Numerical Optimization Methods in Imaging

Part II: Fixed point methods and nonexpansive operators

jean-christophe@pesquet.eu

PhD Summer School MMLIA – Bologna

A first answer

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

Let $\mathcal{H}$ be a Hilbert space. If $T\colon \mathcal{H} \to \mathcal{H}$ is such that

> T is a strict contraction, i.e. there exists $\rho \in [0,1[$ such that
> $$\big(\forall (x,x') \in \mathcal{H}^2\big) \qquad \|Tx - Tx'\| \leq \rho \|x - x'\|.$$

Then T has a unique fixed point $\widehat{x}$.
The sequence $(x_n)_{n\in\mathbb{N}}$ defined as $(\forall n \in \mathbb{N})$ $x_{n+1} = Tx_n$ with $x_0 \in \mathcal{H}$, converges to $\widehat{x}$.

Objective of the remainder of this course

- Extend this theorem to more general operators
  - not necessarily *strictly* contractive
  - possibly dependent on the iteration number n
  - built from composition of simpler operators (*splitting techniques*).

- Apply this to solve minimization problems.
  $\rightsquigarrow$ How to relate T to the objective function f ?

Fixed point algorithms

Fixed point algorithms: 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 algorithms: convergence

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
- $(x_n)_{n\in\mathbb{N}}$ possesses 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 subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ that converges weakly to $\widehat{x}$.

Fixed point algorithms: convergence

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

<u>Illustration</u>:

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

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

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

Fixed point algorithms: Fejér-monotone sequence

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

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

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

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

Fixed point algorithms: Fejér-monotone sequence

Fejér-monotone convergence

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

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

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

Nonexpansive operator: definition

Let C be a nonempty set of a Hilbert space $\mathcal{H}$. Let $T\colon C \to \mathcal{H}$.
The set of fixed points of T is

$$\operatorname{Fix} T = \{x \in C \mid x = Tx\}.$$

Nonexpansive operator: definition

Let C be a nonempty set of a Hilbert space $\mathcal{H}$. Let $T\colon C \to \mathcal{H}$.
The set of fixed points of T is

$$\operatorname{Fix} T = \{x \in C \mid x = Tx\}.$$

Let $C \subset \mathcal{H}$ be a nonempty set.
Let $T\colon C \to \mathcal{H}$.
T is a nonexpansive operator if $\big(\forall (x,y) \in C^2\big) \quad \|Tx - Ty\| \leq \|x - y\|$.

Nonexpansive operator: fixed point algorithm

Demiclosedness principle

Let C be a nonempty closed convex subset of a Hilbert space $\mathcal{H}$.
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 $\widehat{x} \in \operatorname{Fix} T$.

Nonexpansive operator: fixed point algorithm

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

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

Nonexpansive operator: fixed point algorithm

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

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

If $x_n - Tx_n \to 0$,

Nonexpansive operator: fixed point algorithm

Let C be a nonempty closed convex subset of a Hilbert space $\mathcal{H}$.
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 $\operatorname{Fix} T$.

Exercise :

Prove this result by showing that $(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to $\operatorname{Fix} T$.

Nonexpansive operator: fixed point algorithm

Let C be a nonempty closed convex subset of a Hilbert space $\mathcal{H}$.
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 $\operatorname{Fix} T$.

Answer : For every $n \in \mathbb{N}$ and $y \in \operatorname{Fix} T$,
$\|x_{n+1} - y\| = \|Tx_n - Ty\| \leq \|x_n - y\|$.
$(x_n)_{n\in\mathbb{N}}$ is Fejér-monotone with respect to $\operatorname{Fix} T$.
Let $(x_{n_k})_{k\in\mathbb{N}}$ be a subsequence of $(x_n)_{n\in\mathbb{N}}$ such that $x_{n_k} \rightharpoonup \widehat{x}$ where $\widehat{x} \in \mathcal{H}$.
By assumption $x_{n_k} - Tx_{n_k} \to 0$ and thus, according to the demiclosedness principle, $\widehat{x} \in \operatorname{Fix} T$.
This shows the weak convergence of $(x_n)_{n\in\mathbb{N}}$.

