Méthodes d'Éclatement d'Opérateurs Primales-Duales Monotones au Sens de Bregman

Patrick L. Combettes

Department of Mathematics
North Carolina State University
Raleigh, NC 27695, USA

Journées des GdR MOA et MIA, Bordeaux, 20 octobre 2017

PREMIÈRE PARTIE : ÉCLATEMENT DANS LES ESPACES HILBERTIENS

Monotone operators

- $\mathcal{H}$ a real Hilbert space.
- $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone: for every $(x, x^*) \in \mathcal{H}^2$,

$$(x, x^*) \in \operatorname{gra} A \quad \Leftrightarrow \quad (\forall (y, y^*) \in \operatorname{gra} A) \quad \langle x - y \mid x^* - y^* \rangle \geqslant 0$$

- The resolvent of A, $J_A = (\mathrm{Id} + A)^{-1}\colon \mathcal{H} \to \mathcal{H}$, is firmly nonexpansive and $\operatorname{Fix} J_A = \operatorname{zer} A = \{x \in \mathcal{H} \mid 0 \in Ax\}$
- Minty's parametrization:

$$(\forall x \in \mathcal{H}) \quad (J_A x, x - J_A x) = (J_A x, J_{A^{-1}} x) \in \operatorname{gra} A$$

Solving monotone inclusions

- $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ maximally monotone,
- Problem: solve: $0 \in Ax$
- Conceptual solution methods when A is simple:
  - The proximal point algorithm (implicit):

$$x_{n+1} = (\mathrm{Id} + \gamma_n A)^{-1} x_n = J_{\gamma_n A} x_n, \quad \text{where} \quad \gamma_n > 0.$$

  - If $A\colon \mathcal{H} \to \mathcal{H}$ is β-cocoercive ($A^{-1} - \beta\mathrm{Id}$ is monotone), the explicit iteration

$$x_{n+1} = x_n - \gamma_n A x_n, \quad \text{where} \quad 0 < \gamma_n < 2\beta.$$

- For "real" problems **splitting** is required.

Splitting structured problems: 3 basic methods

$A, B\colon \mathcal{H} \to 2^{\mathcal{H}}$ maximally monotone, solve $0 \in A\overline{x} + B\overline{x}$.

- Douglas-Rachford splitting (1979)

$$\left\lfloor \begin{array}{l} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + z_n - y_n \end{array} \right.$$

- $B\colon \mathcal{H} \to \mathcal{H}$ β-cocoercive: forward-backward splitting (1979+)

$$\left\lfloor \begin{array}{l} 0 < \gamma_n < 2\beta \\ y_n = x_n - \gamma_n B x_n \\ x_{n+1} = J_{\gamma_n A} y_n \end{array} \right.$$

- $B\colon \mathcal{H} \to \mathcal{H}$ μ-Lipschitzian: forward-backward-forward splitting (2000)

$$\left\lfloor \begin{array}{l} 0 < \gamma_n < 1/\mu \\ y_n = x_n - \gamma_n B x_n \\ z_n = J_{\gamma_n A} y_n \\ r_n = z_n - \gamma_n B z_n \\ x_{n+1} = x_n - y_n + r_n \end{array} \right.$$

- Spingarn's method (1983) for $0 \in A_1\overline{x} + \cdots + A_n\overline{x}$.

Forward-backward splitting for minimization

- Solution set: $Z = \text{Argmin}\,(f+g)$, where $f \in \Gamma_0(\mathcal{H})$, $g\colon \mathcal{H} \to \mathbb{R}$ convex, differentiable, ∇g is $1/\beta$-Lipschitz-continuous
- The sequence constructed by the algorithm

$$x_{n+1} = \quad \text{prox}_{\gamma \;\; f}\big(x_n - \gamma \quad (\nabla g(x_n) \quad)\big)$$

  - $0 < \gamma < 2\beta$ (Mercier, 1979)

converges weakly to a point in Z

Forward-backward splitting for minimization

- Solution set: $Z = \text{Argmin}\,(f+g)$, where $f \in \Gamma_0(\mathcal{H})$, $g\colon \mathcal{H} \to \mathbb{R}$ convex, differentiable, ∇g is $1/\beta$-Lipschitz-continuous
- The sequence constructed by the algorithm

$$x_{n+1} = \quad \text{prox}_{\gamma_n f}\big(x_n - \gamma_n \quad (\nabla g(x_n) \quad)\big)$$

  - $0 < \gamma < 2\beta$ (Mercier, 1979)
  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leqslant \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$ (Tseng, 1990)

converges weakly to a point in Z

Forward-backward splitting for minimization

- Solution set: $Z = \text{Argmin}\,(f+g)$, where $f \in \Gamma_0(\mathcal{H})$, $g\colon \mathcal{H} \to \mathbb{R}$ convex, differentiable, ∇g is $1/\beta$-Lipschitz-continuous
- The sequence constructed by the algorithm

$$x_{n+1} = \text{prox}_{\gamma_n f}\big(x_n - \gamma_n (\nabla g(x_n) + b_n)\big) + a_n$$

  - $0 < \gamma < 2\beta$ (Mercier, 1979)
  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leqslant \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$ (Tseng, 1990)
  - $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$ (PLC, 2004)

converges weakly to a point in Z

Forward-backward splitting for minimization

- Solution set: $Z = \text{Argmin}\,(f+g)$, where $f \in \Gamma_0(\mathcal{H})$, $g\colon \mathcal{H} \to \mathbb{R}$ convex, differentiable, ∇g is $1/\beta$-Lipschitz-continuous
- The sequence constructed by the algorithm

$$x_{n+1} = x_n + \lambda_n\Big(\text{prox}_{\gamma_n f}\big(x_n - \gamma_n (\nabla g(x_n) + b_n)\big) + a_n - x_n\Big)$$

  - $0 < \gamma < 2\beta$ (Mercier, 1979)
  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leqslant \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$ (Tseng, 1990)
  - $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$ (PLC, 2004)
  - $(\lambda_n)_{n\in\mathbb{N}}$ in $]0,1]$, $\inf_{n\in\mathbb{N}} \lambda_n > 0$ (PLC, 2004)

converges weakly to a point in Z

Forward-backward splitting for minimization

- Solution set: $Z = \text{Argmin}\,(f+g)$, where $f \in \Gamma_0(\mathcal{H})$, $g\colon \mathcal{H} \to \mathbb{R}$ convex, differentiable, ∇g is $1/\beta$-Lipschitz-continuous
- The sequence constructed by the algorithm

$$x_{n+1} = x_n + \lambda_n\Big(\text{prox}^{U_n}_{\gamma_n f}\big(x_n - \gamma_n U_n^{-1}(\nabla g(x_n) + b_n)\big) + a_n - x_n\Big)$$

