Inverse problem and optimization

Nelly PUSTELNIK
CNRS, Laboratoire de Physique de l'ENS Lyon
Laurent CONDAT
CNRS, Gipsa-lab

Course 5 : Nonexpansive operators
January, 19th 2017

Naive answer

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

If

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

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

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

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

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

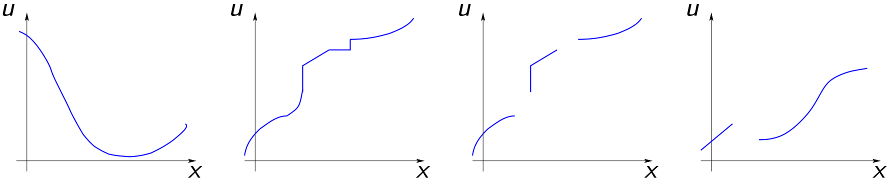
Consequently, $\|x_n - \widehat{x}\| \leq \rho^n \|x_0 - \widehat{x}\|$. Hence, we have proved that $(x_n)_{n\in\mathbb{N}}$ converges linearly to $\widehat{x}$.

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 ?

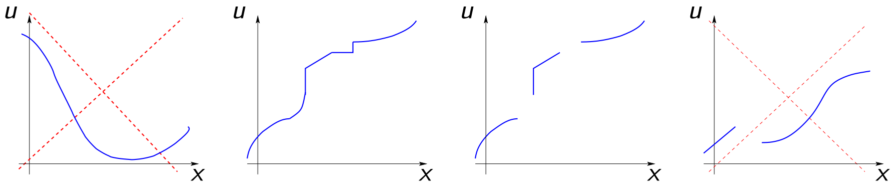


Monotone operator: definition

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 ?

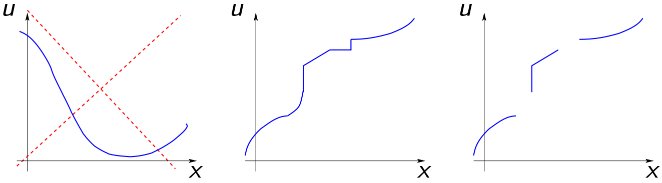


Monotone operator: definition

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 ?

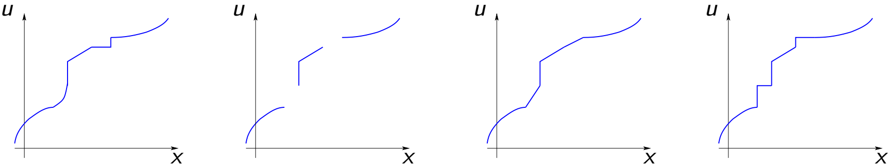


- Example: subdifferential of a convex and proper function

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 $\mathrm{gra} A$.

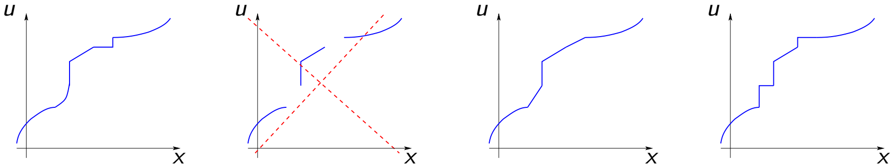
- Maximally monotone operator ?



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 $\mathrm{gra} B$ properly contains $\mathrm{gra} A$.

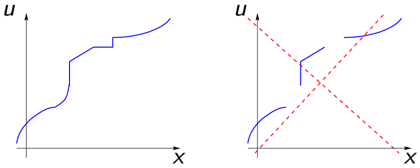
- Maximally monotone operator ?



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 $\mathrm{gra}B$ properly contains $\mathrm{gra}A$.

- Maximally monotone operator ?



- Example: subdifferential of a convex, proper and l.s.c. function.

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 $\big(\forall (x,y) \in C^2\big) \quad \|Ax - Ay\| \leq \|x - y\|$.

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 $\nu \in \left]0, +\infty\right[$

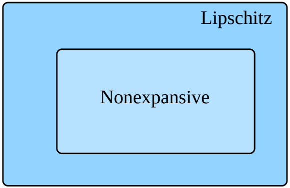
$\nu^{-1}A$ is nonexpansive if $\big(\forall(x,y) \in C^2\big) \quad \|Ax - Ay\| \leq \nu \|x - y\|$.

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

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

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

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

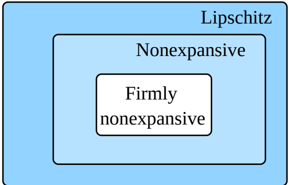
- A is firmly nonexpansive $\Leftrightarrow$ $\mathrm{Id} - A$ is firmly nonexpansive.
- A is firmly nonexpansive $\Leftrightarrow$ $\underbrace{2A - \mathrm{Id}}_{\text{Reflection of A}}$ is nonexpansive.

