A Splitting Algorithm for Coupled System of Primal–Dual Monotone Inclusions

B. C. VŨ

Sestri Levante, Italia, 2014

Outline

Extended set of solution

- $\boldsymbol{A}\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone on a real Hilbert space $\mathcal{H}$.
- $\boldsymbol{C}\colon \mathcal{H} \to 2^{\mathcal{H}}$ is maximally monotone.
- Suppose that $0 \in \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{C}\overline{\boldsymbol{x}}$ then there exists $\overline{\boldsymbol{y}} \in \boldsymbol{C}\overline{\boldsymbol{x}}$ such that $-\boldsymbol{y} \in \boldsymbol{A}\overline{\boldsymbol{x}}$. Define extended set of solution

$$Ext = \big\{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{H}^2 \mid -\boldsymbol{y} \in \boldsymbol{A}\boldsymbol{x} \text{ and } \boldsymbol{y} \in \boldsymbol{C}\boldsymbol{x}\big\} \neq \varnothing$$

- Examples: Forward-backward, Forward-backward-forward, Douglas-Racford.

Extended set of solution (Breceno-Combettes 2011)

- $\boldsymbol{B}\colon \mathcal{G} \to 2^{\mathcal{G}}$ is maximally monotone on Hilbert space $\mathcal{G}$.
- $\boldsymbol{L}\colon \mathcal{H} \to \mathcal{G}$ is a linear and bounded.
- Suppose that $0 \in \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{L}^* \circ \boldsymbol{B} \circ \boldsymbol{L}\overline{\boldsymbol{x}}$. Define extended set of solution

$$Ext = \big\{(\boldsymbol{x}, \boldsymbol{v}) \in \mathcal{H} \times \mathcal{G} \mid -\boldsymbol{L}^*\boldsymbol{v} \in \boldsymbol{A}\boldsymbol{x} \text{ and } \boldsymbol{L}\boldsymbol{x} \in \boldsymbol{B}^{-1}\boldsymbol{v}\big\} \neq \varnothing$$

- Examples: Primal-Dual Forward-Backward-Forward.

Extended set of solution (Combettes-Pesquet 2011)

- $\boldsymbol{D}\colon \mathcal{G} \to 2^{\mathcal{G}}$ is maximally monotone.
- $\boldsymbol{B} \,\square\, \boldsymbol{D} = (\boldsymbol{B}^{-1} + \boldsymbol{D}^{-1})^{-1}$.
- Suppose that $0 \in \boldsymbol{A}\overline{\boldsymbol{x}} + \boldsymbol{L}^* \circ (\boldsymbol{B} \,\square\, \boldsymbol{D}) \circ \boldsymbol{L}\overline{\boldsymbol{x}} + \boldsymbol{C}\overline{\boldsymbol{x}}$. Define extended set of solution

$$E = \big\{(\boldsymbol{x}, \boldsymbol{v}) \in \mathcal{H} \times \mathcal{G} \mid -\boldsymbol{L}^*\boldsymbol{v} \in (\boldsymbol{A} + \boldsymbol{C})\overline{\boldsymbol{x}} \text{ and } \boldsymbol{L}\boldsymbol{x} \in (\boldsymbol{B}^{-1} + \boldsymbol{D}^{-1})\boldsymbol{v}\big\} \neq$$

- Examples: Primal-Dual Forward-Backward-Forward.

System of monotone inclusions

- $\boldsymbol{\mathcal{H}} = \mathcal{H}_1 \times \ldots \times \mathcal{H}_m$.
- Primal coupling: $C_i \colon \boldsymbol{\mathcal{H}} \to \mathcal{H}_i$, $(\boldsymbol{C} \colon \boldsymbol{\mathcal{H}} \to \boldsymbol{\mathcal{H}} \colon \boldsymbol{x} \mapsto (C_i\boldsymbol{x})_{1\leq i\leq m})$

$$\langle \boldsymbol{x} - \boldsymbol{y} \mid \boldsymbol{C}\boldsymbol{x} - \boldsymbol{C}\boldsymbol{y} \rangle \geq \nu_0 \|\boldsymbol{C}\boldsymbol{x} - \boldsymbol{C}\boldsymbol{y}\|^2.$$