  - $0 < \gamma < 2\beta$ (Mercier, 1979)
  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leqslant \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$ (Tseng, 1990)
  - $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$ (PLC, 2004)
  - $(\lambda_n)_{n\in\mathbb{N}}$ in $]0,1]$, $\inf_{n\in\mathbb{N}} \lambda_n > 0$ (PLC, 2004)
  - $(1+\eta_n)U_{n+1} \succcurlyeq U_n = U_n^* \succcurlyeq \alpha\text{Id}$, $\alpha > 0$, $\eta_n \geqslant 0$, $\sum_{n\in\mathbb{N}} \eta_n < +\infty$ (PLC&Vũ, 2012)

converges weakly to a point in Z

Forward-backward splitting for minimization

- Solution set: $Z = \operatorname{Argmin}(f+g)$, where $f \in \Gamma_0(\mathcal{H})$, $g\colon \mathcal{H} \to \mathbb{R}$ convex, differentiable, ∇g is $1/\beta$-Lipschitz-continuous
- The sequence constructed by the algorithm

$$x_{n+1} = x_n + \lambda_n\Big(\operatorname{prox}^{U_n}_{\gamma_n f}\big(x_n - \gamma_n U_n^{-1}(\nabla g(x_n) + b_n)\big) + a_n - x_n\Big)$$

  - $0 < \gamma < 2\beta$ (Mercier, 1979)
  - $0 < \inf_{n\in\mathbb{N}} \gamma_n \leqslant \sup_{n\in\mathbb{N}} \gamma_n < 2\beta$ (Tseng, 1990)
  - $\sum_{n\in\mathbb{N}} \|a_n\| < +\infty$, $\sum_{n\in\mathbb{N}} \|b_n\| < +\infty$ (PLC, 2004)
  - $(\lambda_n)_{n\in\mathbb{N}}$ in $]0,1]$, $\inf_{n\in\mathbb{N}} \lambda_n > 0$ (PLC, 2004)
  - $(1+\eta_n)U_{n+1} \succcurlyeq U_n = U_n^* \succcurlyeq \alpha\mathrm{Id}$, $\alpha > 0$, $\eta_n \geqslant 0$, $\sum_{n\in\mathbb{N}} \eta_n < +\infty$ (PLC&Vũ, 2012)

  converges weakly to a point in Z
- Also: almost surely weakly convergent versions with random block-coordinate sweeping (PLC&Pesquet, SIOPT, July 2015) and/or stochastic approximations (PLC&Pesquet, Pure Appl. Funct. Anal., Jan. 2016)

Further properties of forward-backward splitting

- Solution set: $Z = $Argmin $f + g$
- $x_{n+1} = x_n + \lambda_n\Big(\text{prox}_{\gamma_n f}(x_n - \gamma_n \nabla g(x_n)) - x_n\Big)$, $\varepsilon \leqslant \gamma_n \leqslant (2-\varepsilon)\beta$
- $(\forall n \in \mathbb{N})(\forall z \in Z) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|$: Fejér monotonicity
- $(\forall n \in \mathbb{N}) \quad (f+g)(x_{n+1}) \leqslant (f+g)(x_n)$
- Convergence is only weak
- Even in the finite dimensional or the linear case, no (upper bound on the worst) rate of convergence of $\|x_n - x_\infty\|$ exists
- $\sum_{n\in\mathbb{N}} \big|(f+g)(x_n) - \inf(f+g)(\mathcal{H})\big|^2 < +\infty$
- If $\sum_{n\in\mathbb{N}}(1-\lambda_n) < +\infty$, $(f+g)(x_n) - \inf(f+g)(\mathcal{H}) = o(1/n)$ (PLC, Salzo, Villa, 2017)
- In the case of the projected gradient method, some form of the above results already in:
  - E. S. Levitin and B. T. Polyak, Constrained minimization methods, *Comput. Math. Math. Phys.*, vol. 6, pp. 1–50, 1966

On minimizing sequences

- Let $\Phi \in \Gamma_0(\mathcal{H})$, $Z = \operatorname{Argmin} \Phi \neq \varnothing$ the solution set
- Minimizing sequences may have little to do with actually approaching a point in Z as we can have (even in $\mathbb{R}^2$):
  - $\Phi(x_n) \to \inf \Phi(\mathcal{H})$ and $(\forall z \in Z)\ \|x_n - z\| \geqslant 1$
  - $\Phi(x_n) \to \inf \Phi(\mathcal{H})$ and $(\forall z \in Z)\ \|x_n - z\| \to +\infty$
  - ... and vice versa $\Phi(x_n) \equiv +\infty$ and $x_n \to z \in Z$
- The whole area of metric regularity addresses such issues

Splitting algorithms (1979-2000)

find $\overline{x} \in \mathcal{H}$ such that

$$z^* \in A\overline{x} + B\overline{x}$$

where:

- $z^* \in \mathcal{H}$, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone
- $B\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone

Splitting algorithms (Briceño-Arias-PLC, 2011)

find $\overline{x} \in \mathcal{H}$ such that

$$z^* \in A\overline{x} + L^* B(L\overline{x} - r)$$

where:

- $z^* \in \mathcal{H}$, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone
- $B\colon \mathcal{G} \to 2^{\mathcal{G}}$ is maximally monotone, $r \in \mathcal{G}$, $L \in \mathscr{B}(\mathcal{H}, \mathcal{G})$

Splitting algorithms (Briceño-Arias-PLC, 2011)

find $\overline{x} \in \mathcal{H}$ such that

$$z^* \in A\overline{x} + \sum_{k=1}^{K} L_k^* B_k (L_k \overline{x} - r_k)$$

where:

- $z^* \in \mathcal{H}$, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone
- $B_k : \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, $r_k \in \mathcal{G}_k$, $L_k \in \mathscr{B}(\mathcal{H}, \mathcal{G}_k)$