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 $A : C \to \mathcal{H}$ 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

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

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

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

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

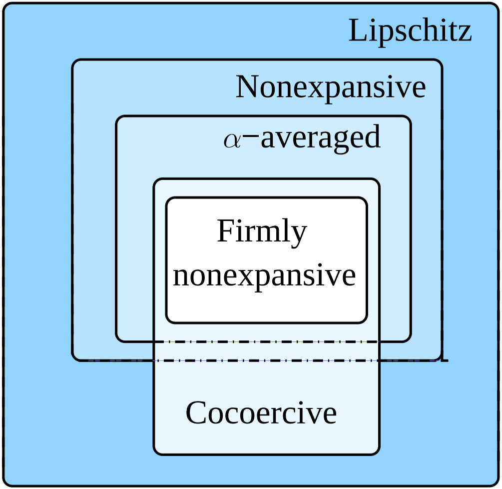
Nonexpansive operator: definition

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

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

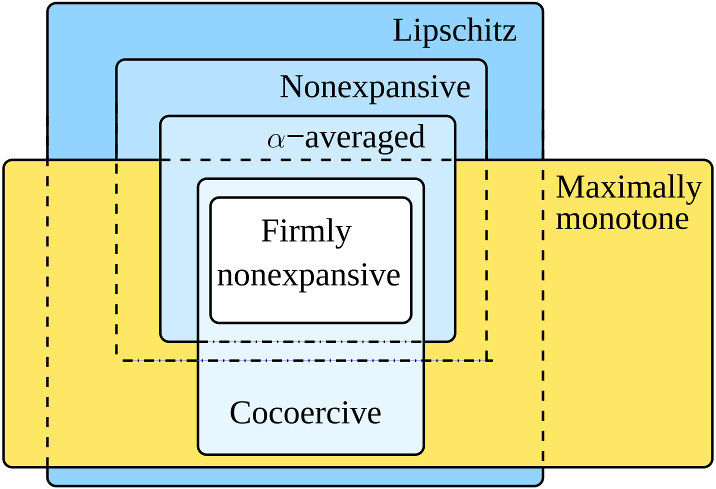
$A : \mathcal{H} \to \mathcal{H}$ is α-averaged with $\alpha \in]0,1/2] \Rightarrow A$ is maximally monotone.

Nonexpansive operator: recap



Nonexpansive operator: recap

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



Nonexpansive operators

Nonexpansive operators → What is their use ?

Nonexpansive operator: example

Descent lemma

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

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

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

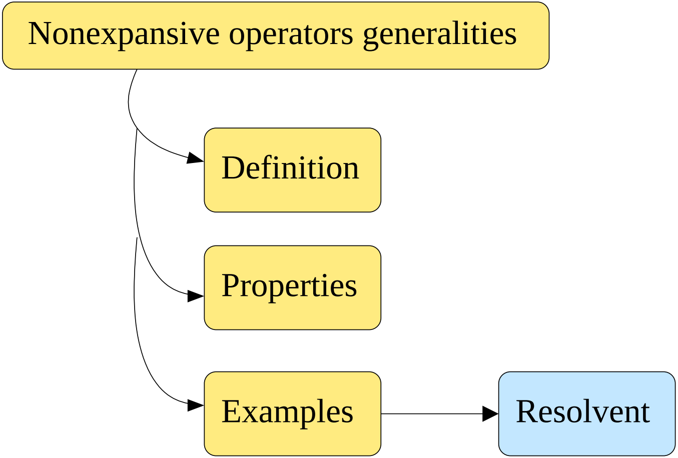
Nonexpansive operator: example

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 its gradient is ν-Lipschitzian, then

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

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 operators generalities
Definition
Properties
Examples
Resolvent

Monotone operator: inversion

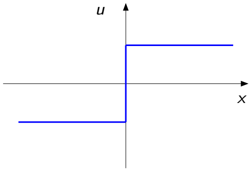
Let $\mathcal{H}$ be a Hilbert space.
Let $A\colon \mathcal{H} \to 2^{\mathcal{H}}$.
A^{-1} is the operator from $\mathcal{H}$ to $2^{\mathcal{H}}$ the graph of which is

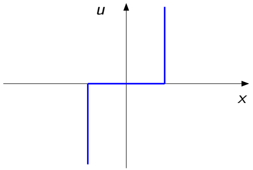
$$\mathrm{gra}(A^{-1}) = \{(u,x) \mid (x,u) \in \mathrm{gra}A\}.$$

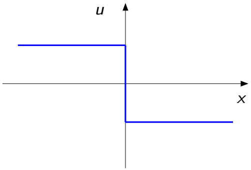
with $\mathrm{gra}A = \{(x,u) \in \mathcal{H}^2 \mid u \in Ax\}$.

Graph of A

Graph of A^{-1} ?