- $\boldsymbol{\mathcal{G}} = \mathcal{G}_1 \times \ldots \times \mathcal{G}_s$.
- Dual coupling: $S_i \colon \boldsymbol{\mathcal{H}} \to \mathcal{H}_i$, $(\boldsymbol{D} \colon \boldsymbol{\mathcal{G}} \to \boldsymbol{\mathcal{G}} \colon \boldsymbol{v} \mapsto (S_i\boldsymbol{v})_{1\leq i\leq s})$

$$\langle \boldsymbol{v} - \boldsymbol{w} \mid \boldsymbol{D}\boldsymbol{v} - \boldsymbol{D}\boldsymbol{w} \rangle \geq \mu_0 \|\boldsymbol{D}\boldsymbol{v} - \boldsymbol{D}\boldsymbol{w}\|^2.$$

System of monotone inclusions

- Linear operators $\boldsymbol{L} = ((L_{k,i})_{1\leq k\leq s})_{1\leq i\leq m}$, where $L_{k,i}\colon \mathcal{H}_i \to \mathcal{G}_k$.
- Maximally monotone $\boldsymbol{A} = \times_{i=1}^m A_i$, where $A_i\colon \mathcal{H}_i \to 2^{\mathcal{H}_i}$.
- Maximally monotone $\boldsymbol{B} = \times_{i=1}^m B_i$, where $B_i\colon \mathcal{G}_i \to 2^{\mathcal{G}_i}$.
- Problem:

$$\begin{cases} z_1 - \sum_{k=1}^{s} L_{k,1}^*\overline{v}_k \in A_1\overline{x}_1 + C_1\overline{\boldsymbol{x}} \\ \vdots \\ z_m - \sum_{k=1}^{s} L_{k,m}^*\overline{v}_k \in A_m\overline{x}_m + C_m\overline{\boldsymbol{x}} \end{cases} \quad \text{and} \quad \begin{cases} \sum_{i=1}^{m} L_{1,i}\overline{x}_i - r_1 \in B_1\overline{v}_1 + S_1\overline{\boldsymbol{v}} \\ \vdots \\ \sum_{i=1}^{m} L_{s,i}\overline{x}_i - r_s \in B_s\overline{v}_s + S_s\overline{\boldsymbol{v}}. \end{cases}$$

- Equivalently: $\boldsymbol{z} - \boldsymbol{L}^*\overline{\boldsymbol{v}} \in (\boldsymbol{A}+\boldsymbol{C})\overline{\boldsymbol{x}}$ and $\boldsymbol{L}\overline{\boldsymbol{x}} - \boldsymbol{r} \in (\boldsymbol{B}+\boldsymbol{D})\overline{\boldsymbol{v}}$.

Ex1: Attouch- Briceño-Arias-Combettes 2010

- Set $\boldsymbol{L} = ((L_{k,i})_{1\leq k\leq s})_{1\leq i\leq m} = 0.$

$$\begin{cases} z_1 \in A_1\overline{x}_1 + C_1\overline{\boldsymbol{x}} \\ \vdots \\ z_m \in A_m\overline{x}_m + C_m\overline{\boldsymbol{x}} \end{cases} \quad \text{and} \quad \begin{cases} -r_1 \in B_1\overline{v}_1 + S_1\overline{\boldsymbol{v}} \\ \vdots \\ -r_s \in B_s\overline{v}_s + S_s\overline{\boldsymbol{v}}. \end{cases}$$

Ex2: Combettes 2013 $C_i(\boldsymbol{x}) = C_i(x_i)$ and $S_k(\boldsymbol{v}) = S_k(v_k)$.