Splitting algorithms (PLC-Pesquet, 2012)

find $\overline{x} \in \mathcal{H}$ such that

$$z^* \in A\overline{x} + \sum_{k=1}^{K} L_k^*(B_k \,\square\, D_k)(L_k\overline{x} - r_k)$$

where:

- $z^* \in \mathcal{H}$, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone
- $B_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, $r_k \in \mathcal{G}_k$, $L_k \in \mathscr{B}(\mathcal{H}, \mathcal{G}_k)$
- $D_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, D_k^{-1} is ν_k-Lipschitzian, $B_k \,\square\, D_k = (B_k^{-1} + D_k^{-1})^{-1}$

Splitting algorithms (PLC-Pesquet, 2012)

find $\overline{x} \in \mathcal{H}$ such that

$$z^* \in A\overline{x} + \sum_{k=1}^{K} L_k^*(B_k \,\square\, D_k)(L_k - r_k\overline{x}) + C\overline{x}$$

where:

- $z^* \in \mathcal{H}$, $A\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone
- $B_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, $r_k \in \mathcal{G}_k$, $L_k \in \mathscr{B}(\mathcal{H}, \mathcal{G}_k)$
- $D_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, D_k^{-1} is ν_k-Lipschitzian, $B_k \,\square\, D_k = (B_k^{-1} + D_k^{-1})^{-1}$
- $C\colon \mathcal{H} \to \mathcal{H}$ is monotone and μ-Lipschtizian

Splitting algorithms (PLC, 2013)

find $\overline{x}_1 \in \mathcal{H}_1, \ldots, \overline{x}_m \in \mathcal{H}_m$ such that

$$\begin{cases} z_1^* \in A_1\overline{x}_1 + \sum_{k=1}^{K} L_{k1}^*\Big((B_k \,\Box\, D_k)\Big(\sum_{i=1}^{m} L_{ki}\overline{x}_i - r_k\Big)\Big) + C_1\overline{x}_1 \\ \quad\vdots \\ z_m^* \in A_m\overline{x}_m + \sum_{k=1}^{K} L_{km}^*\Big((B_k \,\Box\, D_k)\Big(\sum_{i=1}^{m} L_{ki}\overline{x}_i - r_k\Big)\Big) + C_m\overline{x}_m \end{cases}$$

where:

- $z_i^* \in \mathcal{H}_i$, $A_i \colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ is maximally monotone
- $B_k \colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, $r_k \in \mathcal{G}_k$, $L_{ki} \in \mathscr{B}(\mathcal{H}_i, \mathcal{G}_k)$
- $D_k \colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ is maximally monotone, D_k^{-1} is ν_k-Lipschitzian, $B_k \,\Box\, D_k = (B_k^{-1} + D_k^{-1})^{-1}$
- $C_i \colon \mathcal{H}_i \to \mathcal{H}_i$ is monotone and μ_i-Lipschtizian

Splitting algorithms (PLC, 2013)

- $\mathcal{K} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_m \oplus \mathcal{G}_1 \oplus \cdots \oplus \mathcal{G}_p$
- $\boldsymbol{M} \colon \mathcal{K} \to 2^{\mathcal{K}} \colon (x_1, \ldots, x_m, v_1^*, \ldots, v_p^*) \mapsto (-z_1^* + A_1 x_1) \times \cdots \times (-z_m^* + A_m x_m) \times (r_1 + B_1^{-1} v_1^*) \times \cdots \times (r_p + B_p^{-1} v_p^*)$
- $\boldsymbol{Q} \colon \mathcal{K} \to \mathcal{K} \colon (x_1, \ldots, x_m, v_1^*, \ldots, v_p^*) \mapsto \big(C_1 x_1 + \sum_{k=1}^K L_{k1}^* v_k^*, \ldots, C_m x_m + \sum_{k=1}^K L_{km}^* v_k^*, -\sum_{i=1}^m L_{1i} x_i + D_1^{-1} v_1^*, \ldots, \sum_{i=1}^m L_{Ki} x_i + D_K^{-1} v_K^*\big)$
- $\boldsymbol{M}$ and $\boldsymbol{Q}$ are maximally monotone, $\boldsymbol{Q}$ is Lipschitzian, the zeros of $\boldsymbol{M} + \boldsymbol{Q}$ are primal-dual solutions
- Solve $\boldsymbol{0} \in \boldsymbol{M}\boldsymbol{x} + \boldsymbol{Q}\boldsymbol{x}$, where $\boldsymbol{x} = (x_1, \ldots, x_m, v_1^*, \ldots, v_p^*)$ via Tseng's forward-backward-forward splitting algorithm

$$
\left\lfloor
\begin{array}{l}
\boldsymbol{y}_n = \boldsymbol{x}_n - \boldsymbol{Q}\boldsymbol{x}_n \\
\boldsymbol{p}_n = (\mathrm{Id} + \boldsymbol{M})^{-1} \boldsymbol{y}_n \\
\boldsymbol{q}_n = \boldsymbol{p}_n - \boldsymbol{Q}\boldsymbol{p}_n \\
\boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{y}_n + \boldsymbol{q}_n
\end{array}
\right.
$$

in $\mathcal{K}$ to get...

Splitting algorithms (PLC, 2013)

For $n = 0, 1, \ldots$

$$\varepsilon \leqslant \gamma_n \leqslant (1-\varepsilon)/\Big(\max\Big\{\max_{1\leqslant i\leqslant m} \mu_i, \max_{1\leqslant k\leqslant K} \nu_k\Big\} + \sqrt{\textstyle\sum_{k=1}^K \sum_{i=1}^m \|L_{ki}\|^2}\Big)$$