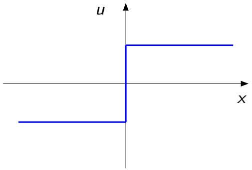
Monotone operator: inversion

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

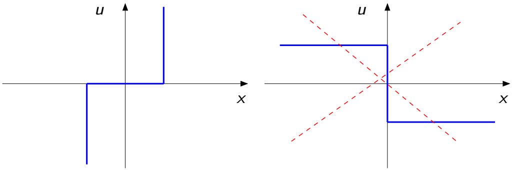
$$\mathrm{gra}(A^{-1}) = \{(u,x) \mid (x,u) \in \mathrm{gra}A\}.$$

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

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

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

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

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

Resolvent: definition

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

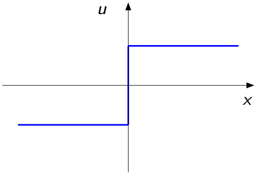
$$J_A = (\mathrm{Id} + A)^{-1}.$$

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 :



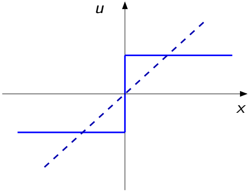
A $\qquad$ $A + \mathrm{Id}$? $\qquad$ J_A ?

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 :



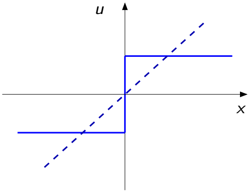
A and Id $\qquad$ $A + \mathrm{Id}$? $\qquad$ J_A ?

Resolvent: definition

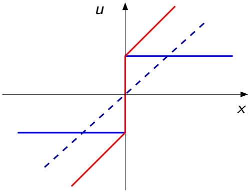
Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

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

- Example :



A and Id



$A + \mathrm{Id}$

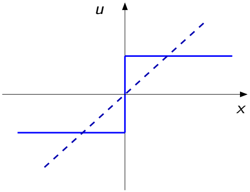
J_A ?

Resolvent: definition

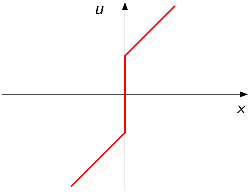
Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is

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

- Example :



A and Id



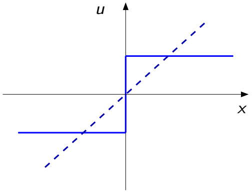
$A + \mathrm{Id}$

J_A ?

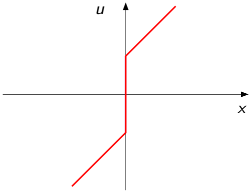
Resolvent: definition

Let $\mathcal{H}$ be a Hilbert space.
Let $A : \mathcal{H} \to 2^{\mathcal{H}}$.
The revolvent of A is
$$J_A = (\mathrm{Id} + A)^{-1}.$$

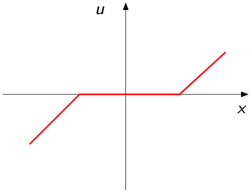
- Example :



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.

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

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.

Proof: A monotone $\Leftrightarrow$ $J_A \colon \operatorname{ran}(\mathrm{Id} + A) \to \mathcal{H}$ 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 $x - p \in \gamma Ap$ and thus $p = J_{\gamma A}x$.

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

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.

Resolvent

Resolvent

Proximity operator

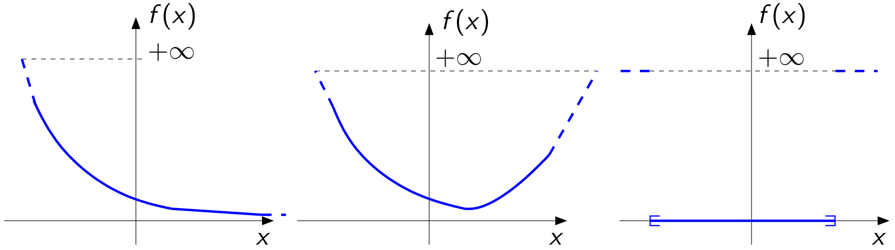
Convex analysis

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

Convex analysis

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

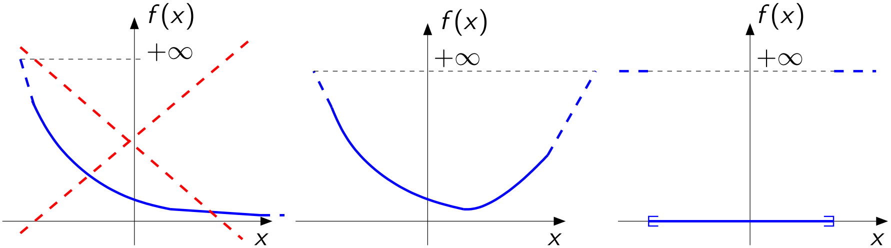
Coercive functions ?



Convex analysis

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

Coercive functions ?



Convex analysis

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

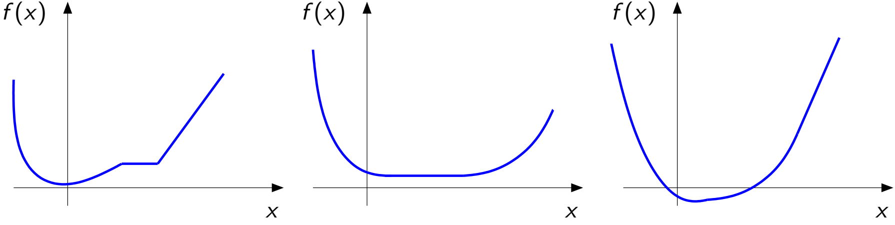
$$(\forall x \in \operatorname{dom} f)(\forall y \in \operatorname{dom} f)(\forall \alpha \in]0,1[)$$
$$x \neq y \quad \Rightarrow \quad f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y).$$

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 ?

