n,i} - z_{n,i}) \\ s_{n,i} = L_i x_n, \quad i \in \{1, \ldots, m\} \\ y_{n+1,i} = \operatorname{prox}_{\frac{g_i}{\gamma}}(z_{n,i} + s_{n,i}), \quad i \in \{1, \ldots, m\} \\ z_{n+1,i} = z_{n,i} + s_{n,i} - y_{n+1,i}, \quad i \in \{1, \ldots, m\}. \end{cases}$$

We have:

- $x_n \rightharpoonup \widehat{x}$ where $\widehat{x} \in \operatorname{Argmin}(\sum_{i=1}^m g_i \circ L_i)$
- $\gamma z_n = \gamma(z_{n,i})_{1 \le i \le m} \rightharpoonup \widehat{v}$ where $\widehat{v} = (\widehat{v}_i)_{1 \le i \le m} \in \underset{\substack{v=(v_i)_{1\le i\le m} \\ \sum_{i=1}^m L_i^* v_i = 0}}{\operatorname{Argmin}} \left(\sum_{i=1}^m g_i^*(v_i)\right)$.

Primal-dual method

FB-based PD algorithm (*Condat-Vũ-...*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two maximally monotone operators.

Let $C\colon \mathcal{H} \to \mathcal{H}$ be a μ-cocoercive operator with $\mu \in \,]0, +\infty[$.

Let $D\colon \mathcal{G} \to 2^{\mathcal{G}}$ be a ν-strongly monotone operator with $\nu \in \,]0, +\infty[$.

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = J_{\tau A}\big(x_n - \tau(Cx_n + L^* v_n)\big) \\ q_n = J_{\sigma B^{-1}}\big(v_n + \sigma\big(L(2p_n - x_n) - D^{-1} v_n\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

Primal-dual method

FB-based PD algorithm (*Condat-Vũ-...*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two maximally monotone operators.

Let $C\colon \mathcal{H} \to \mathcal{H}$ be a μ-cocoercive operator with $\mu \in \,]0, +\infty[$.

Let $D\colon \mathcal{G} \to 2^{\mathcal{G}}$ be a ν-strongly monotone operator with $\nu \in \,]0, +\infty[$.

Let $\beta = \min\{\mu, \nu\}$, $\tau \in \,]0, +\infty[$, $\sigma \in \,]0, +\infty[$, $\rho = \min\{\tau^{-1}, \sigma^{-1}\}(1 - \sqrt{\tau\sigma}\|L\|)$ and $\delta = \min\{1, \rho\beta\} + 1/2$. We assume that $2\rho\beta > 1$.

Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, \delta[$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$.

We assume that $\operatorname{zer}\big(A + C + L^*(B \,\square\, D)L\big) \neq \varnothing$. Let $x_0 \in \mathcal{H}$, $v_0 \in \mathcal{G}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = J_{\tau A}\big(x_n - \tau(Cx_n + L^* v_n)\big) \\ q_n = J_{\sigma B^{-1}}\big(v_n + \sigma\big(L(2p_n - x_n) - D^{-1}v_n\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

We have: $x_n \rightharpoonup \widehat{x} \in \operatorname{zer}\big(A + C + L^*(B \,\square\, D)L\big)$

and $v_n \rightharpoonup \widehat{v} \in \operatorname{zer}\big((-L)(A^{-1} \,\square\, C^{-1})(-L^*) + B^{-1} + D^{-1}\big)$.

Primal-dual method

Proof:

$(\widehat{x}, \widehat{v})$ Kuhn-Tucker point iff $(\widehat{x}, \widehat{v}) \in \mathrm{zer}\,(P + Q)$ where

- P maximally monotone such that $P = M + S$ with

$$M\colon (x, v) \mapsto (Ax, B^{-1}v)$$
$$S\colon (x, v) \mapsto \begin{bmatrix} 0 & L^* \\ -L & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}$$

- $Q\colon (x, v) \mapsto (Cx, D^{-1}v)$ β-cocoercive with $\beta = \min\{\mu, \nu\}$

$\Rightarrow$ Forward-Backward algorithm.

