Proximal Splitting Methods in Signal Recovery

P. L. Combettes

Laboratoire Jacques-Louis Lions
Faculté de Mathématiques
Université Pierre et Marie Curie – Paris 6
75005 Paris, France

Mathematics and Image Analysis, Paris, December 15, 2009

Convex projection methods in signal recovery

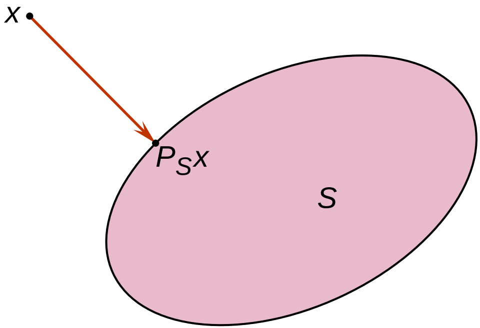
- Mathematical setting: real Hilbert (e.g., Euclidean) space $(\mathcal{H}, \langle \cdot \mid \cdot \rangle)$

Convex projection methods in signal recovery

- Mathematical setting: real Hilbert (e.g., Euclidean) space $(\mathcal{H}, \langle \cdot \mid \cdot \rangle)$
- Signal recovery: restoration, denoising, reconstruction

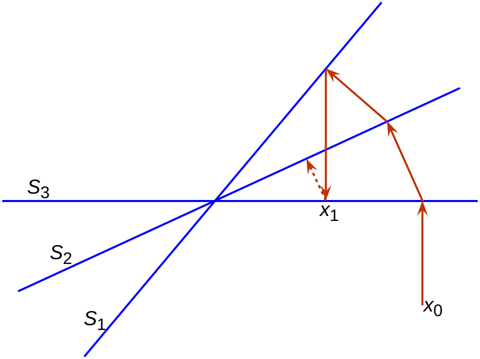
Convex projection methods in signal recovery

- Mathematical setting: real Hilbert (e.g., Euclidean) space $(\mathcal{H}, \langle \cdot \mid \cdot \rangle)$
- Signal recovery: restoration, denoising, reconstruction
- Projection onto a closed convex subset S of $\mathcal{H}$:



Example 1: Algebraic reconstruction techniques (ART) in computer-aided tomography (Herman et al, 1970)

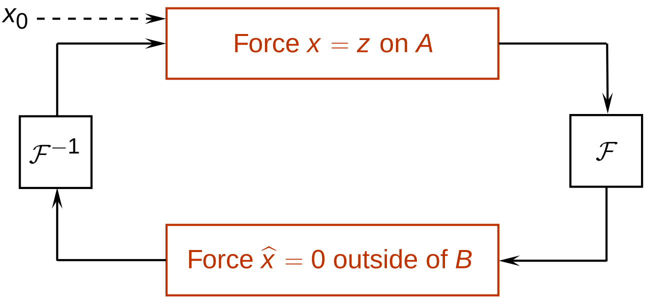
Goal: Reconstruct an image $\overline{x}$ from m scalar measurements $\eta_i = \langle \overline{x} \mid u_i \rangle$. Define $S_i = \big\{ x \in \mathcal{H} \mid \langle x \mid u_i \rangle = \eta_i \big\}$.



This algorithm goes back to Kaczmarz (1937).

Example 2: Band-limited extrapolation (1974-1975)

- The original signal $\overline{x}$ is band-limited (its Fourier transform has compact support B around 0) and it is observed over some region A.
- Gerchberg-Papoulis algorithm:



Example 2: Band-limited extrapolation (1974-1975)

- The set of signals with Fourier support B is the closed vector subspace

$$S_1 = \left\{x \in L^2 \mid \widehat{x}|_{\complement B} = 0\right\}$$

- Projecting x onto S_1 amounts to forcing $\widehat{x}$ to 0 outside of B: $\widehat{P_1 x} = \widehat{x} 1_B$.
- The set of signals which coincide with z on A is the closed affine subspace

$$S_2 = \left\{x \in L^2 \mid x|_A = z\right\}.$$

- Projecting x onto S_2 amounts to forcing $x = z$ on A : $P_2 x = z 1_A + x 1_{\complement A}$.
- Gerchberg-Papoulis is an alternating projection algorithm: $x_{n+1} = P_1 P_2 x_n$ (Youla,1978).

Projection methods in affine feasibility problems

- Given affine subspaces $(S_i)_{1\leq i\leq m}$ of $\mathcal{H}$,

$$\text{Find} \quad x \in S = \bigcap_{i=1}^{m} S_i.$$

Theorem (von Neumann (1933, $m = 2$) - Halperin (1962))

Suppose that $S \neq \varnothing$ and let $x_0 \in \mathcal{H}$. Then

$$x_n = (P_1 \cdots P_m)^n x_0 \longrightarrow P_S x_0.$$

Projection methods in convex feasibility problems

- Given closed convex subsets $(S_i)_{1 \leq i \leq m}$ of $\mathcal{H}$,

$$\text{Find} \quad x \in S = \bigcap_{i=1}^{m} S_i.$$

- Several hundred papers on applications of this convex set theoretic framework in inverse problems.[1]