Fixed point algorithms: Fejér-monotone sequence

Красносе́льский–Mann algorithm

Let C be a nonempty closed convex subset of a Hilbert space $\mathcal{H}$.
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 $(\forall n \in \mathbb{N})\ x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$. Then,

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

Typical choice: $(\forall n \in \mathbb{N})\ \lambda_n = \lambda \in]0,1[$.

Fixed point algorithms: Fejér-monotone sequence

Красносéльский–Mann algorithm

Let C be a nonempty closed convex subset of a Hilbert space $\mathcal{H}$.
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 $(\forall n \in \mathbb{N})\ x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$. Then,

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

Proof: similar to the previous one.

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

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

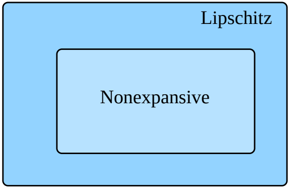
$\nu^{-1}A$ is nonexpansive if $\left(\forall (x, y) \in C^2\right) \quad \|Ax - Ay\| \leq \nu \|x - y\|$.

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

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

Proof: For every $(x,y) \in C^2$,

$$\begin{aligned} & \|Ax - Ay\|^2 + \|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x - y\|^2 \\ \Leftrightarrow & \|Ax - Ay\|^2 + \|x - y\|^2 - 2\langle x - y \mid Ax - Ay\rangle + \|Ax - Ay\|^2 \leq \|x - y\|^2 \\ \Leftrightarrow & \|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle . \end{aligned}$$

Nonexpansive operator: definition

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

  Proof: For every $(x,y) \in C^2$,

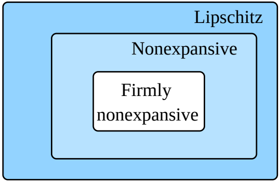
$$\begin{aligned} & \|(2A - \mathrm{Id})x - (2A - \mathrm{Id})y\|^2 \leq \|x - y\|^2 \\ \Leftrightarrow \quad & 4\|Ax - Ay\|^2 - 4\langle Ax - Ay \mid x - y\rangle + \|x - y\|^2 \leq \|x - y\|^2 \\ \Leftrightarrow \quad & \|Ax - Ay\|^2 \leq \langle Ax - Ay \mid x - y\rangle . \end{aligned}$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a real 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 real 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 real 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 real Hilbert spaces, $L \in \mathcal{B}(\mathcal{H},\mathcal{G})$ nonzero, and $A\colon \mathcal{G} \to \mathcal{G}$. A is β-cocoercive $\Rightarrow$ L^*AL is $\|L\|^{-2}\beta$-cocoercive.
  Proof: For every $(x,y) \in \mathcal{H}^2$,
  $$\begin{aligned}\langle L^*ALx - L^*ALy \mid x - y\rangle &= \langle ALx - ALy \mid Lx - Ly\rangle\\ &\geq \beta\|ALx - ALy\|^2\end{aligned}$$
  Furthermore, $\|L^*ALx - L^*ALy\|^2 \leq \|L\|^2\|ALx - ALy\|^2$.
  Then $\langle L^*ALx - L^*ALy \mid x - y\rangle \geq \beta\|L^*ALx - L^*ALy\|^2/\|L\|^2$.