Primal-dual optimization algorithm

FB-based PD algorithm

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.

Let $h \in \Gamma_0(\mathcal{H})$ having a $1/\mu$-Lipschitzian gradient, $\ell \in \Gamma_0(\mathcal{G})$ ν-strongly convex

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \operatorname{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \operatorname{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n) - \nabla \ell^*(v_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

Primal-dual optimization algorithm

FB-based PD algorithm

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.

Let $h \in \Gamma_0(\mathcal{H})$ having a $1/\mu$-Lipschitzian gradient, $\ell \in \Gamma_0(\mathcal{G})$ ν-strongly convex and $\beta = \min\{\mu, \nu\}$.

Let $\tau \in]0, +\infty[$, $\sigma \in]0, +\infty[$, $\rho = \min\{\tau^{-1}, \sigma^{-1}\}(1 - \sqrt{\tau\sigma}\|L\|)$ et $\delta = \min\{1, \rho\beta\} + 1/2$. We assume that $2\rho\beta > 1$.

Let $(\lambda_n)_{n\in\mathbb{N}}$ a sequence in $]0, \delta[$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$.

We assume that $\operatorname{zer}\big(\partial f + \nabla h + L^*(\partial g \,\square\, \partial \ell)L\big) \neq \varnothing$. Let $x_0 \in \mathcal{H}$, $v_0 \in \mathcal{G}$ and

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \operatorname{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \operatorname{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n) - \nabla \ell^*(v_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

We have:

- $x_n \rightharpoonup \widehat{x} \in \operatorname{Argmin}\big(f + h + (g \,\square\, \ell) \circ L\big)$
- $v_n \rightharpoonup \widehat{v} \in \operatorname{Argmin}\big((f^* \,\square\, h^*) \circ (-L^*) + g^* + \ell^*\big)$

Primal-dual optimization algorithm

FB-based PD algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \mathrm{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \mathrm{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n) - \nabla \ell^*(v_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:
  - * No operator inversion.
  - * Allow the use of proximable or/and differentiable functions.

Primal-dual optimization algorithm

FB-based PD algorithm $\Rightarrow$ CP algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \mathrm{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \mathrm{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n) - \nabla \ell^*(v_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

Primal-dual optimization algorithm

FB-based PD algorithm ⇒ CP algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_{n+1} = \operatorname{prox}_{\tau f}\big(x_n - \tau L^* v_n\big) \\ y_n = 2x_{n+1} - x_n \\ v_{n+1} = \operatorname{prox}_{\sigma g^*}\big(v_n + \sigma L y_n\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

Primal-dual optimization algorithm

FB-based PD algorithm $\Rightarrow$ FB algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \operatorname{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \operatorname{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n) - \nabla \ell^*(v_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$, $\ell = \iota_{\{0\}}$ and $L = 0$, this yields the forward-backward algorithm.

Primal-dual optimization algorithm

FB-based PD algorithm $\Rightarrow$ FB algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \operatorname{prox}_{\tau f}\big(x_n - \tau \nabla h(x_n)\big) \\ x_{n+1} = x_n + \lambda_n\big(p_n - x_n\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$, $\ell = \iota_{\{0\}}$ and $L = 0$, this yields the forward-backward algorithm.

Primal-dual optimization algorithm

FB-based PD algorithm $\Rightarrow$ DR algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = \operatorname{prox}_{\tau f}\big(x_n - \tau\big(\nabla h(x_n) + L^* v_n\big)\big) \\ q_n = \operatorname{prox}_{\sigma g^*}\big(v_n + \sigma\big(L(2p_n - x_n) - \nabla \ell^*(v_n)\big)\big) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n\big((p_n, q_n) - (x_n, v_n)\big). \end{cases}$$

- Remark:

  * When $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$, $\ell = \iota_{\{0\}}$ and $L = 0$, this yields the forward-backward algorithm.

  * In the limit case when $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$, $L = \mathrm{Id}$ and $\sigma = 1/\tau$, this yields the Douglas-Rachford algorithm.