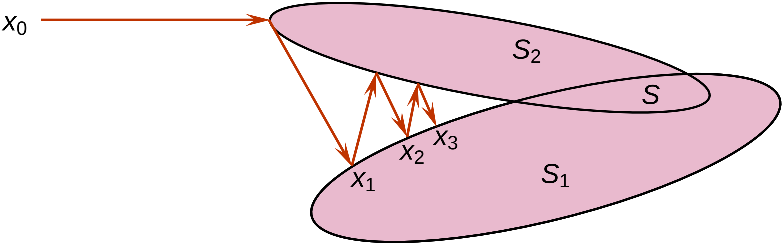
Theorem (Bregman, 1965)

Suppose that $S \neq \varnothing$ and let $x_0 \in \mathcal{H}$. Then (POCS algorithm)

$$x_n = (P_1 \cdots P_m)^n x_0 \rightharpoonup x \in S.$$

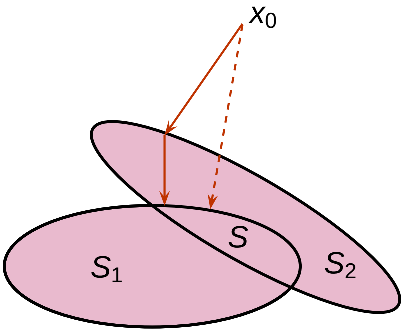
[1] P. L. Combettes, The foundations of set theoretic estimation, *Proc. IEEE* **8**, 182–208 (1993).

Projection methods in convex feasibility problems



- Various variants have been proposed in the form block-iterative parallel algorithms.
- In these convex projection methods the limit is an undetermined feasible point.

Projection methods in convex feasibility problems



The alternating projection algorithm fails to provide the closest point to x_0 in $S = S_1 \cap S_2$.

Projection methods in convex best feasible approximation problems

- Problem: compute $P_S x_0$, i.e.,

$$\min_{x \in S = \bigcap_{i=1}^{m} S_i} \|x - x_0\|$$

- Examples:
  - Minimum energy feasible solution ($x_0 = 0$)
  - Least feasible deviation from a nominal function x_0
  - Constrained signal/image denoising: $x_0 = \overline{x} + w$
- Projection algorithms:
  - Anchor point method
  - Haugazeau's method
  - Boyle-Dykstra's method

Projection methods in convex best feasible approximation problems

Theorem (Boyle-Dykstra, 1986)

Suppose that $S \neq \varnothing$ and let $x_0 \in \mathcal{H}$. Algorithm:

$$
\begin{array}{l}
x_0^m = x_0 \\
\text{for } i = 1, \ldots, m \\
\quad \lfloor \; b_0^i = 0 \\
\text{for } n = 1, 2, \ldots \\
\quad \left\lfloor \begin{array}{l}
x_n^0 = x_{n-1}^m \\
\text{for } i = 1, \ldots, m \\
\quad \left\lfloor \begin{array}{l}
y_n^i = x_n^{i-1} + b_{n-1}^i \\
x_n^i = P_i y_n^i \\
b_n^i = y_n^i - P_i y_n^i
\end{array}\right. \\
x_n = x_n^m
\end{array}\right.
\end{array}
$$

Then $x_n \to P_S x_0$. (Corollary: von-Neumann-Halperin.)

Convex variational formulations in signal recovery

- $\Gamma_0(\mathcal{H})$: lower semicontinuous convex functions $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ such that $\operatorname{dom} f = \{x \in \mathcal{H} \mid f(x) < +\infty\} \neq /emp$.
- General problem:

$$\boxed{\min_{x\in\mathcal{H}} \sum_{i=1}^{m} f_i(x)}, \qquad (1)$$

i.e., we select a point in the feasibility set $S = \bigcap_{i=1}^{m} \operatorname{dom} f_i$.

- Example: $f_1\colon x \mapsto \|x - x_0\|$, $f_i = \iota_{S_i}$ $(2 \leq i \leq m)$ where

$$\iota_{S_i}(x) = 0 \text{ if } x \in S_i;\ \iota_{S_i}(x) = +\infty \text{ if } x \notin S_i.$$

We recover the best feasible approximation problem

$$\min_{x\in\bigcap_{i=2}^{m} S_i} \|x - x_0\|. \qquad (2)$$

- Projections are suitable to solve (2), but not to solve (1).
- Linear/affine projections $\longrightarrow$ convex projections $\longrightarrow$ **???**

J.-J. Moreau's proximity operator (1962)

- Let S be a nonempty closed convex subset of $\mathcal{H}$. Then $\iota_S \in \Gamma_0(\mathcal{H})$ and the projector is defined by

$$P_S \colon x \mapsto \underset{y \in S}{\operatorname{argmin}} \frac{1}{2}\|x - y\|^2 = \underset{y \in \mathcal{H}}{\operatorname{argmin}} \iota_S(y) + \frac{1}{2}\|x - y\|^2.$$

J.-J. Moreau's proximity operator (1962)

- Let S be a nonempty closed convex subset of $\mathcal{H}$. Then $\iota_S \in \Gamma_0(\mathcal{H})$ and the projector is defined by

$$P_S : x \mapsto \underset{y \in S}{\operatorname{argmin}} \frac{1}{2}\|x - y\|^2 = \underset{y \in \mathcal{H}}{\operatorname{argmin}} \, \iota_S(y) + \frac{1}{2}\|x - y\|^2.$$

