Stochastic Classical and Block-Iterative Splitting Methods.

Javier I. Madariaga
– joint work with Patrick L. Combettes –

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

Optimization and Equilibrium Seminar
December 22, 2025

Problem statement

- H is a separable real Hilbert space.
- $\mathsf{M}\colon \mathsf{H} \to 2^{\mathsf{H}}$ is a monotone operator:

$$(\forall \mathsf{x} \in \mathsf{H})(\forall \mathsf{y} \in \mathsf{H})(\forall \mathsf{x}^* \in \mathsf{M}\mathsf{x})(\forall \mathsf{y}^* \in \mathsf{M}\mathsf{y}) \quad \langle x - y \mid x^* - y^* \rangle \geqslant 0.$$

- **Problem:** Find $\bar{\mathsf{x}} \in \mathsf{H}$ such that $0 \in \mathsf{M}\bar{\mathsf{x}}$.
- **Applications:**
  - Optimization
  - PDEs
  - Control theory
  - Machine learning
  - Signal processing
  - Game theory
  - etc.

Splitting methods

- M is the sum of operators with specific attributes, e.g.,
  - $W\colon H \to 2^H$ be maximally monotone.
  - $C\colon H \to H$ be α-cocoercive.
  - $Q\colon H \to H$ be β-Lipschitzian.
- Classical methods for finding $\bar{x} \in \operatorname{zer} M$ are:
  - Proximal-point algo. if $M = W$.
  - Euler method if $M = C$.
  - Forward-backward algo. if $M = W + C$.
  - Forward-backward-forward algo. if $M = W + Q$.
  - Forward-backward-half-forward algo. if $M = W + C + Q$.
  - Among many others.

Splitting methods

As shown in (Combettes, Acta Numer., 2024), all the examples come from the same geometric framework: In the context of $M = W + C$, let $x_0 \in H$ and iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\text{take } (w_n, w_n^*) \in \operatorname{gra} W \text{ and } q_n \in H \\
t_n^* = w_n^* + Cq_n \\
\Delta_n = \langle x_n - w_n \mid t_n^* \rangle - (4\alpha)^{-1} \|w_n - q_n\|^2 \\
\theta_n = \begin{cases} \dfrac{\Delta_n}{\|t_n^*\|^2}, & \text{if } \Delta_n > 0; \\ 0, & \text{otherwise} \end{cases} \\
d_n = \theta_n t_n^* \\
\text{take } \lambda_n \in]0, 2[\\
x_{n+1} = x_n - \lambda_n d_n.
\end{array}
\right.
\end{array}
\tag{1}
$$

Splitting methods

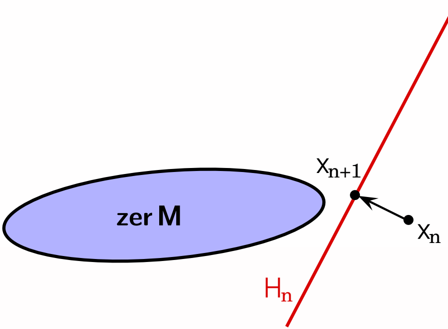

Figure: Geometry of algorithms for finding a point in zer M.

Suppose that

$$\begin{cases} x_n - w_n \rightharpoonup 0; \\ w_n - q_n \to 0; \\ t_n^* = w_n^* + Cq_n \to 0. \end{cases} \tag{2}$$

Then $x_n \rightharpoonup \bar{x} \in \operatorname{zer} M$.

Adding stochasticity

Stochasticity in splitting can be introduced in:

(i) stochastic approximation of operators,

(ii) random block-iterative implementations,

Our objective is to propose a stochastic modification of the deterministic algorithm.

In the context of $\mathsf{M} = \mathsf{W} + \mathsf{C}$, let $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and iterate

$$
\begin{array}{l}
\text{for } \mathsf{n} = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\text{take } \{w_{\mathsf{n}}, w_{\mathsf{n}}^*, e_{\mathsf{n}}, e_{\mathsf{n}}^*\} \subset L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \\
\qquad\qquad \text{such that } (w_{\mathsf{n}} + e_{\mathsf{n}}, w_{\mathsf{n}}^* + e_{\mathsf{n}}^*) \in \operatorname{gra} \mathsf{W} \ \mathsf{P}\text{-a.s.} \\
\text{take } \{q_{\mathsf{n}}, c_{\mathsf{n}}^*, f_{\mathsf{n}}^*\} \subset L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \\
\qquad\qquad \text{such that } c_{\mathsf{n}}^* + f_{\mathsf{n}}^* = \mathsf{C} q_{\mathsf{n}} \ \mathsf{P}\text{-a.s.} \\
t_{\mathsf{n}}^* = w_{\mathsf{n}}^* + c_{\mathsf{n}}^* \\
\Delta_{\mathsf{n}} = \langle x_{\mathsf{n}} - w_{\mathsf{n}} \mid t_{\mathsf{n}}^* \rangle - (4\alpha)^{-1} \|w_{\mathsf{n}} - q_{\mathsf{n}}\|^2 \\
\theta_{\mathsf{n}} = \dfrac{1_{[t_{\mathsf{n}}^* \neq 0]} 1_{[\Delta_{\mathsf{n}} > 0]} \Delta_{\mathsf{n}}}{\|t_{\mathsf{n}}^*\|^2 + 1_{[t_{\mathsf{n}}^* = 0]}} \\
d_{\mathsf{n}} = \theta_{\mathsf{n}} t_{\mathsf{n}}^* \\
\text{take } \lambda_{\mathsf{n}} \in \left]0, 2\right[\\
x_{\mathsf{n}+1} = x_{\mathsf{n}} - \lambda_{\mathsf{n}} d_{\mathsf{n}}.
\end{array}
\right.
\end{array}
\tag{3}
$$