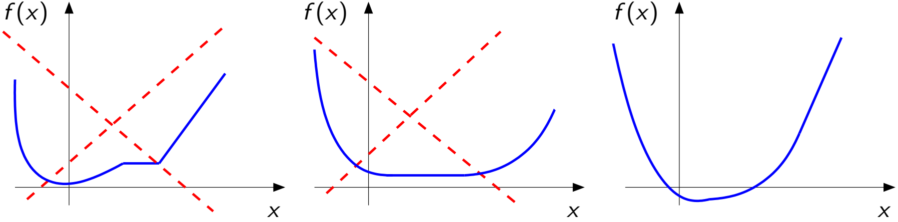


Convex analysis

Let $\mathcal{H}$ be a Hilbert space. 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 $\operatorname{dom} f \cap C \neq \varnothing$.
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

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

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

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

Proximity operator: definition

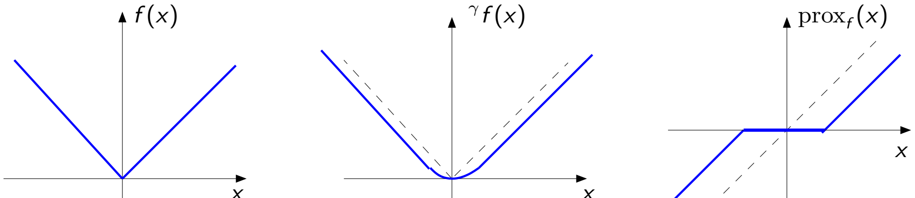
Let $\mathcal{H}$ be a Hilbert space. Let $f \in \Gamma_0(\mathcal{H})$.

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

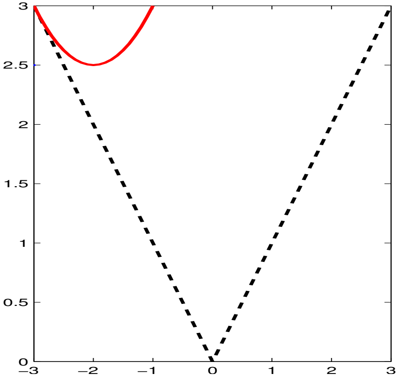
$$^{\gamma}f\colon \mathcal{H} \to \mathbb{R}\colon x \mapsto \inf_{y\in\mathcal{H}} f(y) + \frac{1}{2\gamma}\|y-x\|^2.$$