- More generally, for any function $f \in \Gamma_0(\mathcal{H})$, the proximity operator of f is defined by

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

J.-J. Moreau's proximity operator (1962)

- Let S be a nonempty closed convex subset of $\mathcal{H}$. Then $\iota_S \in \Gamma_0(\mathcal{H})$ and the projector is defined by

$$P_S \colon x \mapsto \underset{y \in S}{\operatorname{argmin}} \frac{1}{2}\|x-y\|^2 = \underset{y \in \mathcal{H}}{\operatorname{argmin}}\, \iota_S(y) + \frac{1}{2}\|x-y\|^2.$$

- More generally, for any function $f \in \Gamma_0(\mathcal{H})$, the proximity operator of f is defined by

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

- Basic properties:
  - $\operatorname{Fix}\, \operatorname{prox}_f = \operatorname{Argmin} f$.
  - $\|\operatorname{prox}_f x - \operatorname{prox}_f y\| \leq \|x-y\|$

J.-J. Moreau's proximity operator (1962)

- Let S be a nonempty closed convex subset of $\mathcal{H}$. Then $\iota_S \in \Gamma_0(\mathcal{H})$ and the projector is defined by

$$P_S : x \mapsto \underset{y \in S}{\operatorname{argmin}} \frac{1}{2}\|x - y\|^2 = \underset{y \in \mathcal{H}}{\operatorname{argmin}} \iota_S(y) + \frac{1}{2}\|x - y\|^2.$$

- More generally, for any function $f \in \Gamma_0(\mathcal{H})$, the proximity operator of f is defined by

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

- Basic properties:
  - $\operatorname{Fix} \operatorname{prox}_f = \operatorname{Argmin} f$.
  - $\|\operatorname{prox}_f x - \operatorname{prox}_f y\|^2 \leq \|x - y\|^2 - \|(\operatorname{Id} - \operatorname{prox}_f)x - (\operatorname{Id} - \operatorname{prox}_f)y\|^2$.

More properties of proximity operators...

property	$\psi(x)$	$\operatorname{prox}_{\psi} x$
shift	$\varphi(x - z),\ z \in \mathcal{H}$	$z + \operatorname{prox}_{\varphi}(x - z)$
scaling	$\varphi(x/\rho),\ \rho \in \mathbb{R} \smallsetminus \{0\}$	$\rho \operatorname{prox}_{\varphi/\rho^2}(x/\rho)$
reflection	$\varphi(-x)$	$-\operatorname{prox}_{\varphi}(-x)$
quadratic perturbation	$\varphi(x) + \alpha\\|x\\|^2/2 + \beta\langle x \mid u\rangle + \gamma$ $u \in \mathcal{H},\ \alpha > 0,\ (\beta,\gamma) \in \mathbb{R}^2$	$\operatorname{prox}_{\varphi/(\alpha+1)}\big((x - \beta u)/(\alpha + 1)\big)$
conjugation	$\varphi^*(x)$	$x - \operatorname{prox}_{\varphi} x$
squared distance	$d_C^2(x)/2$	$(x + P_C x)/2$
Moreau envelope	$\widetilde{\varphi}(x) = \inf_{y\in\mathcal{H}} \varphi(y) + \\|x - y\\|^2/2$	$\big(x + \operatorname{prox}_{2\varphi} x\big)/2$
decomposition in an orthonormal basis	$\sum_{k=1}^{N} \phi_k(\langle b_k \mid x\rangle)$ $\phi_k \in \Gamma_0(\mathbb{R}),\ (b_k)_{1\leq k\leq N}$ orthonormal basis of $\mathcal{H}$	$\sum_{k=1}^{N} \operatorname{prox}_{\phi_k}(\langle b_k \mid x\rangle) b_k$
semi-orthogonal linear transform	$\varphi(Lx)$ $L \in \mathbb{R}^{M\times N},\ LL^* = \nu I,\ \nu > 0$	$x + \nu^{-1}L^*\big(\operatorname{prox}_{\nu\varphi}(Lx) - Lx\big)$
quadratic function	$\gamma\\|Lx - z\\|^2/2$ $L \in \mathbb{R}^{M\times N},\ \gamma > 0,\ z \in \mathcal{G}$	$(I + \gamma L^*L)^{-1}(x + \gamma L^* z)$
indicator function	$\iota_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$	$P_C x$
distance function	$\gamma d_C(x),\ \gamma > 0$	$\begin{cases} x + \dfrac{\gamma}{d_C(x)}(P_C x - x) & \text{if } d_C(x) > \gamma \\ P_C x & \text{otherwise} \end{cases}$
function of distance	$\phi(d_C(x))$ $\phi \in \Gamma_0(\mathbb{R})$ even, differentiable at 0 with $\phi'(0) = 0$	$\begin{cases} x + \left(1 - \dfrac{\operatorname{prox}_{\phi} d_C(x)}{d_C(x)}\right)(P_C x - x) & \text{if } x \notin C \\ x & \text{otherwise} \end{cases}$
support function	$\sigma_C(x)$	$x - P_C x$
extended thresholding	$\sigma_C(x) + \phi(\\|x\\|)$ $\phi \in \Gamma_0(\mathbb{R})$ even and not constant	$\begin{cases} \dfrac{\operatorname{prox}_{\phi} d_C(x)}{d_C(x)}(x - P_C x) & \text{if } d_C(x) > \max \operatorname{Argmin} \phi \\ x - P_C x & \text{otherwise} \end{cases}$