Proposed method

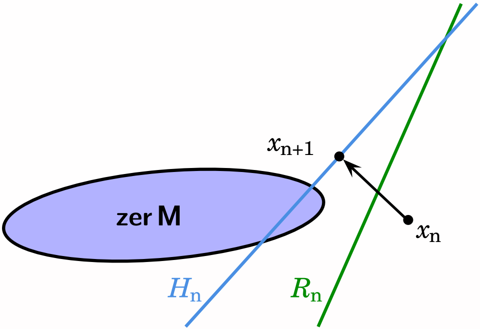

Figure: Geometry of proposed algorithm for finding a point in zer M. R_n denotes the true and unknown half-space.

Convergence

Theorem 1

For every $n \in \mathbb{N}$ *and every* $z \in \operatorname{zer} M$*, set* $\mathcal{X}_n = \sigma(x_0, \ldots, x_n)$ *and*

$$\varepsilon_n(\cdot, z) = \max\left\{0, \mathsf{E}\left(\theta_n\left(\langle w_n - z \mid e_n^* + f_n^*\rangle + \langle e_n \mid w_n^* + Cz\rangle + \langle e_n \mid e_n^*\rangle\right) \middle| \mathcal{X}_n\right)\right\}.$$

Suppose that, for every $z \in \operatorname{zer} M$, $\sum_{n\in\mathbb{N}} \lambda_n \mathsf{E}\varepsilon_n(\cdot, z) < +\infty$*. Then the following are satisfied:*

1. *Suppose that* $x_n - w_n - e_n \rightharpoonup 0$ P-a.s., $w_n + e_n - q_n \to 0$ P-a.s.*, and* $w_n^* + e_n^* + Cq_n \to 0$ P-a.s. *Then* $x_n \rightharpoonup \bar{x} \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \operatorname{zer} M)$*.*
2. *Suppose that* $\dim H < +\infty$, $x_n - w_n - e_n \xrightarrow{\mathsf{P}} 0$, $w_n + e_n - q_n \xrightarrow{\mathsf{P}} 0$*, and* $w_n^* + e_n^* + Cq_n \xrightarrow{\mathsf{P}} 0$*. Then* $x_n \to \bar{x} \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \operatorname{zer} M)$*.*

Application: Stochastic proximal-point algorithm

Stochastic proximal-point algorithm

Theorem 2

Let $\mathsf{W}\colon \mathsf{H} \to 2^{\mathsf{H}}$ *be a maximally monotone operator such that* $\operatorname{zer}\mathsf{W} \neq \varnothing$, *let* $(\gamma_n)_{n\in\mathbb{N}}$ *be a sequence in* $]0,+\infty[$, *and let* $x_0 \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$. *Iterate*

$$
\begin{array}{l}
\text{for } n = 0,1,\ldots \\
\left\lfloor
\begin{array}{l}
\text{take } e_n \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H}) \text{ and } \lambda_n \in\,]0,2[\\
x_{n+1} = x_n + \lambda_n\big(\mathsf{J}_{\gamma_n\mathsf{W}}x_n - e_n - x_n\big).
\end{array}
\right.
\end{array}
\tag{4}
$$

Suppose that one of the following holds:

1. $\sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty$, $\sum_{n\in\mathbb{N}} \lambda_n\sqrt{\mathsf{E}\|e_n\|^2} < +\infty$, $(\mathsf{E}\|e_n\|^2)_{n\in\mathbb{N}}$ *is bounded, and* $(\forall n\in\mathbb{N})\ \gamma_n = 1$.
2. $\inf_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) > 0$, $\inf_{n\in\mathbb{N}} \gamma_n > 0$, *and* $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}\|e_n\|^2} < +\infty$.
3. $\sum_{n\in\mathbb{N}} \gamma_n^2 = +\infty$, $\sum_{n\in\mathbb{N}} \sqrt{\mathsf{E}\|e_n\|^2} < +\infty$, *and* $(\forall n\in\mathbb{N})\ \lambda_n = 1$ P-a.s.

Then $x_n \rightharpoonup \bar{x} \in L^2(\Omega,\mathcal{F},\mathsf{P};\operatorname{zer}\mathsf{W})$.