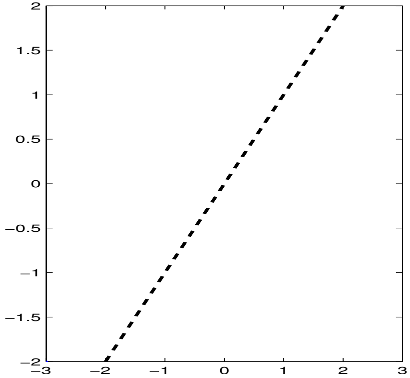
- The **proximity operator** of f is

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

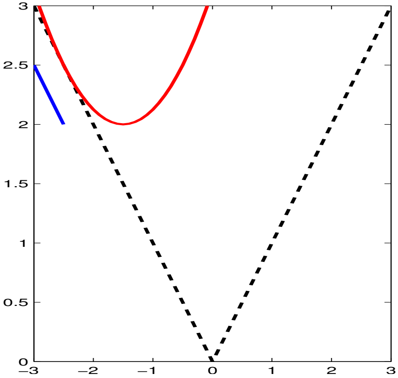


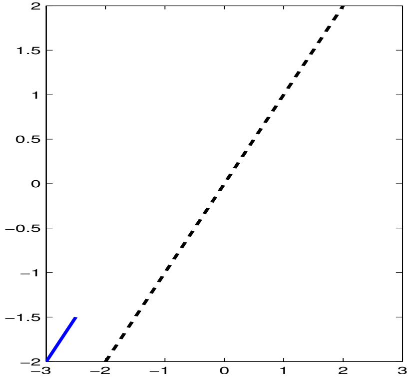
Proximity operator: definition



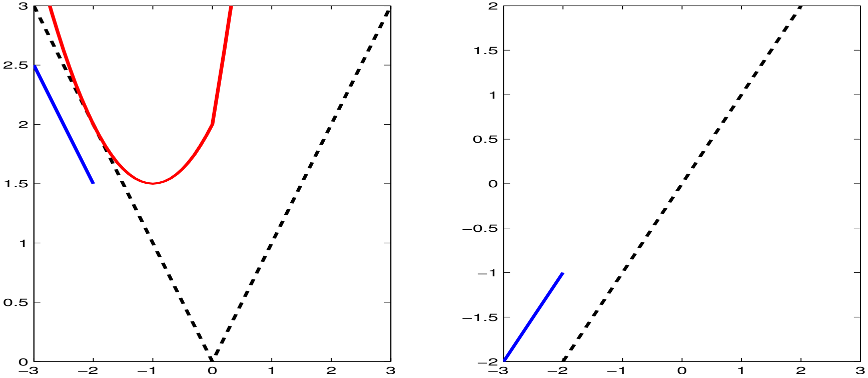


Proximity operator: definition

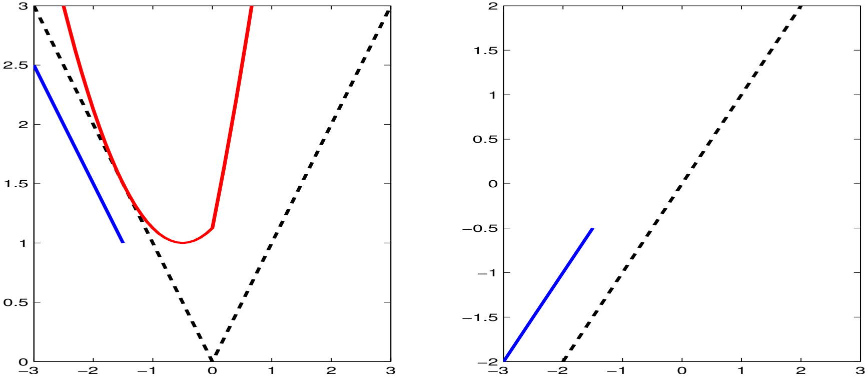




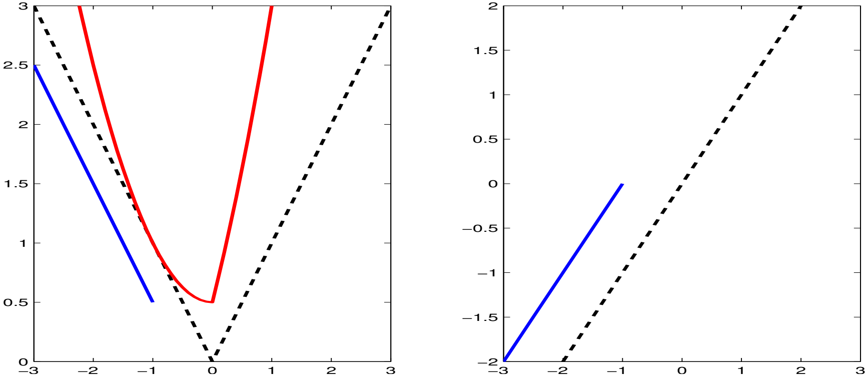
Proximity operator: definition



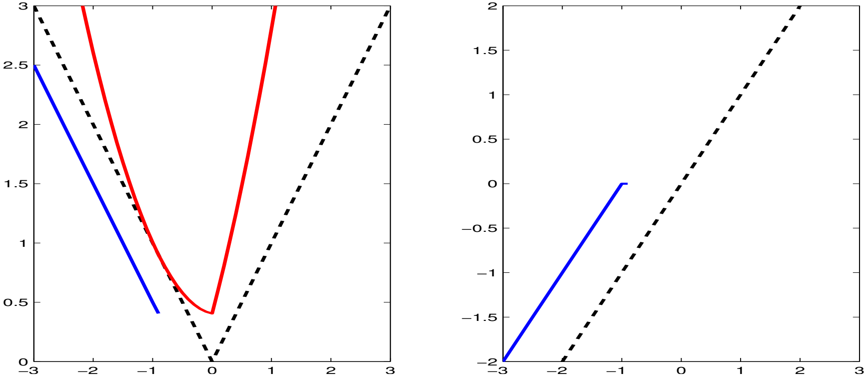
Proximity operator: definition



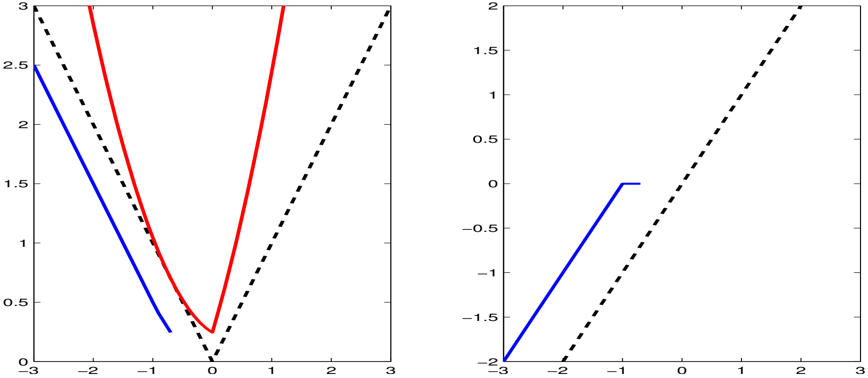
Proximity operator: definition



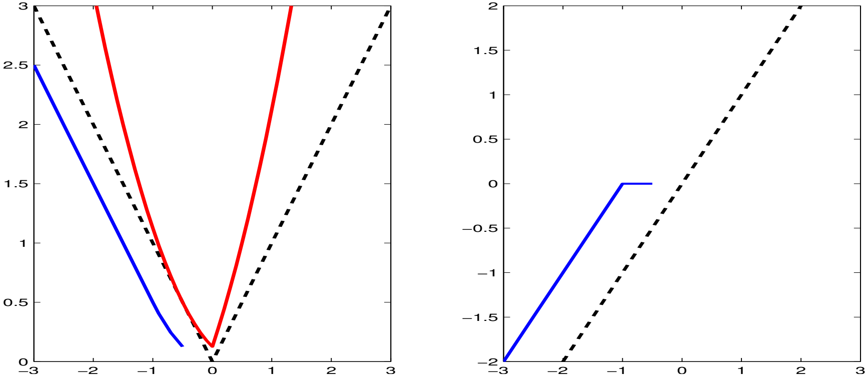
Proximity operator: definition



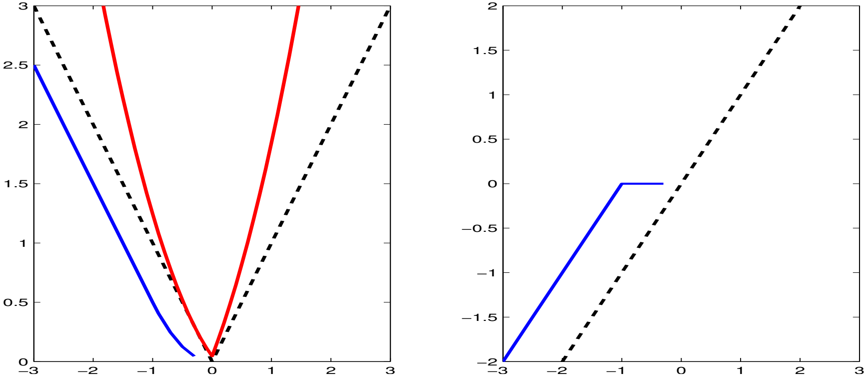
Proximity operator: definition



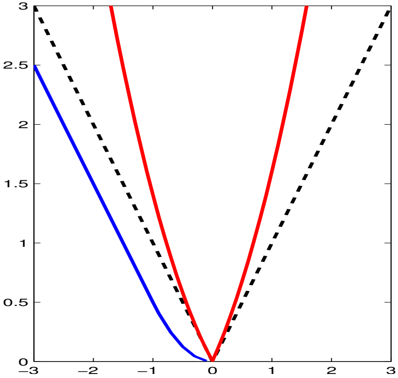
Proximity operator: definition

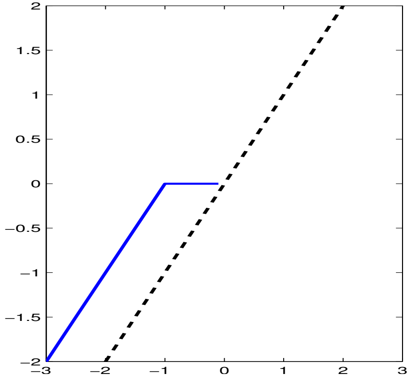


Proximity operator: definition

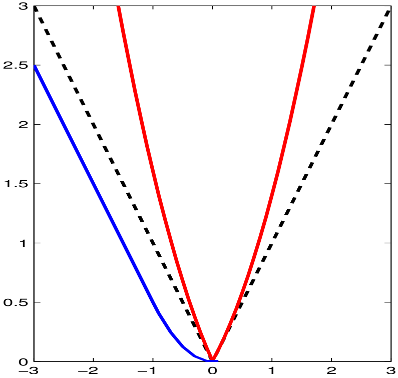


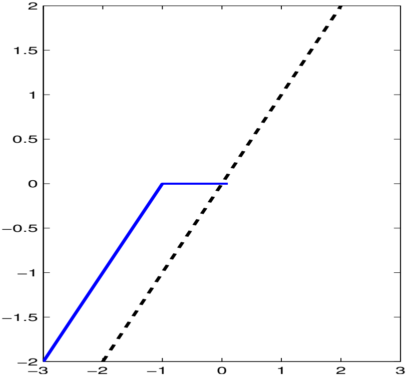
Proximity operator: definition



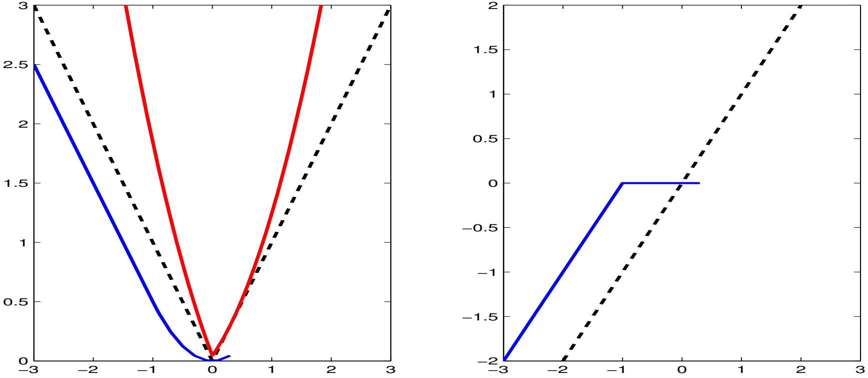


Proximity operator: definition

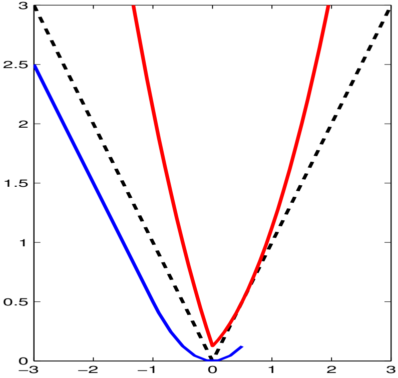


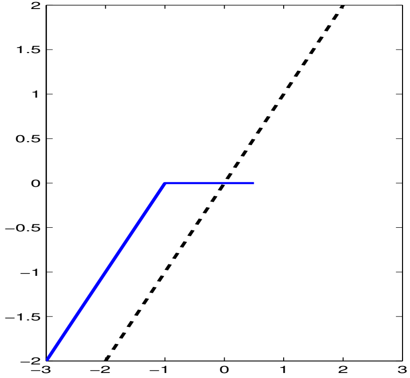