Primal-dual optimization algorithm

FB-based PD algorithm $\Rightarrow$ DR algorithm

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} x_{n+1} = \mathrm{prox}_{\tau f}\big(x_n - \tau v_n\big) \\ s_n = \mathrm{prox}_{\tau g}\big(2x_{n+1} - (x_n - \tau v_n)\big) \\ x_{n+1} - \tau v_{n+1} = \big(x_n - \tau v_n\big) + s_n - x_{n+1} \end{cases}$$

- Remark:

  * When $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$ and $\sigma\tau\|L\|^2 < 1$, this yields the Chambolle-Pock algorithm.

  * When $g = 0$, $\ell = \iota_{\{0\}}$ and $L = 0$, this yields the forward-backward algorithm.

  * In the limit case when $h = 0$, $\ell = \iota_{\{0\}}$, $\lambda_n \equiv 1$, $L = \mathrm{Id}$ and $\sigma = 1/\tau$, this yields the Douglas-Rachford algorithm.

FB-based PD algorithm

Image restoration : Variational approach

$$\widehat{x} \in \operatorname*{Argmin}_{x\in\mathbb{R}^N} \|Ax - z\|_2^2 + \eta \|[W_1^* H_1^* \dots W_J^* H_J^*]^* x\|_{2,1} + \iota_C(x)$$

with
$$\begin{cases} \eta > 0 \\ H_1, \dots, H_J \in \mathbb{R}^{N\times N} & \text{(e.g. operators associated with filters)} \\ W_1, \dots, W_J \in \mathbb{R}^{N\times N} & \text{(e.g. weights)} \\ C = [0, 255]^N \end{cases}$$

- PPXA+ : $(2\mathrm{Id} + W_1^* H_1^* H_1 W_1 + \dots + W_J^* H_J^* H_J W_J)^{-1}$ complicated
- FB-based PD algorithm : implementation without inversion

FB-based PD algorithm

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \|[W_1^* H_1^* \ldots W_J^* H_J^*]^* x\|_{2,1} + \iota_C(x)$$

Degraded image z

Restored image
[PD – NLTV]

FB-based PD algorithm

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[W_1^* H_1^* \dots W_J^* H_J^*]^* x\|_{2,1} + \iota_C(x)$$

Degraded image z

Restored image
[FB – DTT]

FB-based PD algorithm

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \|[W_1^* H_1^* \dots W_J^* H_J^*]^* x\|_{2,1} + \iota_C(x)$$

Degraded image z

Restored image [PPXA – TV]

FB-based PD algorithm

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[W_1^* H_1^* \dots W_J^* H_J^*]^* x\|_{2,1} + \iota_C(x)$$

Degraded image z

Restored image
[DR – DWT]

Primal-dual method

Parallel FB-based PD algorithm

Let $\mathcal{H}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_m$ be Hilbert spaces. For every $i \in \{1, \ldots, m\}$, let $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and, for every $\in \{1, \ldots, m\}$, $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ maximally monotone operators.

Let $C\colon \mathcal{H} \to \mathcal{H}$ be a μ-cocoercive operator where $\mu \in \left]0, +\infty\right[$,

$(\forall i \in \{1, \ldots, m\})$ $D_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be ν_i-strongly monotone where $\nu_i \in \left]0, +\infty\right[$.

$$
(\forall n \in \mathbb{N}) \qquad \begin{cases} p_n = J_{\tau A}\big(x_n - \tau\big(C x_n + \sum_{i=1}^m L_i^* v_{n,i}\big)\big) \\ q_{n,i} = J_{\sigma B_i^{-1}}\big(v_{n,i} + \sigma\big(L_i(2p_n - x_n) - D_i^{-1} v_{n,i}\big)\big), \ i \in \{1, \ldots, m\} \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ v_{n+1,i} = v_{n,i} + \lambda_n(q_{n,i} - v_{n,i}), \ i \in \{1, \ldots, m\}. \end{cases}
$$

Primal-dual method

Parallel FB-based PD algorithm

Let $\mathcal{H}$ and $\mathcal{G}_1, \ldots, \mathcal{G}_m$ be Hilbert spaces. For every $i \in \{1, \ldots, m\}$, let $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$.