Examples of proximity operators

$\phi(\xi)$	$\mathrm{prox}_{\phi}\,\xi$
$\sigma_{[\underline{\omega},\overline{\omega}]}(\xi)=\begin{cases}\underline{\omega}\xi & \text{if } \xi<0\\ 0 & \text{if } \xi=0\\ \overline{\omega}\xi & \text{otherwise}\end{cases}$	$\mathrm{soft}_{[\underline{\omega},\overline{\omega}]}(\xi)=\begin{cases}\xi-\underline{\omega} & \text{if } \xi<\underline{\omega}\\ 0 & \text{if } \xi\in[\underline{\omega},\overline{\omega}]\\ \xi-\overline{\omega} & \text{if } \xi>\overline{\omega}\end{cases}$
$\psi(\xi)+\sigma_{[\underline{\omega},\overline{\omega}]}(\xi)$ $\psi\in\Gamma_0(\mathbb{R})$ differentiable at 0 with $\psi'(0)=0$	$\mathrm{prox}_{\psi}\left(\mathrm{soft}_{[\underline{\omega},\overline{\omega}]}(\xi)\right)$
$\tau\lvert\xi\rvert^2$	$\xi/(2\tau+1)$
$\kappa\lvert\xi\rvert^p$	$\mathrm{sign}(\xi)\varrho$, where $\varrho\geq 0$ and $\varrho+p\kappa\varrho^{p-1}=\lvert\xi\rvert$
$\begin{cases}\tau\xi^2 & \text{if } \lvert\xi\rvert\leq\omega/\sqrt{2\tau}\\ \omega\sqrt{2\tau}\lvert\xi\rvert-\omega^2/2 & \text{otherwise}\end{cases}$	$\begin{cases}\xi/(2\tau+1) & \text{if } \lvert\xi\rvert\leq\omega(2\tau+1)/\sqrt{2\tau}\\ \xi-\omega\sqrt{2\tau}\,\mathrm{sign}(\xi) & \text{otherwise}\end{cases}$
$\omega\lvert\xi\rvert+\tau\lvert\xi\rvert^2+\kappa\lvert\xi\rvert^p$	$\mathrm{sign}(\xi)\,\mathrm{prox}_{\kappa\lvert\cdot\rvert^p/(2\tau+1)}\left(\max\{\lvert\xi\rvert-\omega,0\}/(2\tau+1)\right)$
$\omega\lvert\xi\rvert-\ln(1+\omega\lvert\xi\rvert)$, $\omega>0$	$(2\omega)^{-1}\mathrm{sign}(\xi)\left(\omega\lvert\xi\rvert-\omega^2-1+\sqrt{\left\lvert\omega\lvert\xi\rvert-\omega^2-1\right\rvert^2+4\omega\lvert\xi\rvert}\right)$
$\begin{cases}\omega\xi & \text{if } \xi\geq 0\\ +\infty & \text{otherwise}\end{cases}$	$\begin{cases}\xi-\omega & \text{if } \xi\geq\omega\\ 0 & \text{otherwise}\end{cases}$
$\begin{cases}-\kappa\ln(\xi)+\omega\xi & \text{if } \xi>0\\ +\infty & \text{otherwise}\end{cases}$	$\left(\xi-\omega+\sqrt{\lvert\xi-\omega\rvert^2+4\kappa}\right)/2$
$\begin{cases}-\kappa\ln(\xi)+\xi^2/2 & \text{if } \xi>0\\ +\infty, & \text{otherwise}\end{cases}$	$\left(\xi+\sqrt{\xi^2+8\kappa}\right)/4$
$\iota_{[\underline{\omega},\overline{\omega}]}(\xi)$	$P_{[\underline{\omega},\overline{\omega}]}\,\xi$
$\begin{cases}-\ln(\xi-\underline{\omega})+\ln(-\underline{\omega}) & \text{if } \xi\in\,]\underline{\omega},0]\\ -\ln(\overline{\omega}-\xi)+\ln(\overline{\omega}) & \text{if } \xi\in\,]0,\overline{\omega}[\\ +\infty & \text{otherwise}\end{cases}$ $\underline{\omega}<0<\overline{\omega}$	$\begin{cases}\left(\xi+\underline{\omega}+\sqrt{\lvert\xi-\underline{\omega}\rvert^2+4}\right)/2 & \text{if } \xi<1/\underline{\omega}\\ \left(\xi+\overline{\omega}-\sqrt{\lvert\xi-\overline{\omega}\rvert^2+4}\right)/2 & \text{if } \xi>1/\overline{\omega}\\ 0 & \text{otherwise}\end{cases}$
$\begin{cases}-\kappa\ln(\xi)+\omega\xi^p & \text{if } \xi>0\\ +\infty & \text{otherwise}\end{cases}$	$\pi>0$ such that $p\omega\pi^p+\pi^2-\xi\pi=\kappa$
$\begin{cases}-\kappa\ln(\xi)+\omega\xi+\rho/\xi & \text{if } \xi>0\\ +\infty, & \text{otherwise}\end{cases}$	$\pi>0$ such that $\pi^3+(\omega-\xi)\pi^2-\kappa\pi=\rho$
$\begin{cases}-\underline{\kappa}\ln(\xi-\underline{\omega})-\overline{\kappa}\ln(\overline{\omega}-\xi) & \text{if } \xi\in\,]\underline{\omega},\overline{\omega}[\\ +\infty & \text{otherwise}\end{cases}$	$\pi\in\,]\underline{\omega},\overline{\omega}[$ such that $\pi^3-(\underline{\omega}+\overline{\omega}+\xi)\pi^2+\left(\underline{\omega}\overline{\omega}-\underline{\kappa}-\overline{\kappa}+(\underline{\omega}+\overline{\omega})\xi\right)\pi=\underline{\omega}\overline{\omega}\xi-\underline{\omega}\overline{\kappa}-\overline{\omega}\underline{\kappa}$