- Primal inclusion ($R_k = S_k^{-1}$), ($A \Box B = (A^{-1} + B^{-1})^{-1}$)

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

- Dual inclusion:

$$\begin{cases} -r_1 \in B_1\overline{v}_1 + \sum_{i=1}^{m} L_{1,i}\Big((A_i \,\Box\, C_i)^{-1}\Big(z_i - \sum_{i=1}^{m} L_{k,i}^{*}\overline{v}_i\Big)\Big) + S_1\overline{v}_1 \\ \vdots \\ -r_s \in B_s\overline{v}_s + \sum_{i=1}^{m} L_{s,i}\Big((A_i \,\Box\, C_i)^{-1}\Big(z_i - \sum_{i=1}^{m} L_{k,i}^{*}\overline{v}_i\Big)\Big) + S_s\overline{v}_s. \end{cases}$$

Ex3: Boţ-Csetnek-Nagy 2013

- Solve system

$$\begin{cases} 0 \in L_{1,1}^*\big((R_1 \,\square\, B_1)(L_{1,1}\overline{x}_1)\big) + C_1\overline{\boldsymbol{x}} \\ \vdots \\ 0 \in L_{m,m}^*\big((R_m \,\square\, B_m)(L_{m,m}\overline{x}_m)\big) + C_m\overline{\boldsymbol{x}}. \end{cases}$$

Some features of the algorithm proposed

- Use the approximations of the resolvents of the set-valued operators: $J_{UA} = (\mathrm{Id} + UA)^{-1}$ (see Combettes-Vu).
- Use the approximations of the single-valued operators C_i and S_k in the sense of Attouch- Briceño-Arias-Combettes 2010.
  - $(C_{i,n})_{n\in\mathbb{N}}$ are operators from $\boldsymbol{\mathcal{H}}$ to $\mathcal{H}_i$ such that
    (b1) $C_{i,n} - C_i$ is κ_n-Lipschitz continuous satisfying
    $$\sum_{n\in\mathbb{N}} \kappa_{i,n} < +\infty \text{ and } (\exists \overline{\boldsymbol{s}} \in \boldsymbol{\mathcal{H}})(\forall n \in \mathbb{N}) \quad C_{i,n}\overline{\boldsymbol{s}} = C_i\overline{\boldsymbol{s}}.$$
  - $(S_{k,n})_{n\in\mathbb{N}}$ are operators from $\boldsymbol{\mathcal{G}}$ to $\mathcal{G}_k$ such that
    (c1) $S_{k,n} - S_k$ is η_n-Lipschitz continuous satisfying
    $$\sum_{n\in\mathbb{N}} \eta_{k,n} < +\infty \text{ and } (\exists \overline{\boldsymbol{w}} \in \boldsymbol{\mathcal{G}})(\forall n \in \mathbb{N}) \quad S_{k,n}\overline{\boldsymbol{w}} = S_k\overline{\boldsymbol{w}}.$$
- Allow the metric vary over course of iterations (see Combettes-Vu): $(\forall n \in \mathbb{N})$ $U_{i,n+1} \succeq U_{i,n}$ and $V_{k,n+1} \succeq V_{k,n}$, and
  $$\mu := \sup_{n\in\mathbb{N}}\{\|U_{1,n}\|, \ldots, \|U_{m,n}\|, \|V_{1,n}\|, \ldots, \|V_{s,n}\|\} < +\infty.$$

Algorithms

- Write $a_n \approx b_n \iff \sum_{n\in\mathbb{N}} \|b_n - a_n\| < +\infty$.
- Algorithm

$$
\left\lfloor
\begin{array}{l}
(i)\ \text{For } i = 1, \ldots, m \\
\quad \left\lfloor
\begin{array}{l}
1.\ t_{i,n} \approx \sum_{k=1}^{s} L_{k,i}^{*} v_{k,n} + C_{i,n}(x_{1,n}, \ldots, x_{m,n}) \\
2.\ p_{i,n} \approx J_{U_{i,n}A_i}\Big(x_{i,n} - U_{i,n}(t_{i,n} - z_i)\Big) \\
3.\ y_{i,n} \approx 2p_{i,n} - x_{i,n} \\
4.\ x_{i,n+1} \approx x_{i,n} + \lambda_n(p_{i,n} - x_{i,n})
\end{array}
\right. \\
(ii)\ \text{For } k = 1, \ldots, s \\
\quad \left\lfloor
\begin{array}{l}
1.\ w_{k,n} \approx \sum_{i=1}^{m} L_{k,i} y_{i,n} - S_{k,n}(v_{1,n}, \ldots, v_{s,n}) \\
2.\ q_{k,n} \approx J_{V_{k,n}B_k}\Big(v_{k,n} + V_{k,n}(w_{k,n} - r_k)\Big) \\
3.\ v_{k,n+1} \approx v_{k,n} + \lambda_n(q_{k,n} - v_{k,n}),
\end{array}
\right.
\end{array}
\right.
$$

- This is a non-trivial extention of the algorithm in Combettes-Vu (2013) from univariate to multi-variate C_i, S_k, from C_i, S_k to $C_{i,n}$, $S_{k,n}$.