For $i = 1, \ldots, m$

$$s_{1,i,n} = x_{i,n} - \gamma_n\Big(C_i x_{i,n} + \textstyle\sum_{k=1}^K L_{ki}^* v_{k,n}^*\Big)$$

$$p_{1,i,n} = J_{\gamma_n A_i}(s_{1,i,n} + \gamma_n z_i)$$

For $k = 1, \ldots, K$

$$s_{2,k,n} = v_{k,n}^* - \gamma_n\Big(D_k^{-1} v_{k,n}^* - \textstyle\sum_{i=1}^m L_{ki} x_{i,n}\Big)$$

$$p_{2,k,n} = s_{2,k,n} - \gamma_n\big(r_k + J_{\gamma_n^{-1} B_k}(\gamma_n^{-1} s_{2,k,n} - r_k)\big)$$

$$q_{2,k,n} = p_{2,k,n} - \gamma_n\Big(D_k^{-1} p_{2,k,n} - \textstyle\sum_{i=1}^m L_{ki} p_{1,i,n}\Big)$$

$$v_{k,n+1}^* = v_{k,n}^* - s_{2,k,n} + q_{2,k,n}$$

For $i = 1, \ldots, m$

$$q_{1,i,n} = p_{1,i,n} - \gamma_n\Big(C_i p_{1,i,n} + \textstyle\sum_{k=1}^K L_{ki}^* p_{2,k,n}\Big)$$

$$x_{i,n+1} = x_{i,n} - s_{1,i,n} + q_{1,i,n}$$

Some limitations of the state-of-the-art

We present a new framework that circumvents simultaneously the limitations of current methods, which require:

- inversions of linear operators or knowledge of bounds on norms of all the L_{ki}
- the proximal parameters must be the same for all the monotone operators
- activation of the resolvents of all the monotone operators: impossible in huge-scale problems
- synchronicity: all resolvent operator evaluations must be computed and used during the current iteration

and, in general,

- converge only weakly

Asynchronous, block-iterative splitting

- For every $i \in I$ (finite), $\mathcal{H}_i$ a Hilbert space, $A_i\colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ maximally monotone, $z_i^* \in \mathcal{H}_i$
- For every $k \in K$ (finite), $\mathcal{G}_k$ a Hilbert space, $B_k\colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ maximally monotone, $r_k \in \mathcal{G}_k$, $L_{ki} \in \mathscr{B}(\mathcal{H}_i, \mathcal{G}_k)$
- **Initial problem:** find $(\overline{x}_i)_{i\in I} \in \boldsymbol{\mathcal{H}} = \bigoplus_{i\in I} \mathcal{H}_i$ such that

$$(\forall i \in I) \quad z_i^* \in A_i\overline{x}_i + \sum_{k\in K} L_{ki}^*\Bigg(B_k\Bigg(\sum_{j\in I} L_{kj}\overline{x}_j - r_k\Bigg)\Bigg)$$

Asynchronous, block-iterative splitting

- For every $i \in I$ (finite), $\mathcal{H}_i$ a Hilbert space, $A_i \colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$ maximally monotone, $z_i^* \in \mathcal{H}_i$
- For every $k \in K$ (finite), $\mathcal{G}_k$ a Hilbert space, $B_k \colon \mathcal{G}_k \to 2^{\mathcal{G}_k}$ maximally monotone, $r_k \in \mathcal{G}_k$, $\mathsf{L}_{ki} \in \mathscr{B}(\mathcal{H}_i, \mathcal{G}_k)$
- **Initial problem:** find $(\overline{x}_i)_{i \in I} \in \boldsymbol{\mathcal{H}} = \bigoplus_{i \in I} \mathcal{H}_i$ such that

$$(\forall i \in I) \quad z_i^* \in A_i \overline{x}_i + \sum_{k \in K} L_{ki}^* \Bigg(B_k \Bigg(\sum_{j \in I} L_{kj} \overline{x}_j - r_k \Bigg) \Bigg)$$

- **Dual problem:** find $(\overline{v}_k^*)_{k \in K} \in \boldsymbol{\mathcal{G}} = \bigoplus_{k \in K} \mathcal{G}_k$ such that

$$(\forall k \in K) \quad -r_k \in -\sum_{i \in I} L_{ki} \Bigg(A_i^{-1} \Bigg(z_i^* - \sum_{l \in K} L_{li}^* \overline{v}_l^* \Bigg) \Bigg) + B_k^{-1} \overline{v}_k^*$$

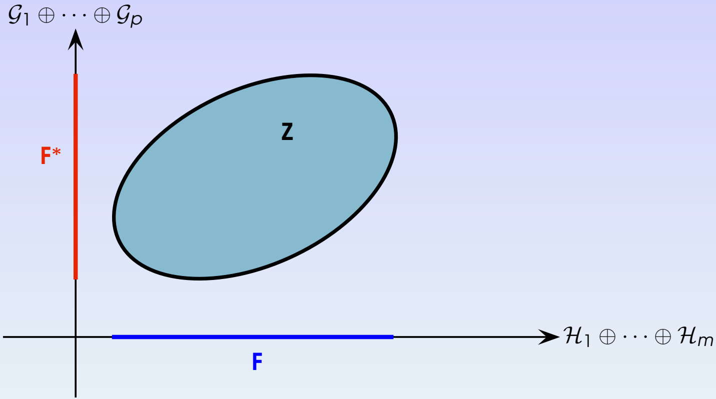
Asynchronous, block-iterative splitting

- **Solutions set:** the associated Kuhn-Tucker set

$$\boldsymbol{Z} = \Big\{ \big((\overline{x}_i)_{i\in I}, (\overline{v}_k^*)_{k\in K}\big) \;\Big|\; \overline{x}_i \in \mathcal{H}_i \;\text{ and }\; z_i^* - \sum_{k\in K} L_{ki}^* \overline{v}_k^* \in A_i \overline{x}_i, \quad \overline{v}_k^* \in \mathcal{G}_k \;\text{ and }\; \sum_{i\in I} L_{ki}\overline{x}_i - r_k \in B_k^{-1}\overline{v}_k^* \Big\}$$

- $\boldsymbol{Z}$ is a closed convex set
- The projection of $\boldsymbol{Z}$ onto $\boldsymbol{\mathcal{H}}$ is the set $\boldsymbol{F}$ of primal solutions
- The projection of $\boldsymbol{Z}$ onto $\boldsymbol{\mathcal{G}}$ is the set $\boldsymbol{F}^*$ of dual solutions

The Kuhn-Tucker set



With proper CQ, this framework includes..