Formal problem statement for $m = 2$

- $\Gamma_0(\mathcal{H})$: proper lower semicontinuous convex functions from $\mathcal{H}$ to $]-\infty, +\infty]$.
- f_1, f_2 in $\Gamma_0(\mathcal{H})$ such that

$$0 \in \operatorname{sri}(\operatorname{dom} f_1 - \operatorname{dom} f_2).$$

- Problem:

$$\boxed{\underset{x \in \mathcal{H}}{\text{minimize}} \;\; f_1(x) + f_2(x)}$$

Forward-backward splitting ($m = 2$)

- f_2 finite, differentiable, ∇f_2 $1/\beta$-Lipschitz-continuous.
- Examples:
  - Noisy linear observations: $z_i = T_i\overline{x} + w_i$, $1 \leq i \leq p$.
  - Closed convex *a priori* constraint sets: $(S_j)_{1\leq j\leq q}$.
  - Functional: $f_2 \colon x \mapsto \sum_{i=1}^{p} \mu_i \|T_i x - z_i\|^2 + \sum_{j=1}^{q} \rho_j d_{S_j}^2(x)$.
- Characterization of solutions: x minimizes $f_1 + f_2 \Leftrightarrow$ $x = \operatorname{prox}_{\gamma f_1}(x - \gamma\nabla f_2(x))$, $\gamma > 0$.
- Algorithm:

$$x_{n+1} = \operatorname{prox}_{\gamma\ f_1}\big(x_n - \gamma\ (\nabla f_2(x_n)\qquad)\big)\qquad ,$$

where

Forward-backward splitting ($m = 2$)

- f_2 finite, differentiable, ∇f_2 $1/\beta$-Lipschitz-continuous.
- Examples:
  - Noisy linear observations: $z_i = T_i\overline{x} + w_i$, $1 \leq i \leq p$.
  - Closed convex *a priori* constraint sets: $(S_j)_{1\leq j\leq q}$.
  - Functional: $f_2 : x \mapsto \sum_{i=1}^{p} \mu_i \|T_i x - z_i\|^2 + \sum_{j=1}^{q} \rho_j d_{S_j}^2(x)$.
- Characterization of solutions: x minimizes $f_1 + f_2 \Leftrightarrow$ $x = \text{prox}_{\gamma f_1}(x - \gamma\nabla f_2(x))$, $\gamma > 0$.
- Algorithm:

$$x_{n+1} = \text{prox}_{\gamma_n f_1}\big(x_n - \gamma_n(\nabla f_2(x_n))\big),$$

where
  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leq \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$.

Forward-backward splitting ($m = 2$)

- f_2 finite, differentiable, ∇f_2 $1/\beta$-Lipschitz-continuous.
- Examples:
  - Noisy linear observations: $z_i = T_i\overline{x} + w_i$, $1 \leq i \leq p$.
  - Closed convex *a priori* constraint sets: $(S_j)_{1\leq j\leq q}$.
  - Functional: $f_2 : x \mapsto \sum_{i=1}^{p} \mu_i \|T_i x - z_i\|^2 + \sum_{j=1}^{q} \rho_j d_{S_j}^2(x)$.
- Characterization of solutions: x minimizes $f_1 + f_2 \Leftrightarrow$ $x = \operatorname{prox}_{\gamma f_1}(x - \gamma\nabla f_2(x))$, $\gamma > 0$.
- Algorithm:

$$x_{n+1} = \operatorname{prox}_{\gamma_n f_1}\big(x_n - \gamma_n(\nabla f_2(x_n) + b_n)\big) + a_n ,$$

where

  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leq \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$.
  - $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$.