Let $A \colon \mathcal{H} \to 2^{\mathcal{H}}$ and, for every $\in \{1, \ldots, m\}$, $B_i \colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ maximally monotone operators.

Let $C \colon \mathcal{H} \to \mathcal{H}$ be a μ-cocoercive operator where $\mu \in \left]0, +\infty\right[$,

$(\forall i \in \{1, \ldots, m\})$ $D_i \colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$ be ν_i-strongly monotone where $\nu_i \in \left]0, +\infty\right[$.

Let $\beta = \min\{\mu, \nu_1, \ldots, \nu_m\}$, $\tau > 0$, $\sigma > 0$, $\rho = \min\{\tau^{-1}, \sigma^{-1}\}(1 - \sqrt{\tau\sigma \sum_{i=1}^m \|L_i\|^2})$
and $\delta = \min\{1, \rho\beta\} + 1/2$. We assume that $2\rho\beta > 1$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\left]0, \delta\right[$ such that $\sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$.
We assume that $\mathrm{zer}\left(A + C + \sum_{i=1}^m L_i^*(B_i \,\square\, D_i)L_i\right) \neq \varnothing$.
Let $x_0 \in \mathcal{H}$, $(v_{0,1}, \ldots, v_{0,m}) \in \mathcal{G}_1 \times \cdots \times \mathcal{G}_m$ and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = J_{\tau A}\big(x_n - \tau(Cx_n + \sum_{i=1}^m L_i^* v_{n,i})\big) \\ q_{n,i} = J_{\sigma B_i^{-1}}\big(v_{n,i} + \sigma\big(L_i(2p_n - x_n) - D_i^{-1} v_{n,i}\big)\big), \; i \in \{1, \ldots, m\} \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ v_{n+1,i} = v_{n,i} + \lambda_n(q_{n,i} - v_{n,i}), \; i \in \{1, \ldots, m\}. \end{cases}$$

We have: $x_n \rightharpoonup \widehat{x} \in \mathrm{zer}\left(A + C + \sum_{i=1}^m L_i^*(B_i \,\square\, D_i)L_i\right)$
and $v_{n,i} \rightharpoonup \widehat{v}_i \in \mathrm{zer}\left((-L_i)(A^{-1} \,\square\, C^{-1})(-L_i^*) + B_i^{-1} + D_i^{-1}\right)$, $i \in \{1, \ldots, m\}$.

Primal-dual method

M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two maximally monotone operators.

Let $C\colon \mathcal{H} \to \mathcal{H}$ be a monotone operator μ^{-1}-Lipschitzian with $\mu \in \,]0,+\infty[$,

$D\colon \mathcal{G} \to 2^{\mathcal{G}}$ be ν-strongly monotone where $\nu \in \,]0,+\infty[$ and $\beta^{-1} = \max\{\mu^{-1},\nu^{-1}\} + \|L\|$.

$$
(\forall n \in \mathbb{N}) \qquad
\begin{cases}
y_{1,n} = x_n - \gamma_n(Cx_n + L^* v_n)\\
y_{2,n} = v_n + \gamma_n(Lx_n - D^{-1}v_n)\\
p_{1,n} = J_{\gamma_n A} y_{1,n},\, p_{2,n} = J_{\gamma_n B^{-1}} y_{2,n}\\
q_{1,n} = p_{1,n} - \gamma_n(Cp_{1,n} + L^* p_{2,n})\\
q_{2,n} = p_{2,n} + \gamma_n(Lp_{1,n} - D^{-1}p_{2,n})\\
(x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n},\, v_n - y_{2,n} + q_{2,n}).
\end{cases}
$$

Primal-dual method

M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ and $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ be two maximally monotone operators.

Let $C\colon \mathcal{H} \to \mathcal{H}$ be a monotone operator μ^{-1}-Lipschitzian with $\mu \in \,]0, +\infty[$,

$D\colon \mathcal{G} \to 2^{\mathcal{G}}$ be ν-strongly monotone where $\nu \in \,]0, +\infty[$ and $\beta^{-1} = \max\{\mu^{-1}, \nu^{-1}\} + \|L\|$.