- Let $\boldsymbol{F}$ be the set of solutions to the problem

$$\underset{(x_i)_{i\in I}\in\boldsymbol{\mathcal{H}}}{\text{minimize}} \sum_{i\in I}\big(f_i(x_i)-\langle x_i \mid z_i^*\rangle\big)+\sum_{k\in K} g_k\Big(\sum_{i\in I} L_{ki}x_i - r_k\Big)$$

where $f_i \in \Gamma_0(\mathcal{H}_i)$, $g_k \in \Gamma_0(\mathcal{G}_k)$, $L_{ki} \in \mathscr{B}(\mathcal{H}_i, \mathcal{G}_k)$

- Let $\boldsymbol{F}^*$ be the set of solutions to the dual problem

$$\underset{(v_k^*)_{k\in K}\in\bigoplus_{k\in K}\mathcal{G}_k}{\text{minimize}} \sum_{i\in I} f_i^*\Big(z_i^* - \sum_{k\in K} L_{ki}^* v_k^*\Big)+\sum_{k\in K}\big(g_k^*(v_k^*)+\langle v_k^* \mid r_k\rangle\big)$$

- Associated Kuhn-Tucker set

$$\boldsymbol{Z}=\Big\{\big((\overline{x}_i)_{i\in I},(\overline{v}_k^*)_{k\in K}\big)\;\Big|\;\overline{x}_i\in\mathcal{H}_i \text{ and } z_i^*-\sum_{k\in K}L_{ki}^*\overline{v}_k^*\in\partial f_i(\overline{x}_i),\quad \overline{v}_k^*\in\mathcal{G}_k \text{ and } \sum_{i\in I}L_{ki}\overline{x}_i - r_k \in \partial g_k^*(\overline{v}_k^*)\Big\}$$

Methodology for 2 operators

- $0 \in A\overline{x} + L^*\big(BL\overline{x}\big)$ and $0 \in -L\big(A^{-1}\big(-L^*\overline{v}^*\big)\big) + B^{-1}\overline{v}^*$

Methodology for 2 operators

- $0 \in A\overline{x} + L^*\big(BL\overline{x}\big)$ and $0 \in -L\big(A^{-1}\big(-L^*\overline{v}^*\big)\big) + B^{-1}\overline{v}^*$
- Take $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$. Then $-L^*\overline{v}^* \in A\overline{x}$ and $L\overline{x} \in B^{-1}\overline{v}^*$, i.e.,

$$\big(\overline{x}, -L^*\overline{v}^*\big) \in \operatorname{gra} A \quad \text{and} \quad \big(L\overline{x}, \overline{v}^*\big) \in \operatorname{gra} B$$

Methodology for 2 operators

- $0 \in A\overline{x} + L^*(BL\overline{x})$ and $0 \in -L(A^{-1}(-L^*\overline{v}^*)) + B^{-1}\overline{v}^*$
- Take $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$. Then $-L^*\overline{v}^* \in A\overline{x}$ and $L\overline{x} \in B^{-1}\overline{v}^*$, i.e.,

$$(\overline{x}, -L^*\overline{v}^*) \in \operatorname{gra} A \quad \text{and} \quad (L\overline{x}, \overline{v}^*) \in \operatorname{gra} B$$

- Suppose that at iteration n

$$(a_n, a_n^*) \in \operatorname{gra} A \quad \text{and} \quad (b_n, b_n^*) \in \operatorname{gra} B$$

Methodology for 2 operators

- $0 \in A\overline{x} + L^*(BL\overline{x})$ and $0 \in -L(A^{-1}(-L^*\overline{v}^*)) + B^{-1}\overline{v}^*$
- Take $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$. Then $-L^*\overline{v}^* \in A\overline{x}$ and $L\overline{x} \in B^{-1}\overline{v}^*$, i.e.,

$$(\overline{x}, -L^*\overline{v}^*) \in \operatorname{gra} A \quad \text{and} \quad (L\overline{x}, \overline{v}^*) \in \operatorname{gra} B$$

- Suppose that at iteration n

$$(a_n, a_n^*) \in \operatorname{gra} A \quad \text{and} \quad (b_n, b_n^*) \in \operatorname{gra} B$$

- By monotonicity of A and B,

$$\langle a_n - \overline{x} \mid a_n^* + L^*\overline{v}^* \rangle + \langle b_n - L\overline{x} \mid b_n^* - \overline{v}^* \rangle \geqslant 0$$

i.e.,

$$\langle (\overline{x}, \overline{v}) \mid (a_n^* + L^* b_n^*, b_n - La_n) \rangle \leqslant \langle a_n \mid a_n^* \rangle + \langle b_n \mid b_n^* \rangle$$

Methodology for 2 operators

- $0 \in A\overline{x} + L^*\big(BL\overline{x}\big)$ and $0 \in -L\big(A^{-1}\big(-L^*\overline{v}^*\big)\big) + B^{-1}\overline{v}^*$
- Take $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$. Then $-L^*\overline{v}^* \in A\overline{x}$ and $L\overline{x} \in B^{-1}\overline{v}^*$, i.e.,

$$\big(\overline{x}, -L^*\overline{v}^*\big) \in \operatorname{gra} A \quad \text{and} \quad \big(L\overline{x}, \overline{v}^*\big) \in \operatorname{gra} B$$

- Suppose that at iteration n

$$(a_n, a_n^*) \in \operatorname{gra} A \quad \text{and} \quad (b_n, b_n^*) \in \operatorname{gra} B$$

- By monotonicity of A and B,

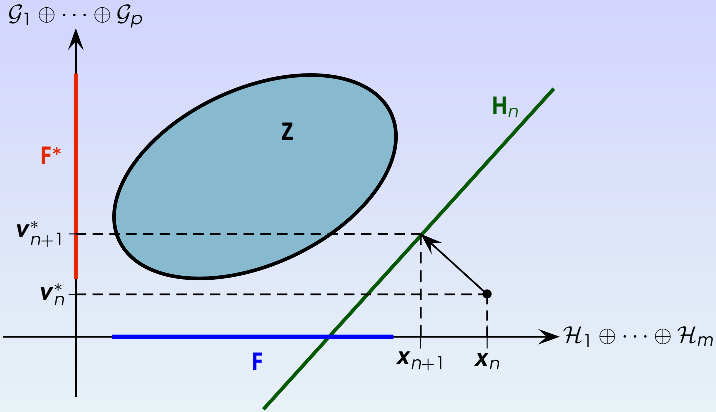
$$\langle a_n - \overline{x} \mid a_n^* + L^*\overline{v}^*\rangle + \langle b_n - L\overline{x} \mid b_n^* - \overline{v}^*\rangle \geqslant 0$$

i.e.,

$$\langle(\overline{x}, \overline{v}) \mid (a_n^* + L^*b_n^*, b_n - La_n)\rangle \leqslant \langle a_n \mid a_n^*\rangle + \langle b_n \mid b_n^*\rangle$$

- This places $\boldsymbol{Z}$ in the closed affine half-space with outer normal $(a_n^* + L^*b_n^*, b_n - La_n)$!