Forward-backward splitting ($m = 2$)

- f_2 finite, differentiable, ∇f_2 $1/\beta$-Lipschitz-continuous.
- Examples:
  - Noisy linear observations: $z_i = T_i\overline{x} + w_i$, $1 \leq i \leq p$.
  - Closed convex *a priori* constraint sets: $(S_j)_{1\leq j\leq q}$.
  - Functional: $f_2 : x \mapsto \sum_{i=1}^{p} \mu_i \|T_i x - z_i\|^2 + \sum_{j=1}^{q} \rho_j d_{S_j}^2(x)$.
- Characterization of solutions: x minimizes $f_1 + f_2 \Leftrightarrow$ $x = \operatorname{prox}_{\gamma f_1}(x - \gamma\nabla f_2(x))$, $\gamma > 0$.
- Algorithm:

$$x_{n+1} = x_n + \lambda_n\Big(\operatorname{prox}_{\gamma_n f_1}\big(x_n - \gamma_n(\nabla f_2(x_n) + b_n)\big) + a_n - x_n\Big),$$

where

  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leq \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$.
  - $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$.
  - $(\lambda_n)_{n\in\mathbb{N}}$ in $]0,1]$, $\inf_{n\in\mathbb{N}} \lambda_n > 0$.

Forward-backward splitting ($m = 2$)

Theorem

Suppose that $\operatorname{Argmin} f_1 + f_2 \neq \varnothing$. *Then any sequence* $(x_n)_{n\in\mathbb{N}}$ *generated by the forward-backward algorithm converges weakly to a point in* $\operatorname{Argmin} f_1 + f_2$.

This result covers and extends:

- Alternating projection method, parallel projection method;
- Parallel projection methods for hard constrained inconsistent feasibility problems;
- Projected Landweber method, split feasibility methods;
- Iterative soft-thresholding method;
- Variational geometry/texture decomposition methods; etc.

Douglas-Rachford splitting ($m = 2$)

- f_2 is no longer assumed to be smooth.
- For instance, f_1 and f_2 are any of the following:
  - ι_C, $C \subset \mathcal{H}$ closed and convex.
  - d_C, $C \subset \mathcal{H}$ closed and convex.
  - $\|\cdot\|_1$ in $\mathcal{H} = \mathbb{R}^N$.
  - $x \mapsto \sum_{i=1}^m \mu_i \|T_i x - z_i\|_1$ in $\mathcal{H} = \mathbb{R}^N$.
  - $x \mapsto \sum_{i=1}^m \mu_i \|T_i x - z_i\|_{L^1}$ in $\mathcal{H} = L^2(\Omega)$, $\Omega \subset \mathbb{R}^N$ open, bounded.
  - Total variation.
  - $x \mapsto \int_\Omega \phi(x(t), \nabla x(t))dt$ in $\mathcal{H} = H^1(\Omega)$, $\Omega \subset \mathbb{R}^N$ open, bounded, and $\phi \in \Gamma_0(\mathbb{R}^{m+1})$ nonsmooth.
  - $\max_{1 \leq i \leq m} \varphi_i$, $\varphi_i \in \Gamma_0(\mathcal{H})$.
  - etc...

Douglas-Rachford splitting ($m = 2$)

Characterization of solutions: Let $x \in \mathcal{H}$ and $\gamma \in \,]0, +\infty[$. Then the following are equivalent.

- $x \in \operatorname{Argmin} f_1 + f_2$.
- $x = \operatorname{prox}_{\gamma f_2} y$, where y satisfies

$$\operatorname{prox}_{\gamma f_2} y = \operatorname{prox}_{\gamma f_1}\big(2\operatorname{prox}_{\gamma f_2} y - y\big).$$

Douglas-Rachford splitting ($m = 2$)

Algorithm:

- Let $\gamma \in \,]0, +\infty[$, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $]0, 2[$, and let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathcal{H}$.
- $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$.
- $\sum_{n\in\mathbb{N}} \lambda_n(\|a_n\| + \|b_n\|) < +\infty$.
- Iterations: Take $x_0 \in \mathcal{H}$ and set, for every $n \in \mathbb{N}$,

$$\begin{cases} x_{n+\frac{1}{2}} = \operatorname{prox}_{\gamma f_2} x_n + b_n \\ x_{n+1} = x_n + \lambda_n\Big(\operatorname{prox}_{\gamma f_1}\big(2x_{n+\frac{1}{2}} - x_n\big) + a_n - x_{n+\frac{1}{2}}\Big). \end{cases}$$

Douglas-Rachford splitting ($m = 2$)

Theorem

Suppose that $\operatorname{Argmin} f_1 + f_2 \neq \varnothing$ *and let* $(x_n)_{n\in\mathbb{N}}$ *be an arbitrary sequence generated by the Douglas-Rachford algorithm. Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to some point* $y \in \mathcal{H}$ *and* $x = \operatorname{prox}_{\gamma f_2} y \in \operatorname{Argmin} f_1 + f_2$.

Corollary

Suppose that $\operatorname{Argmin} f_1 + f_2 \neq \varnothing$, *that* $\mathcal{H}$ *is finite-dimensional, and that* $b_n \to 0$. *Let* $(x_n)_{n\in\mathbb{N}}$ *be an arbitrary sequence generated by the Douglas-Rachford algorithm. Then* $(x_{n+\frac{1}{2}})_{n\in\mathbb{N}}$ *converges to a point in* $\operatorname{Argmin} f_1 + f_2$.