Let $\varepsilon \in]0, \beta/(1+\beta)[$ and $(\gamma_n)_{n\in\mathbb{N}}$ a sequence in $[\varepsilon, (1-\epsilon)\beta]$.

We assume $\operatorname{zer}\big(A + C + L^*(B \,\square\, D)L\big) \neq \varnothing$. Let $x_0 \in \mathcal{H}$, $v_0 \in \mathcal{G}$ and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{1,n} = x_n - \gamma_n(Cx_n + L^* v_n) \\ y_{2,n} = v_n + \gamma_n(Lx_n - D^{-1} v_n) \\ p_{1,n} = J_{\gamma_n A} y_{1,n}, \, p_{2,n} = J_{\gamma_n B^{-1}} y_{2,n} \\ q_{1,n} = p_{1,n} - \gamma_n(Cp_{1,n} + L^* p_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma_n(Lp_{1,n} - D^{-1} p_{2,n}) \\ (x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n}, v_n - y_{2,n} + q_{2,n}). \end{cases}$$

We have: $x_n \rightharpoonup \widehat{x}$ and $p_{1,n} \rightharpoonup \widehat{x}$ where $\widehat{x} \in \operatorname{zer}\big(A + C + L^*(B \,\square\, D)L\big)$
and $v_n \rightharpoonup \widehat{v}$ and $p_{2,n} \rightharpoonup \widehat{v}$ where $\widehat{v} \in \operatorname{zer}\big((-L)(A^{-1} \,\square\, C^{-1})(-L^*) + B^{-1} + D^{-1}\big)$.

Primal-dual optimization method

M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$.

Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.

Let $h \in \Gamma_0(\mathcal{H})$ having a μ^{-1}-Lipschitz gradient with $\mu \in\,]0,+\infty[$,

$\ell \in \Gamma_0(\mathcal{G})$ ν-strongly convex

$$(\forall n \in \mathbb{N}) \qquad \begin{cases} y_{1,n} = x_n - \gamma_n\big(\nabla h(x_n) + L^* v_n\big) \\ y_{2,n} = v_n + \gamma_n\big(L x_n - \nabla \ell^*(v_n)\big) \\ p_{1,n} = \mathrm{prox}_{\gamma_n f}\, y_{1,n},\; p_{2,n} = \mathrm{prox}_{\gamma_n g^*}\, y_{2,n} \\ q_{1,n} = p_{1,n} - \gamma_n\big(\nabla h(p_{1,n}) + L^* p_{2,n}\big) \\ q_{2,n} = p_{2,n} + \gamma_n\big(L p_{1,n} - \nabla \ell^*(p_{2,n})\big) \\ (x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n},\, v_n - y_{2,n} + q_{2,n}). \end{cases}$$

Primal-dual optimization method

M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let $\mathcal{H}$ and $\mathcal{G}$ be two Hilbert spaces. Let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$.

Let $f \in \Gamma_0(\mathcal{H})$ and $g \in \Gamma_0(\mathcal{G})$.

Let $h \in \Gamma_0(\mathcal{H})$ having a μ^{-1}-Lipschitz gradient with $\mu \in\ ]0, +\infty[$,

$\ell \in \Gamma_0(\mathcal{G})$ ν-strongly convex where $\nu \in\ ]0, +\infty[$ and $\beta^{-1} = \max\{\mu^{-1}, \nu^{-1}\} + \|L\|$.

Let $\varepsilon \in]0, \beta/(1+\beta)[$ and $(\gamma_n)_{n\in\mathbb{N}}$ a sequence in $[\varepsilon, (1-\epsilon)\beta]$.