Fejér monotone scheme in the general case



- Choose suitable points in the graphs of $(A_i)_{i\in I}$ and $(B_k)_{k\in K}$ to construct a half-space H_n containing Z
- Algorithm: $(\boldsymbol{x}_{n+1}, \boldsymbol{v}^*_{n+1}) = P_{\mathsf{H}_n}(\boldsymbol{x}_n, \boldsymbol{v}^*_n) \rightharpoonup (\boldsymbol{x}, \boldsymbol{v}^*) \in \mathsf{Z} \subset \mathsf{F} \times \mathsf{F}^*$

Main novelties

- **Block iterations:** At iteration n, we require calculation of new points in the graphs of only some the operators $(A_i)_{i \in I_n \subset I}$ and $(B_k)_{k \in K_n \subset K}$. The control sequences $(I_n)_{n \in \mathbb{N}}$ and $(K_n)_{n \in \mathbb{N}}$ dictate how frequently the various operators are used.
- **Asynchronicity:** A new point $(a_{i,n}, a^*_{i,n}) \in \operatorname{gra} A_i$ being incorporated into the calculations at iteration n may be based on data $x_{i,c_i(n)}$ and $(v^*_{k,c_i(n)})_{k \in K}$ available at some possibly earlier iteration $c_i(n) \leqslant n$. Therefore, the calculation of $(a_{i,n}, a^*_{i,n})$ could have been initiated at iteration $c_i(n)$, with its results becoming available only at iteration n. Likewise, for $(b_{k,n}, b^*_{k,n}) \in \operatorname{gra} B_k$.

Also:

- No knowledge of the $\|L_{ki}\|$s is required
- No linear operator inversion is required
- No bounds required on the proximal parameters

Asynchronous block-iterative proximal splitting I

for $n = 0, 1, \ldots$

for every $i \in I_n$

$l_{i,n}^* = \sum_{k\in K} L_{ki}^* v_{k,c_i(n)}^*$

$(a_{i,n}, a_{i,n}^*) = \Big(J_{\gamma_{i,c_i(n)} A_i}\big(x_{i,c_i(n)} + \gamma_{i,c_i(n)}(z_i - l_{i,n}^*)\big), \gamma_{i,c_i(n)}^{-1}(x_{i,c_i(n)} - a_{i,n}) - l_{i,n}^*\Big)$

for every $i \in I \smallsetminus I_n$

$(a_{i,n}, a_{i,n}^*) = (a_{i,n-1}, a_{i,n-1}^*)$

for every $k \in K_n$

$l_{k,n} = \sum_{i\in I} L_{ki} x_{i,d_k(n)}$

$(b_{k,n}, b_{k,n}^*) = \Big(r_k + J_{\mu_{k,d_k(n)} B_k}\big(l_{k,n} + \mu_{k,d_k(n)} v_{k,d_k(n)}^* - r_k\big), v_{k,d_k(n)}^* + \mu_{k,d_k(n)}^{-1}(l_{k,n} - b_{k,n})\Big)$

for every $k \in K \smallsetminus K_n$

$(b_{k,n}, b_{k,n}^*) = (b_{k,n-1}, b_{k,n-1}^*)$

$\big((t_{i,n}^*)_{i\in I}, (t_{k,n})_{k\in K}\big) = \big((a_{i,n}^* + \sum_{k\in K} L_{ki}^* b_{k,n}^*)_{i\in I}, (b_{k,n} - \sum_{i\in I} L_{ki} a_{i,n})_{k\in K}\big)$

$\tau_n = \sum_{i\in I} \|t_{i,n}^*\|^2 + \sum_{k\in K} \|t_{k,n}\|^2$

if $\tau_n > 0$

$$\theta_n = \frac{\lambda_n}{\tau_n} \max\Big\{0, \sum_{i\in I} \big(\langle x_{i,n} \mid t_{i,n}^*\rangle - \langle a_{i,n} \mid a_{i,n}^*\rangle\big) + \sum_{k\in K} \big(\langle t_{k,n} \mid v_{k,n}^*\rangle - \langle b_{k,n} \mid b_{k,n}^*\rangle\big)\Big\}$$

else $\theta_n = 0$

for every $i \in I$

$x_{i,n+1} = x_{i,n} - \theta_n t_{i,n}^*$

for every $k \in K$

$v_{k,n+1}^* = v_{k,n}^* - \theta_n t_{k,n}$

Convergence

- $(I_n)_{n\in\mathbb{N}}$ is a sequence of nonempty subsets of I, and $(K_n)_{n\in\mathbb{N}}$ is a sequence of nonempty subsets of K such that $I_0 = I$, $K_0 = K$, and

$$(\forall n \in \mathbb{N}) \quad \left(\bigcup_{j=n}^{n+M-1} I_j = I \quad \text{and} \quad \bigcup_{j=n}^{n+M-1} K_j = K \right). \tag{1}$$

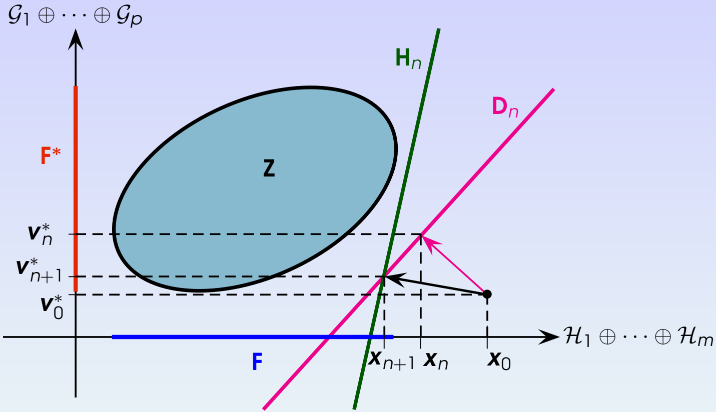
- $(c_i(n))_{n\in\mathbb{N}}$ and $(d_k(n))_{n\in\mathbb{N}}$ are sequences in $\mathbb{N}$ such that

$$(\forall i \in I) \;\; n - D \leqslant c_i(n) \leqslant n \quad \text{and} \quad (\forall k \in K) \;\; n - D \leqslant d_k(n) \leqslant n$$

- $\varepsilon \in \left]0, 1\right[$ and $(\gamma_{i,n})_{n\in\mathbb{N}}$ and $(\mu_{k,n})_{n\in\mathbb{N}}$ are sequences in $[\varepsilon, 1/\varepsilon]$.