Proximal splitting in the general case

- The functions $(f_i)_{1\leq i\leq m}$ are in $\Gamma_0(\mathcal{H})$.
- CQ: $0 \in \operatorname{sri}\big\{(x - x_1, \ldots, x - x_m) \mid x \in \mathcal{H}, x_i \in \operatorname{dom} f_i\big\}$.
- Problem:

$$\min_{x\in\mathcal{H}} \sum_{i=1}^{m} f_i(x)$$

Proximal splitting in the general case

- The functions $(f_i)_{1\leq i\leq m}$ are in $\Gamma_0(\mathcal{H})$.
- CQ: $0 \in \operatorname{sri}\big\{(x - x_1, \ldots, x - x_m) \mid x \in \mathcal{H},\, x_i \in \operatorname{dom} f_i\big\}$.
- Problem:

$$\min_{x\in\mathcal{H}} \sum_{i=1}^{m} f_i(x), \quad i.e., \underset{x_1=\cdots=x_m}{\min} \sum_{i=1}^{m} f_i(x_i).$$

Proximal splitting in the general case

- The functions $(f_i)_{1\leq i\leq m}$ are in $\Gamma_0(\mathcal{H})$.
- CQ: $0 \in \operatorname{sri}\big\{(x-x_1,\ldots,x-x_m) \mid x \in \mathcal{H},\, x_i \in \operatorname{dom} f_i\big\}$.
- Problem:

$$\min_{x\in\mathcal{H}} \sum_{i=1}^{m} f_i(x), \quad i.e., \min_{x_1=\cdots=x_m} \sum_{i=1}^{m} f_i(x_i).$$

- Equivalent problem in the product space $\mathcal{H}^m$:

$$\min_{\boldsymbol{x}\in\mathcal{H}^m} \iota_{\boldsymbol{D}}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}), \quad \text{with} \quad \begin{cases} \boldsymbol{D} = \big\{(x,\ldots,x) \mid x \in \mathcal{H}\big\} \\ \boldsymbol{f} : \boldsymbol{x} = (x_i)_{1\leq i\leq m} \mapsto \sum_{i=1}^{m} f_i(x_i). \end{cases}$$

Proximal splitting in the general case

- The functions $(f_i)_{1\leq i\leq m}$ are in $\Gamma_0(\mathcal{H})$.
- CQ: $0 \in \operatorname{sri}\big\{(x-x_1,\ldots,x-x_m) \mid x\in\mathcal{H},\, x_i\in\operatorname{dom} f_i\big\}$.
- Problem:

$$\min_{x\in\mathcal{H}} \sum_{i=1}^{m} f_i(x), \quad i.e., \quad \min_{x_1=\cdots=x_m} \sum_{i=1}^{m} f_i(x_i).$$

- Equivalent problem in the product space $\mathcal{H}^m$:

$$\min_{\boldsymbol{x}\in\mathcal{H}^m} \iota_{\boldsymbol{D}}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}), \quad \text{with} \quad \begin{cases} \boldsymbol{D} = \{(x,\ldots,x) \mid x\in\mathcal{H}\} \\ \boldsymbol{f} : \boldsymbol{x} = (x_i)_{1\leq i\leq m} \mapsto \sum_{i=1}^{m} f_i(x_i). \end{cases}$$

- Apply Douglas-Rachford to $\iota_{\boldsymbol{D}}$ and $\boldsymbol{f}$ in the product space!

Parallel proximal algorithm (PPXA)

Initialization

$$\gamma \in \,]0,+\infty[$$
$$(\omega_i)_{1\le i\le m} \in \,]0,1]^m \text{ satisfy } \sum_{i=1}^m \omega_i = 1$$
$$(y_{i,0})_{1\le i\le m} \in \mathcal{H}^m$$
$$x_0 = \sum_{i=1}^m \omega_i y_{i,0}$$

For $n = 0, 1, \ldots$

For $i = 1, \ldots, m$

$$p_{i,n} = \operatorname{prox}_{\gamma f_i/\omega_i} y_{i,n} + a_{i,n}$$

$$p_n = \sum_{i=1}^m \omega_i p_{i,n}$$
$$\lambda_n \in \,]0,2[$$

For $i = 1, \ldots, m$

$$y_{i,n+1} = y_{i,n} + \lambda_n\big(2p_n - x_n - p_{i,n}\big)$$

$$x_{n+1} = x_n + \lambda_n(p_n - x_n).$$

Parallel proximal algorithm (PPXA)

Theorem

Suppose that $\operatorname{Argmin} f_1 + \cdots + f_m \neq \varnothing$ *and let* $(x_n)_{n\in\mathbb{N}}$ *be an arbitrary sequence generated by the PPXA algorithm. Then* $(x_n)_{n\in\mathbb{N}}$ *converges weakly to some point in* $\operatorname{Argmin} f_1 + \cdots + f_m$.