We assume that $\mathrm{Argmin}\big(f + g + (h \,\square\, \ell) \circ L\big) \neq \varnothing$. Lett $x_0 \in \mathcal{H}$, $v_0 \in \mathcal{G}$ and

$$
(\forall n \in \mathbb{N}) \quad
\begin{cases}
y_{1,n} = x_n - \gamma_n\big(\nabla h(x_n) + L^* v_n\big)\\
y_{2,n} = v_n + \gamma_n\big(L x_n - \nabla \ell^*(v_n)\big)\\
p_{1,n} = \mathrm{prox}_{\gamma_n f}\, y_{1,n},\ p_{2,n} = \mathrm{prox}_{\gamma_n g^*}\, y_{2,n}\\
q_{1,n} = p_{1,n} - \gamma_n\big(\nabla h(p_{1,n}) + L^* p_{2,n}\big)\\
q_{2,n} = p_{2,n} + \gamma_n\big(L p_{1,n} - \nabla \ell^*(p_{2,n})\big)\\
(x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n}, v_n - y_{2,n} + q_{2,n}).
\end{cases}
$$

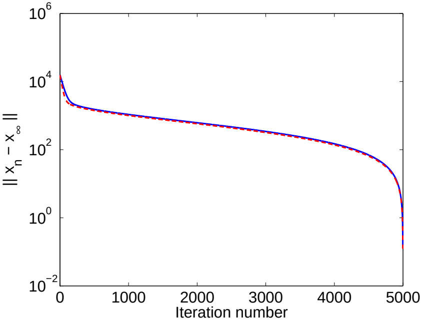
We have: $x_n \rightharpoonup \widehat{x}$ and $p_{1,n} \rightharpoonup \widehat{x}$ where $\widehat{x} \in \mathrm{Argmin}\big(f + h + (g \,\square\, h) \circ L\big)$

and $v_n \rightharpoonup \widehat{v}$ et $p_{2,n} \rightharpoonup \widehat{v}$ where $\widehat{v} \in \mathrm{Argmin}\big((f^* \,\square\, h^*) \circ (-L^*) + g^* + \ell^*\big)$.

Primal-dual optimization method

Image restoration : Variational approach

$$\widehat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|[W_1^* H_1^* \ldots W_J^* H_J^*]^* x\|_{2,1} + \iota_C(x)$$



→ FB-based PD (red)

→ M+LFBF (blue)

Conclusions

- Flexible framework unifying several problems:
  - regularized approaches

$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; f(x) + \sum_{i=1}^{m} g_i(x)$$

where $f\colon \mathcal{H} \to \mathbb{R}$, $(\forall i \in \{1,\ldots,m\})$ $g_i\colon \mathcal{H} \to \mathbb{R}$ are convex.

  - feasibility approaches

$$\text{Find } x \in \mathcal{H} \text{ such that } \quad x \in \bigcap_{i=1}^{m} C_i$$

where $(\forall i \in \{1,\ldots,m\})$ C_i is a nonempty closed convex subset of $\mathcal{H}$.

Conclusions

- Flexible framework unifying several problems:
  - sparse problems
$$\underset{x\in\mathcal{H}}{\text{minimize}} \;\; f(x) + \sigma_C(x)$$
where $f \in \Gamma_0(\mathcal{H})$ and C is a nonempty closed convex subset of $\mathcal{H}$.
  - constrained problems
$$\underset{x\in C}{\text{minimize}} \;\; g(x)$$
where $g \in \Gamma_0(\mathcal{H})$.

Conclusions

- Results in infinite dimension (continuous problems).
- Splitting is the key.
  Parallel methods adapted to multi-core architectures.
  Can be extended to distributed/stochastic methods.
- Robustness to computational errors.

$$\mathrm{prox}_f x_n \to \mathrm{prox}_f x_n + e_n$$

  where $(e_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{H}$ such that $\sum_{n\in\mathbb{N}} \|e_n\| < +\infty$.
- Applications to other fields: game theory, PDEs,...
  Find maximally monotone operators in your favorite area !
- Extension to the nonconvex case.

A few references

P.L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing", in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer-Verlag, pp. 185–212, 2010.

F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties", *Foundations and Trends in Machine Learning*, vol. 4, no. 1, pp. 1–106, 2012.

N. Parikh and S. Boyd, "Proximal algorithms", *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 123-231, 2013.