Set $x_n = (x_{i,n})_{i\in I}$ and $v_n^* = (v_{k,n}^*)_{k\in K}$. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point $\overline{x} \in \boldsymbol{F}$, $(v_n^*)_{n\in\mathbb{N}}$ converges weakly to a point $\overline{v} \in \boldsymbol{F}^*$, and $(\overline{x}, \overline{v}^*) \in \boldsymbol{Z}$.

Asynchronous block-iterative proximal splitting II



- Construct $\mathbf{H}_n$ as before
- The half-space $\mathbf{D}_n$ satisfies $(\boldsymbol{x}_n, \boldsymbol{v}_n^*) = P_{\mathbf{D}_n}(\boldsymbol{x}_0, \boldsymbol{v}_0^*)$
- Algorithm: $(\boldsymbol{x}_{n+1}, \boldsymbol{v}_{n+1}^*) = P_{\mathbf{H}_n \cap \mathbf{D}_n}(\boldsymbol{x}_0, \boldsymbol{v}_0^*) \to P_{\mathbf{Z}}(\boldsymbol{x}_0, \boldsymbol{v}_0^*) \in \mathbf{F} \times \mathbf{F}^*$

DEUXIÈME PARTIE : ÉCLATEMENT DANS LES ESPACES DE BANACH

Legendre functions (Bauschke-Borwein-PLC, 2001)

- $\mathcal{X}$ is a reflexive real Banach space.
- $f \in \Gamma_0(\mathcal{X})$ is:
  - essentially smooth, if ∂f is both locally bounded and single-valued on its domain;
  - essentially strictly convex, if $(\partial f)^{-1}$ is locally bounded on its domain and f is strictly convex on every convex subset of $\operatorname{dom} \partial f$;
  - a Legendre function, if it is both essentially smooth and essentially strictly convex.

Legendre functions

If f is a Legendre function, then

- f^* is a Legendre function.
- $\operatorname{dom} \partial f = \operatorname{int} \operatorname{dom} f \neq \varnothing$ and f is Gâteaux differentiable on $\operatorname{int} \operatorname{dom} f$.
- $\nabla f \colon \operatorname{int} \operatorname{dom} f \to \operatorname{int} \operatorname{dom} f^*$ is bijective with inverse

$$\nabla f^* \colon \operatorname{int} \operatorname{dom} f^* \to \operatorname{int} \operatorname{dom} f.$$

- For every sequence $(x_n)_{n \in \mathbb{N}}$ in $\operatorname{int} \operatorname{dom} f$ that converges to some point in $\operatorname{bdry} \operatorname{dom} f$, we have

$$\|\nabla f(x_n)\| \to +\infty,$$

Legendre functions

- Let $f\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be a Legendre function.
- The associated Bregman distance is

$$D^f\colon \mathcal{X}\times\mathcal{X} \to [0,+\infty]$$
$$(x,y) \mapsto \begin{cases} f(x)-f(y)-\langle x-y, \nabla f(y)\rangle, & \text{if } \; y\in \operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{otherwise.}\end{cases}$$

- A sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{X}$ is Bregman-monotone with respect to $Z\subset\mathcal{X}$ if (Bauschke-Borwein-PLC, 2003)
  - $Z\cap\operatorname{dom} f\neq\varnothing$.
  - $(x_n)_{n\in\mathbb{N}}$ lies in $\operatorname{int}\operatorname{dom} f$.
  - $(\forall x\in Z\cap\operatorname{dom} f)(\forall n\in\mathbb{N})\;\; D(x,x_{n+1})\leqslant D(x,x_n)$.

How to generate a Bregman monotone sequence?

- For every x and y in $\operatorname{int}\operatorname{dom} f$, set

$$H(x,y) = \big\{z \in \mathcal{X} \mid \langle z - y, \nabla f(x) - \nabla f(y)\rangle \leqslant 0\big\}.$$

- Call $\mathfrak{B}$ the class of all operators $T\colon \mathcal{X} \to 2^{\mathcal{X}}$ such that

$$\operatorname{ran} T \subset \operatorname{dom} T = \operatorname{intdom} f, \text{ and } (\forall (x,u) \in \operatorname{gra} T)\ \operatorname{Fix} T \subset H(x,u).$$

- Set

$$x_0 \in \operatorname{int}\operatorname{dom} f \quad \text{and} \quad (\forall n \in \mathbb{N})\ \ x_{n+1} \in T_n x_n, \quad \text{where} \quad T_n \in \mathfrak{B}.$$

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

$$Z = \bigcap_{n\in\mathbb{N}} \operatorname{Fix} T_n.$$

Bregman Projection

- Let $f\colon \mathcal{X} \to \,]-\infty, +\infty]$ be a Legendre function.
- Let C be a closed convex subset of $\mathcal{X}$ such that $C \cap \operatorname{int}\operatorname{dom} \neq \varnothing$.
- The Bregman projector onto C induced by f is

$$
\begin{aligned}
P_C^f \colon \operatorname{int}\operatorname{dom} f &\to C \cap \operatorname{int}\operatorname{dom} f \\
y &\mapsto \underset{x \in C}{\operatorname{argmin}}\, D^f(x, y).
\end{aligned}
$$

- $P_C^f \in \mathfrak{B}$.

Bregman Resolvent

- Let $f\colon \mathcal{X} \to]-\infty, +\infty]$ be a Legendre function.
- Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone.
- The Bregman resolvent of A induced by f is

$$R_A^f = (\nabla f + A)^{-1} \circ \nabla f\colon \operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f.$$

- $R_A^f \in \mathfrak{B}$.
- The Bregman proximal operator of $\varphi \in \Gamma_0(\mathcal{X})$ induced by f is

$$\operatorname{prox}_\varphi^f = R_{\partial\varphi}^f\colon \operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f$$
$$y \mapsto \underset{x\in\mathcal{X}}{\operatorname{argmin}}\, \varphi(x) + D^f(x,y).$$

Note: Even in Euclidean spaces, it may be easier to evaluate $\operatorname{prox}_\varphi^f$ than the usual Moreau proximity operator $\operatorname{prox}_\varphi = (\operatorname{Id} + \gamma\partial\varphi)^{-1}$. Here are a few examples...