Dykstra-like parallel proximal algorithm

- $(f_i)_{1 \le i \le m}$ functions in $\Gamma_0(\mathcal{H})$ such that (no CQ)

$$\operatorname{dom} f_1 \cap \cdots \cap \operatorname{dom} f_m \neq \varnothing.$$

- $\omega_i > 0$, $\sum_{i=1}^m \omega_i = 1$.
- Problem (extending the best feasible approximation problem):

$$\min_{x \in \mathcal{H}} \sum_{i=1}^m \omega_i f_i(x) + \frac{1}{2}\|x - x_0\|^2.$$

Dykstra-like parallel proximal algorithm

Initialization

$z_{1,0} = x_0, \ldots, z_{m,0} = x_0$

For $n = 0, 1, \ldots$

For $i = 1, \ldots, m$

$p_{i,n} = \text{prox}_{f_i} z_{i,n}$

$$x_{n+1} = \sum_{i=1}^{m} \omega_i p_{i,n}$$

For $i = 1, \ldots, m$

$z_{i,n+1} = x_{n+1} + z_{i,n} - p_{i,n}.$

Theorem

$x_n \longrightarrow argmin\ \omega_1 f_1 + \cdots + \omega_m f_m + \| \cdot - x_0\|^2/2.$

Alternating-direction method of multipliers (ADMM)

- $f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{G})$, $L\colon \mathcal{H} \to \mathcal{G}$ linear and bounded, $L^* \circ L$ invertible, $0 \in \operatorname{sri}(L(\operatorname{dom} f) - \operatorname{dom} g)$.
- Problem:

$$\underset{x\in\mathcal{H}}{\operatorname{minimize}}\; f(x) + g(Lx), \; i.e., \; \underset{\substack{x\in\mathcal{H},\, y\in\mathcal{G}\\ Lx=y}}{\operatorname{minimize}}\; f(x) + g(y).$$

- Augmented Lagrangian:

$$\mathcal{L}_\gamma\colon \mathcal{H} \times \mathcal{G} \times \mathcal{G} \to \left]-\infty, +\infty\right]$$

$$(x, y, z) \mapsto f(x) + g(y) + \frac{1}{\gamma}\langle (Lx - y) \mid z\rangle + \frac{1}{2\gamma}\|Lx - y\|^2.$$

Alternating-direction method of multipliers (ADMM)

Denote by prox_f^L the operator which maps $y \in \mathcal{G}$ to the unique minimizer of $x \mapsto f(x) + \|Lx - y\|^2/2$.

Initialization

$$\left\lfloor \begin{array}{l} \gamma > 0 \\ y_0 \in \mathcal{G} \\ z_0 \in \mathcal{G} \end{array} \right.$$

For $n = 0, 1, \ldots$

$$\left\lfloor \begin{array}{l} x_n = \mathrm{prox}_{\gamma f}^L (y_n - z_n) \\ s_n = L x_n \\ y_{n+1} = \mathrm{prox}_{\gamma g}(s_n + z_n) \\ z_{n+1} = z_n + s_n - y_{n+1} \end{array} \right.$$

Alternating-direction method of multipliers (ADMM)

Denote by prox_f^L the operator which maps $y \in \mathcal{G}$ to the unique minimizer of $x \mapsto f(x) + \|Lx - y\|^2/2$.

Initialization
$$\left\lfloor \begin{array}{l} \gamma > 0 \\ y_0 \in \mathcal{G} \\ z_0 \in \mathcal{G} \end{array} \right.$$
For $n = 0, 1, \ldots$
$$\left\lfloor \begin{array}{l} x_n = \mathrm{prox}_{\gamma f}^L(y_n - z_n) \\ s_n = Lx_n \\ y_{n+1} = \mathrm{prox}_{\gamma g}(s_n + z_n) \\ z_{n+1} = z_n + s_n - y_{n+1} \end{array} \right.$$

~~Alternating split Bregman algorithm~~

Simultaneous-direction method of multipliers (SDMM)

- $g_i \in \Gamma_0(\mathcal{G}_i)$, $1 \leq i \leq m$
- $L_i \colon \mathcal{H} \to \mathcal{G}_i$ linear and bounded, $Q = \sum_{i=1}^{m} L_i^* \circ L_i$ invertible
- $0 \in \operatorname{sri}\big\{(L_1 x - y_1, \ldots, L_m x - y_m) \mid x \in \mathcal{H}, y_i \in \operatorname{dom} g_i\big\}$.
- Problem:

$$\underset{x \in \mathcal{H}}{\text{minimize}} \;\; g_1(L_1 x) + \cdots + g_m(L_m x).$$

Simultaneous-direction method of multipliers (SDMM)

Initialization

$$\left\lfloor \begin{array}{l} \gamma > 0 \\ y_{1,0} \in \mathcal{G}_1, \ldots, y_{m,0} \in \mathcal{G}_m \\ z_{1,0} \in \mathcal{G}_1, \ldots, z_{m,0} \in \mathcal{G}_m \end{array} \right.$$

For $n = 0, 1, \ldots$

$$\left\lfloor \begin{array}{l} x_n = Q^{-1} \sum_{i=1}^{m} L_i^* (y_{i,n} - z_{i,n}) \\ \text{For } i = 1, \ldots, m \\ \left\lfloor \begin{array}{l} s_{i,n} = L_i x_n \\ y_{i,n+1} = \operatorname{prox}_{\gamma g_i}(s_{i,n} + z_{i,n}) \\ z_{i,n+1} = z_{i,n} + s_{i,n} - y_{i,n+1} \end{array} \right. \end{array} \right.$$

References

- http://www.ann.jussieu.fr/∼plc/publist.html
- arxiv