Weak convergence

- Assume that the set of solution is non empty and $\exists L_{k_0,i_0} \neq 0$.
- Define

$$\delta_n = \left(\sqrt{\sum_{i=1}^{m}\sum_{k=1}^{s}\left\|\sqrt{V_{k,n}}L_{k,i}\sqrt{U_{i,n}}\right\|^2}\right)^{-1} - 1,$$

- Set $\beta = \min\{\mu_0, \nu_0\}$ and suppose that

$$\zeta_n := \frac{\delta_n}{(1+\delta_n)\max_{1\leq i\leq m, 1\leq k\leq s}\{\|U_{i,n}\|, \|V_{k,n}\|\}} \geq \frac{1}{2\beta - \varepsilon}.$$

- Then, $(x_{1,n}, \dots x_{m,n}, v_{1,n}, \dots v_{s,n})_{n\in\mathbb{N}}$ coverges strongly to a solution $(\overline{x}_1, \dots, \overline{x}_m, \overline{v}_1, \dots, \overline{v}_s)$.

Strong convergence

- Recall (Attouch-Briceno-Combettes): A set-valued operator A is said to be demiregular at $x \in \operatorname{dom} A$ if for every $v \in Ax$ and

$$\begin{cases} x_n \rightharpoonup x \\ Ax_n \ni v_n \to v \end{cases} \implies x_n \to x.$$

- If $\boldsymbol{C}$ is demiregular at $(\overline{x}_1, \ldots, \overline{x}_m)$, $(x_{1,n}, \ldots x_{m,n}) \to (\overline{x}_1, \ldots, \overline{x}_m)$.
- If $\boldsymbol{D}$ is demiregular at $(\overline{v}_1, \ldots, \overline{v}_s)$, $(v_{1,n}, \ldots v_{s,n}) \to (\overline{v}_1, \ldots, \overline{v}_s)$.
- Suppose that there exists $j \in \{1, \ldots, m\}$ and $C\colon \mathcal{H}_j \to \mathcal{H}_j$ such that $(\forall (x_i)_{1\leq i\leq m} \in (\mathcal{H}_i)_{1\leq i\leq m})\ C_j(x_1, \ldots, x_m) = Cx_j$ and C is demiregular at $\overline{x}_j$, then $x_{j,n} \to \overline{x}_j$.
- Suppose that there exists $j \in \{1, \ldots, s\}$ and $D\colon \mathcal{G}_j \to \mathcal{G}_j$ such that $(\forall (v_k)_{1\leq k\leq s} \in (\mathcal{G}_k)_{1\leq k\leq s})\ S_j(v_1, \ldots, v_s) = Dv_j$ and D is demiregular at $\overline{v}_j$, then $v_{j,n} \to \overline{v}_j$.

non-smooth coupling + smooth coupling + prior information

- Smooth coupling (see also Attouch-Breceno-Combettes): $\varphi\colon \mathcal{H}_1 \times \ldots \times \mathcal{H}_m \to \mathbb{R}$ is a convex differentiable function with ν_0^{-1}-Lipschitz continuous gradient.
- Non-smooth coupling (Combettes 2013): $\ell_k, g_k \in \Gamma_0(\mathcal{G}_k)$, $\ell_k \,\square\, g_k : x \mapsto \inf_{y\in\mathcal{G}}(\ell_k(x-y)+g_k(y))$

$$\psi\colon (x_1,\ldots,x_m) \mapsto \sum_{k=1}^{s} (\ell_k \,\square\, g_k)\Big(\sum_{i=1}^{m} L_{k,i}x_i - r_k\Big)$$