Examples of Bregman proximity operators (PLC-Nguyen, 2016)

Suppose that f is the Fermi-Dirac entropy function on $\mathbb{R}^m$.

- Let $\omega \in \mathbb{R}$ and suppose that each component of φ is

$$\phi\colon \xi \mapsto \begin{cases} \xi \ln \xi - \omega\xi, & \text{if } \xi \in \left]0,+\infty\right[; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases}$$

  Then $\eta_i = -e^{\xi_i+\omega-1}/2 + \sqrt{e^{2(\xi_i+\omega-1)}/4 + e^{\xi_i+\omega-1}}$.

- Suppose that

$$\phi\colon \xi \mapsto \begin{cases} (1-\xi)\ln(1-\xi) + \xi, & \text{if } \xi \in \left]-\infty,1\right[; \\ 1, & \text{if } \xi = 1; \\ +\infty, & \text{otherwise.} \end{cases}$$

  Then $\eta_i = 1 + e^{-\xi_i}/2 - \sqrt{e^{-\xi_i} + e^{-2\xi_i}/4}$.

Problem

- Let $\mathcal{X}$ and $\mathcal{Y}$ be reflexive real Banach spaces, let $\boldsymbol{\mathcal{X}} = \mathcal{X} \times \mathcal{Y}^*$ be normed with $(x, y^*) \mapsto \sqrt{\|x\|^2 + \|y^*\|^2}$, and let $\boldsymbol{\mathcal{X}}^*$ be its topological dual, that is, $\mathcal{X}^* \times \mathcal{Y}$ equipped with the norm $(x^*, y) \mapsto \sqrt{\|x^*\|^2 + \|y\|^2}$.
- Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $B\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ be maximally monotone, and let $L\colon \mathcal{X} \to \mathcal{Y}$ be linear and bounded.
- Primal problem:

$$\text{find} \quad x \in \mathcal{X} \quad \text{such that} \quad 0 \in Ax + L^*BLx.$$

- Dual problem:

$$\text{find} \quad y^* \in \mathcal{Y}^* \quad \text{such that} \quad 0 \in -LA^{-1}(-L^*y^*) + B^{-1}y^*.$$

- Kuhn-Tucker set:

$$\boldsymbol{Z} = \big\{(x, y^*) \in \boldsymbol{\mathcal{X}} \mid -L^*y^* \in Ax \quad \text{and} \quad Lx \in B^{-1}y^*\big\}.$$

Problem

- Let $f\colon \mathcal{X} \to]-\infty,+\infty]$ and $g\colon \mathcal{Y} \to]-\infty,+\infty]$ be Legendre functions, set

$$\boldsymbol{f}\colon \mathcal{X} \to]-\infty,+\infty]\colon (x,y^*) \mapsto f(x)+g^*(y^*),$$

  let $x_0 \in \operatorname{int}\operatorname{dom} f$, let $y_0^* \in \operatorname{int}\operatorname{dom} g^*$.
- Suppose that $\boldsymbol{Z} \cap \operatorname{int}\operatorname{dom} \boldsymbol{f} \neq \varnothing$.
- The problem is to find the best Bregman approximation $(\overline{x},\overline{y}^*) = P^{\boldsymbol{f}}_{\boldsymbol{Z}}(x_0,y_0^*)$ to (x_0,y_0^*) from $\boldsymbol{Z}$.

Methodology

- Use the same geometrical intuition as in the hilbertian case.
- Thus, in the fully parallel case we iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
(\gamma_n, \mu_n) \in [\varepsilon, \sigma] \times [\delta, \sigma] \\
a_n = R^h_{\gamma_n A}\big(\nabla h(x_n) - \gamma_n L^* y_n^*\big) \\
a_n^* = \gamma_n^{-1}\big(\nabla h(x_n) - \nabla h(a_n)\big) - L^* y_n^* \\
b_n = R^j_{\mu_n B}\big(\nabla j(L x_n) + \mu_n y_n^*\big) \\
b_n^* = \mu_n^{-1}\big(\nabla j(L x_n) - \nabla j(b_n)\big) + y_n^* \\
\boldsymbol{H}_n = \Big\{(x, y^*) \in \mathfrak{X} \,\Big|\, \langle x, a_n^* + L^* b_n^* \rangle + \langle b_n - L a_n, y^* \rangle \\
\qquad\qquad\qquad\qquad\qquad \leqslant \langle a_n, a_n^* \rangle + \langle b_n, b_n^* \rangle \Big\} \\
(x_{n+1/2}, y_{n+1/2}^*) = P^{\boldsymbol{f}}_{\boldsymbol{H}_n}(x_n, y_n^*) \\
(x_{n+1}, y_{n+1}^*) = Q^{\boldsymbol{f}}\big((x_0, y_0^*), (x_n, y_n^*), (x_{n+1/2}, y_{n+1/2}^*)\big).
\end{array}
\right.
\end{array}
$$

where

$$
Q^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}, \boldsymbol{y}) = P^{\boldsymbol{f}}_{H^{\boldsymbol{f}}(\boldsymbol{x}_0, \boldsymbol{x}) \cap H^{\boldsymbol{f}}(\boldsymbol{x}, \boldsymbol{y})} \boldsymbol{x}_0.
$$

References

- PLC and J. Eckstein, Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions, *Math. Program.*, 2017.
- PLC and Q. V. Nguyen, Solving composite monotone inclusions in reflexive Banach spaces by constructing best Bregman approximations from their Kuhn-Tucker set, *J. Convex Anal.*, vol. 23, 2016.
- H. Attouch, L. M. Briceño-Arias, PLC, A strongly convergent primal-dual method for nonoverlapping domain decomposition, *Numer. Math.*, vol. 133, 2016.
- PLC, Systems of structured monotone inclusions: Duality, algorithms, and applications, *SIAM J. Optim.*, vol. 23, 2013.
- H. H. Bauschke, J. M. Borwein, and PLC, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, 2003.
- H. H. Bauschke and PLC, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* **2nd ed.**, Springer, New York, 2017.