Nonexpansive operator: definition

Let $\mathcal{H}$ be a real 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.,

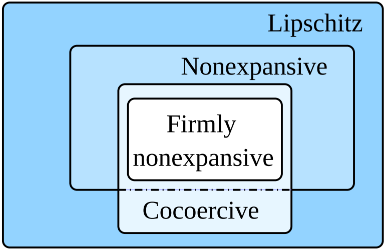
$$(\forall x \in C)(\forall y \in C) \quad \beta\|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay\rangle .$$

- Let $\mathcal{H}$ and $\mathcal{G}$ be two real Hilbert spaces, $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ nonzero, and $A\colon \mathcal{G} \to \mathcal{G}$. 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 real 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 real 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 real 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 real 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

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

Proof: For every $(x,y) \in C^2$,

$$\begin{aligned}
& \|Ax - Ay\|^2 + \frac{1-\alpha}{\alpha}\|(\mathrm{Id} - A)x - (\mathrm{Id} - A)y\|^2 \leq \|x-y\|^2 \\
\Leftrightarrow \quad & \alpha\|Ax - Ay\|^2 + (1-\alpha)\big(\|Ax - Ay\|^2 - 2\langle x-y \mid Ax - Ay\rangle + \|x-y\|^2\big) \\
& \qquad \leq \alpha\|x-y\|^2 \\
\Leftrightarrow \quad & \|Ax - Ay\|^2 - 2(1-\alpha)\langle x-y \mid Ax - Ay\rangle + (1-2\alpha)\|x-y\|^2 \leq 0 \\
\Leftrightarrow \quad & \|Ax - Ay - (1-\alpha)(x-y)\|^2 \leq \alpha^2\|x-y\|^2 \\
\Leftrightarrow \quad & R = \frac{A - (1-\alpha)\mathrm{Id}}{\alpha} \text{ nonexpansive.}
\end{aligned}$$

Nonexpansive operator: definition

Let $\mathcal{H}$ be a real 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 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 real 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.
  Proof: If A is α-averaged, there exists a nonexpansive operator R such that $A = (1-\alpha)\mathrm{Id} + \alpha R$. We have thus

$$\begin{aligned}(1-\lambda)\mathrm{Id} + \lambda A =& (1-\lambda)\mathrm{Id} + \lambda((1-\alpha)\mathrm{Id} + \alpha R) \\ =& (1-\lambda\alpha)\mathrm{Id} + \lambda\alpha R.\end{aligned}$$

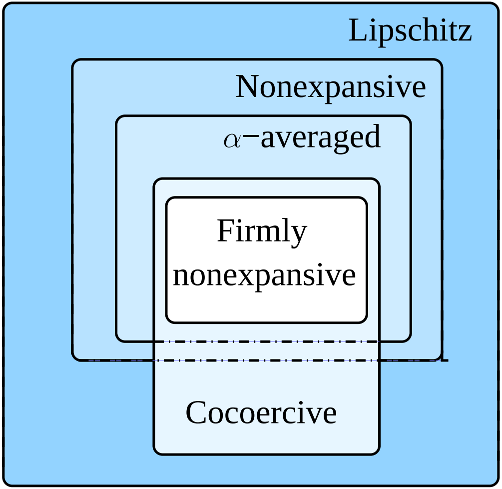
Nonexpansive operator: definition

Let $\mathcal{H}$ be a real 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

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

- 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 $i \in \{1,\ldots,n\}$, $A_i \colon C \to \mathcal{H}$ is α_i-averaged, then $\sum_{i=1}^n \omega_i A_i$ is α-averaged with $\alpha = \sum_{i=1}^m \omega_i \alpha_i$.
- Let $(\alpha_i)_{1\leq i\leq n} \in]0,1[^n$. If, for every $i \in \{1,\ldots,n\}$, $A_i \colon C \to C$ is α_i-averaged, then $A_1 \cdots A_n$ is α-averaged with $\alpha = \dfrac{1}{1 + 1/(\sum_{i=1}^n \frac{\alpha_i}{1-\alpha_i})}$.