- f_i model the prior information on $\overline{x}_i$.
- Primal problem:

$$\underset{x_1\in\mathcal{H}_1,\ldots,x_m\in\mathcal{H}_m}{\text{minimize}} \sum_{i=1}^{m}\big(f_i(x_i) - \langle x_i \mid z_i\rangle\big) + \psi(x_1,\ldots,x_m) + \varphi(x_1,\ldots,x_m),$$

Non-smooth coupling + smooth coupling + prior information

- $f^*\colon a \mapsto \sup_x(\langle a \mid x\rangle - f(x))$.
- Dual problem:

$$\underset{v_1\in\mathcal{G}_1,\ldots,v_s\in\mathcal{G}_s}{\text{minimize}} \left(\varphi^* \,\square\, \left(\sum_{i=1}^m f_i^*\right)\right)\left(\left(z_i - \sum_{k=1}^s L_{k,i}^* v_k\right)_{1\leq i\leq m}\right) + \sum_{k=1}^s \left(\ell_k^*(v_k) + g_k^*(v_k) + \langle v_k \mid r_k\rangle\right).$$

- Assumption: $\exists \overline{\boldsymbol{x}} = (\overline{x}_1,\ldots,\overline{x}_m)$ such that, for every $i \in \{1,\ldots,m\}$,

$$z_i \in \partial f_i(\overline{x}_i) + \sum_{k=1}^s L_{k,i}^* \circ \left(\partial\ell_k \,\square\, \partial g_k\right) \circ \left(\sum_{j=1}^m L_{k,j}\overline{x}_j - r_k\right) + \nabla_i\varphi(\overline{\boldsymbol{x}}),$$

where $\nabla_i\varphi$ is the *ith* component of the gradient $\nabla\varphi$,

Non-smooth coupling + smooth coupling + prior information

- Operator proximal:

$$J_{U^{-1}\partial f} := \mathrm{prox}_f^U : \mathcal{H} \to \mathcal{H} : x \mapsto \underset{y\in\mathcal{H}}{\arg\min} \left(f(y) + \frac{1}{2}\|x-y\|_U^2\right)$$

- $(\varphi_n)_{n\in\mathbb{N}}$ are differentiable functions in $\Gamma_0(\mathcal{H}_1 \times \ldots \times \mathcal{H}_m)$ such that

  1. $\nabla_i\varphi_n - \nabla_i\varphi$ is κ_n Lipschitz continuous satisfying $\sum_{n\in\mathbb{N}} \kappa_{i,n} < +\infty$ and $(\exists \overline{\boldsymbol{s}} \in \boldsymbol{\mathcal{H}})(\forall n \in \mathbb{N})\nabla_i\varphi_n(\overline{\boldsymbol{s}}) = \nabla_i\varphi(\overline{\boldsymbol{s}})$.

- $(\widetilde{\ell}_{k,n})_{n\in\mathbb{N}}$ are differentiable functions in $\Gamma_0(\mathcal{G}_k)$ such that

  1. $(\nabla\widetilde{\ell}_{k,n} - \nabla\ell_k^*)_{n\in\mathbb{N}}$ is $\eta_{k,n}$ Lipschitz continuous satisfying $\sum_{n\in\mathbb{N}} \eta_{k,n} < +\infty$ and $(\exists \overline{\boldsymbol{w}} := (\overline{w}_j)_{1\leq j\leq s} \in \boldsymbol{\mathcal{G}})(\forall n \in \mathbb{N})\nabla\widetilde{\ell}_{k,n}(\overline{w}_k) = \nabla\ell_k^*(\overline{w}_k)$.