Proximity operator: definition

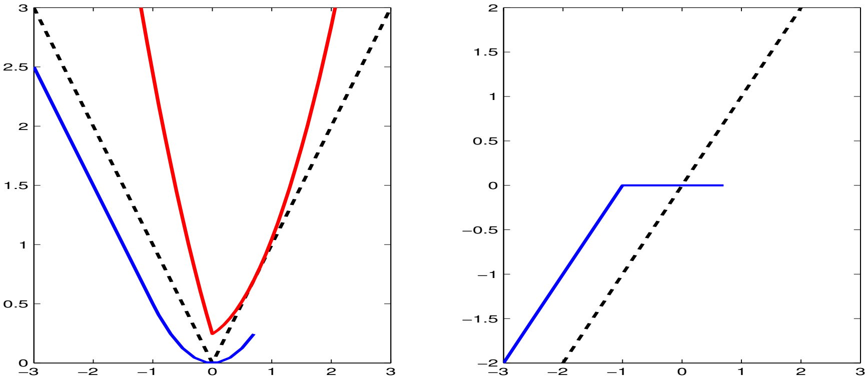


Proximity operator: definition

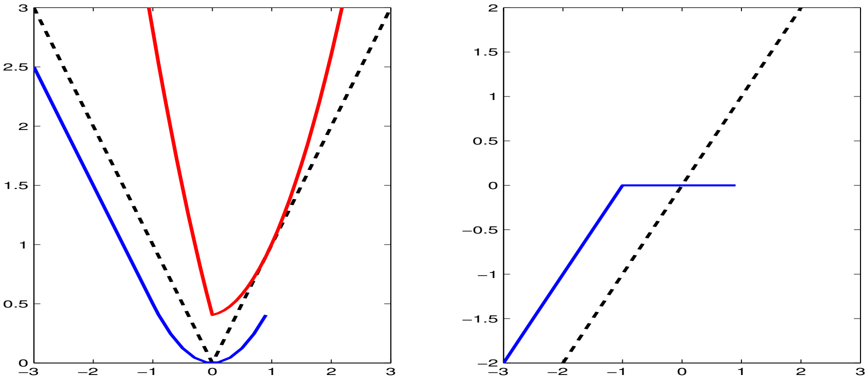




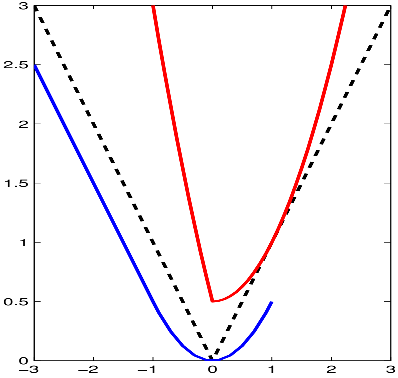
Proximity operator: definition

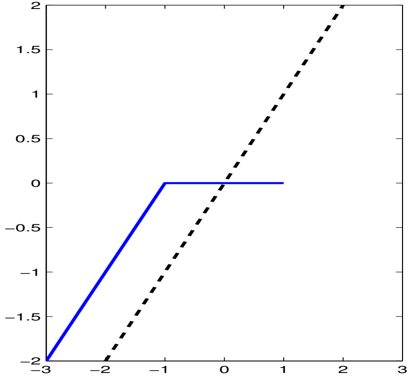


Proximity operator: definition

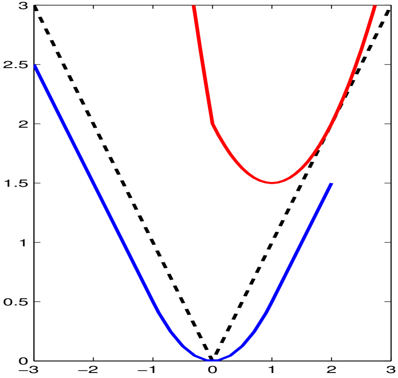


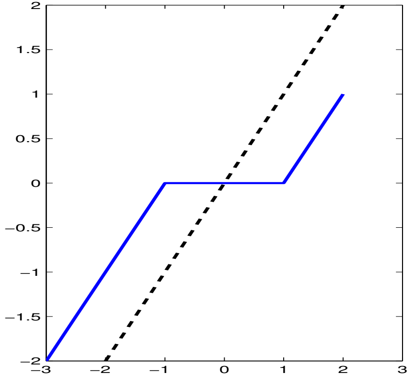
Proximity operator: definition



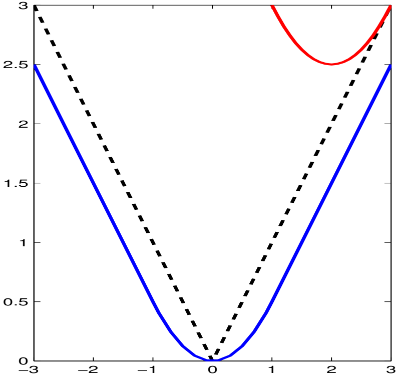


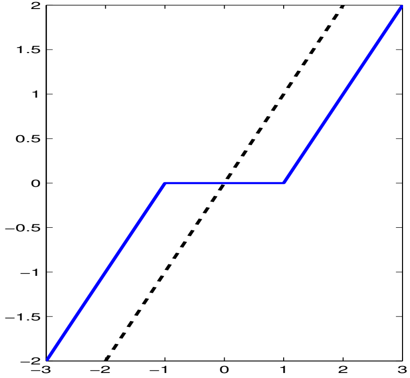
Proximity operator: definition





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 $\mathrm{dom}\, f \cap \mathrm{int}\,(\mathrm{dom}\, g) \neq \varnothing$ then $\partial(f+g) = \partial f + \partial g$.

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

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

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

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

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 $(x, p) \in \mathcal{H}^2$.

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

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

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