Nonexpansive operator: recap



Nonexpansive operator: main property

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

Fixed point algorithms: α-averaged operator

Let $T\colon \mathcal{H} \to \mathcal{H}$ be an α-averaged operator with $\alpha \in]0,1[$ such that $\operatorname{Fix} T \neq \varnothing$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be 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 $(\forall n \in \mathbb{N})\ x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$.
The following properties are satisfied:

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

Fixed point algorithms: α-averaged operator

Let $T\colon \mathcal{H} \to \mathcal{H}$ be an α-averaged operator with $\alpha \in]0,1[$ such that $\operatorname{Fix} T \neq \varnothing$.
Let $(\lambda_n)_{n\in\mathbb{N}}$ be 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 $(\forall n \in \mathbb{N})\ x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$.
The following properties are satisfied:

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

Remark: since $\alpha < 1$, one can choose $(\forall n \in \mathbb{N})\ \lambda_n = 1$, that is

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

Fixed point algorithms: α-averaged operator

<u>Proof</u> :
Since T is α-averaged, there exists a non expansive operator R such that $T = (1-\alpha)\mathrm{Id} + \alpha R$.
Let $(\forall n \in \mathbb{N})\ \mu_n = \alpha\lambda_n \in [0,1]$.

$$\sum_{n\in\mathbb{N}} \lambda_n(1-\alpha\lambda_n) = +\infty \quad \Leftrightarrow \quad \sum_{n\in\mathbb{N}} \mu_n(1-\mu_n) = +\infty.$$

The iterations can be written as

$$(\forall n \in \mathbb{N}) \qquad \begin{aligned} x_{n+1} &= x_n + \lambda_n(Tx_n - x_n) \\ &= x_n + \mu_n(Rx_n - x_n). \end{aligned}$$

Moreover, for every $x \in \mathcal{H}$,

$$x \in \mathrm{Fix}\,T \quad \Leftrightarrow \quad x = (1-\alpha)x + \alpha Rx \quad \Leftrightarrow \quad x \in \mathrm{Fix}R,$$

that is $\mathrm{Fix}R = \mathrm{Fix}\,T$.
+ Krasnosel'skii-Mann algorithm.

Nonexpansive operators

Nonexpansive operators → What is their use ?

Descent lemma

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

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

Descent lemma

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

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

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

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

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

In addition, according to the Cauchy-Schwarz inequality,

$$\begin{aligned}&\langle y-x \mid \nabla f(x+t(y-x)) - \nabla f(x)\rangle \\ \leq\; &\|y-x\|\|\nabla f(x+t(y-x)) - \nabla f(x)\| \leq t\nu\|y-x\|^2.\end{aligned}$$

This leads to

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

Baillon-Haddad theorem

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

Baillon-Haddad theorem

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

Proof : Assume that ∇f is ν-Lipschitzian. According to Fenchel-Young inequality, for every $(x, y, z) \in \mathcal{H}^3$,

$$f^*(\nabla f(y)) \geq \langle z \mid \nabla f(y) \rangle - f(z).$$

From the descent lemma,

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

Moreover, using again the Fenchel-Young result,

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

Thus,

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

Baillon-Haddad theorem

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

Proof : Thus,

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

Taking the supremum with respect to z yields

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

Consequently,

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

Baillon-Haddad theorem

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

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

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

and symmetrically

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

By summing, we finally obtain

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

which shows that ∇f is $1/\nu$-cocoercive.

Nonexpansive operator: example

Baillon-Haddad theorem

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

α-averaged operator: example

Let $f \in \Gamma_0(\mathcal{H})$.
prox_f is a firmly nonexpansive (i.e., $1/2$-averaged operator).

α-averaged operator: example

Proof :
Let $u_1 \in \partial f(x_1)$ and $u_2 \in \partial f(x_2)$. By monotonicity of ∂f,

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

If we consider $u_1' \in (\mathrm{Id} + \partial f)x_1$ et $u_2' \in (\mathrm{Id} + \partial f)x_2$, it results that

$$\langle x_1 - x_2 \mid u_1' - u_2' \rangle \geq \|x_1 - x_2\|^2.$$

Then, from the definition of the proximity operator,

$$\langle \mathrm{prox}_f u_1' - \mathrm{prox}_f u_2' \mid u_1' - u_2' \rangle \geq \|\mathrm{prox}_f u_1' - \mathrm{prox}_f u_2'\|^2.$$


Is there a fixed point?