Algorithms

- Algorithm

$$
\left|\begin{array}{l}
(i)\ \text{For } i = 1,\ldots,m \\
\quad \left|\begin{array}{l}
1.\ t_{i,n} \approx \sum_{k=1}^{s} L_{k,i}^{*} v_{k,n} + \nabla_i \varphi_n(x_{1,n},\ldots,x_{m,n}) \\
2.\ p_{i,n} \approx \operatorname{prox}_{f_i}^{U_{i,n}^{-1}} \big(x_{i,n} - U_{i,n}(t_{i,n} - z_i)\big) \\
3.\ y_{i,n} := 2p_{i,n} - x_{i,n} \\
4.\ x_{i,n+1} := x_{i,n} + \lambda_n(p_{i,n} - x_{i,n})
\end{array}\right. \\
(ii)\ \text{For } k = 1,\ldots,s \\
\quad \left|\begin{array}{l}
1.\ w_{k,n} \approx \sum_{i=1}^{m} L_{k,i} y_{i,n} - \nabla \widetilde{\ell}_{k,n}(v_{k,n}) \\
2.\ q_{k,n} \approx \operatorname{prox}_{g_k^{*}}^{V_{k,n}^{-1}} \big(v_{k,n} + V_{k,n}(w_{k,n} - r_k)\big) \\
3.\ v_{k,n+1} := v_{k,n} + \lambda_n(q_{k,n} - v_{k,n}),
\end{array}\right.
\end{array}\right.
$$

Weak convergence

- Assume that the set of solution is non empty and $\exists L_{k_0,i_0} \neq 0$.
- Define

$$\delta_n = \left(\sqrt{\sum_{i=1}^{m}\sum_{k=1}^{s}\left\|\sqrt{V_{k,n}}L_{k,i}\sqrt{U_{i,n}}\right\|^2}\right)^{-1} - 1,$$

- Set $\beta = \min\{\mu_0, \nu_0\}$ and suppose that

$$\zeta_n := \frac{\delta_n}{(1+\delta_n)\max_{1\leq i\leq m, 1\leq k\leq s}\{\|U_{i,n}\|, \|V_{k,n}\|\}} \geq \frac{1}{2\beta - \varepsilon}.$$

- Then, $(x_{1,n}, \ldots x_{m,n}, v_{1,n}, \ldots v_{s,n})_{n\in\mathbb{N}}$ coverges strongly to a solution $(\overline{x}_1, \ldots, \overline{x}_m, \overline{v}_1, \ldots, \overline{v}_s)$.

Strong convergence

- If φ is uniformly convex at $(\overline{x}_1, \ldots, \overline{x}_m)$, then $(x_{1,n}, \ldots x_{m,n}) \to (\overline{x}_1, \ldots, \overline{x}_m)$.
- If $\varphi(x_1, \ldots, x_m) = h_1(x_1) + \ldots h_m(x_m)$ for some convex differentiable h_i and h_j is uniformly convex at $\overline{x}_j$, then $x_{j,n} \to \overline{x}_j$.
- Suppose that ℓ_j^* is uniformly convex at $\overline{v}_j$, for some $j \in \{1, \ldots, s\}$, then $v_{j,n} \to \overline{v}_j$.

Open problem and future work

- Implementation of Algorithm recovery of signal over dictionarry with non trivial $U_{i,n}$ and $V_{k,n}$?
- Rate of convergence in expectation of the iterations and of the function-valued for strongly smooth coupling and their unbiased estimations are known (See also, Lorenzo-Silva-Vu, Combettes-Pesquet, Pesquet-Repetti)?
- $U_{i,n}$ and $V_{k,n}$ are random vectors with values in $\mathcal{P}_\alpha(\mathcal{H}_i)$ and $\mathcal{P}_\alpha(\mathcal{G}_k)$, respectively, and with with unbiased of smooth coupling. (see Bennar 2006 for the case of stochastic projected gradient)

Main references

- Attouch, H., Briceño-Arias, L.M., Combettes, P.L.: A parallel splitting method for coupled monotone inclusions. SIAM J. Control Optim. **48**, 3246–3270 (2010)
- Combettes, P.L., Vũ, B.C.: Variable metric forward-backward splitting with applications to monotone inclusions in duality. Optimization, **63**, pp. 1289–1318 (2014)
- Combettes, P.L.: Systems of structured monotone inclusions: duality, algorithms, and applications. SIAM J. Optim., vol. 23, pp. 24202447.
- VŨ, B. C, A splitting algorithm for coupled system of primaldual monotone inclusions. Journal of Optimization Theory and Applications, 1-33, 2014.