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

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

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

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

then

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

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

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

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

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

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

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

Proximity operator: properties

Moreau decomposition formula

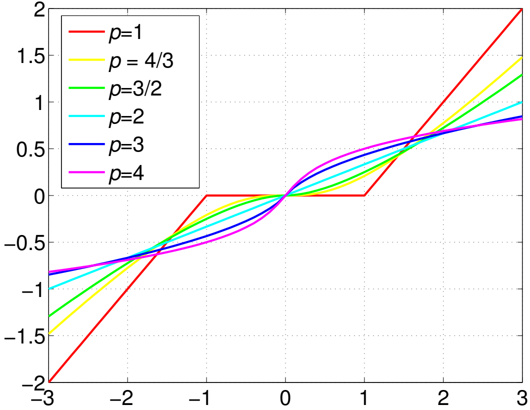
Let $\mathcal{H}$ be a Hilbert space, $f \in \Gamma_0(\mathcal{H})$ and $\gamma \in \,]0,+\infty[$.

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

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

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

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 $\operatorname{ran} L = \mathcal{H}$. 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 $\mathrm{ran}\, L = \mathcal{H}$. 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

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

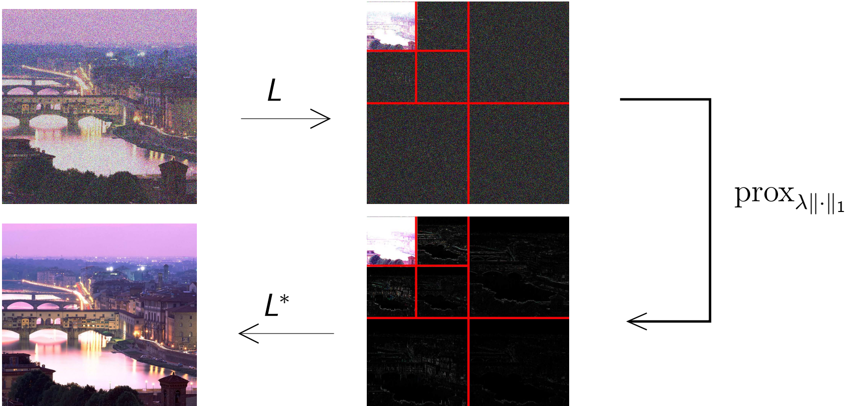
Remark :

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

Proximity operator: properties

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

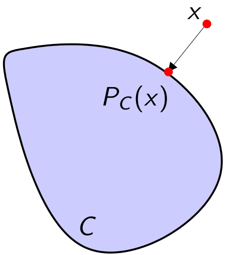
- Illustration: denoising using an ℓ_1 penalty on the coefficients resulting from an orthogonal wavelet transform L.



Proximity operator: examples

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

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



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

ℓ_2-norm :

Let $f \in \Gamma_0(\mathcal{H})$ such that $f = \gamma\| \cdot \|_2$, its proximity operator is

$$(\forall x \in \mathcal{H}) \quad \mathrm{prox}_{\gamma f} x = \max\left(0, \frac{1-\gamma}{\|x\|_2}\right) x$$

Some take-home messages

- Gradient descent is α-averaged.
- $\mathrm{prox}_f = J_{\partial f}$ with $f \in \Gamma_0(\mathcal{H})$ is firmly non-expansive, thus α-averaged.
- The reflected resolvent is nonexpansive.
- Closed form expressions form several functions.
- Next course: design algorithms.