Application:
Multivariate monotone inclusion

Multivariate monotone inclusion

I and K are finite. For every $\mathsf{i} \in I$ and every $\mathsf{k} \in K$,

$$
\begin{cases}
\mathsf{H}_\mathsf{i} \text{ is a Euclidean space,} \\
\mathsf{A}_\mathsf{i} \colon \mathsf{H}_\mathsf{i} \to 2^{\mathsf{H}_\mathsf{i}} \text{ is maximally monotone,} \\
\mathsf{C}_\mathsf{i} \colon \mathsf{H}_\mathsf{i} \to \mathsf{H}_\mathsf{i} \text{ is cocoercive,} \\
\mathsf{Q}_\mathsf{i} \colon \mathsf{H}_\mathsf{i} \to \mathsf{H}_\mathsf{i} \text{ is monotone and Lipschitzian,} \\
\\
\mathsf{G}_\mathsf{k} \text{ is a Euclidean space,} \\
\mathsf{B}_\mathsf{k}^m \colon \mathsf{G}_\mathsf{k} \to 2^{\mathsf{G}_\mathsf{k}} \text{ and } \mathsf{D}_\mathsf{k}^m \colon \mathsf{G}_\mathsf{k} \to 2^{\mathsf{G}_\mathsf{k}} \text{ are maximally monotone,} \\
\mathsf{B}_\mathsf{k}^c \colon \mathsf{G}_\mathsf{k} \to \mathsf{G}_\mathsf{k} \text{ and } \mathsf{D}_\mathsf{k}^c \colon \mathsf{G}_\mathsf{k} \to \mathsf{G}_\mathsf{k} \text{ are cocoercive,} \\
\mathsf{B}_\mathsf{k}^l \colon \mathsf{G}_\mathsf{k} \to \mathsf{G}_\mathsf{k} \text{ and } \mathsf{D}_\mathsf{k}^l \colon \mathsf{G}_\mathsf{k} \to \mathsf{G}_\mathsf{k} \text{ are monotone and Lipschitzian,} \\
\\
\mathsf{L}_{\mathsf{ki}} \colon \mathsf{H}_\mathsf{i} \to \mathsf{G}_\mathsf{k} \text{ is linear.}
\end{cases}
$$

We set $\mathbf{H} = \bigoplus_{i\in I} \mathsf{H}_i$ and $\mathbf{G} = \bigoplus_{k\in K} \mathsf{G}_k$.

$$(\forall i \in I) \quad R_i : \mathbf{H} \to \mathsf{H}_i, \tag{5}$$

and we assume that $\mathbf{R} \colon \mathbf{H} \to \mathbf{H} \colon \mathbf{x} \mapsto (R_i\mathbf{x})_{i\in I}$ is monotone and Lipschitzian.

Problem 1

The objective is to

$$\text{find } \overline{\mathbf{x}} \in \mathbf{H} \text{ such that } (\forall i \in I) \;\; 0 \in A_i\overline{x}_i + C_i\overline{x}_i + Q_i\overline{x}_i + R_i\overline{\mathbf{x}} \\ + \sum_{k\in K} L_{ki}^*\left(\left((B_k^m + B_k^c + B_k^l) \,\square\, (D_k^m + D_k^c + D_k^l)\right)\left(\sum_{j\in I} L_{kj}\overline{x}_j\right)\right),$$

where $M_1 \,\square\, M_2 = (M_1^{-1} + M_2^{-1})^{-1}$ denotes the parallel sum of M_1 and M_2.

Multivariate monotone inclusion

Modern large-scale problems require complex models.

Multivariate monotone inclusion

Modern large-scale problems require complex models.

To solve Problem 1, we consider the following saddle form.

The saddle form

We set $\underline{\mathbf{X}} = \mathbf{H} \oplus \mathbf{G} \oplus \mathbf{G} \oplus \mathbf{G}$. The *saddle operator* is

$$\underline{\mathcal{S}} \colon \underline{\mathbf{X}} \to 2^{\underline{\mathbf{X}}} \colon (\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{v}^*) \mapsto$$

$$\Bigg(\bigtimes_{i \in I} \Big(A_i x_i + C_i x_i + Q_i x_i + R_i \mathbf{x} + \sum_{k \in K} L_{ki}^* v_k^* \Big),$$

$$\bigtimes_{k \in K} \big(B_k^m y_k + B_k^c y_k + B_k^l y_k - v_k^* \big),$$

$$\bigtimes_{k \in K} \big(D_k^m z_k + D_k^c z_k + D_k^l z_k - v_k^* \big),$$

$$\bigtimes_{k \in K} \Big\{ y_k + z_k - \sum_{i \in I} L_{ki} x_i \Big\} \Bigg),$$

and the *saddle form* is to

$$\text{find } \underline{\overline{\mathbf{x}}} \in \underline{\mathbf{X}} \text{ such that } \underline{\mathbf{0}} \in \underline{\mathcal{S}} \underline{\overline{\mathbf{x}}}. \tag{6}$$

The saddle form

- It follows from (Bùi-Combettes, Math. Oper. Res., 2022) that

$$(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{v}^*) \in \operatorname{zer} \underline{\mathcal{S}} \Rightarrow \mathbf{x} \text{ solves Problem 1.}$$

The saddle form

- It follows from (Bùi-Combettes, Math. Oper. Res., 2022) that

$$(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{v}^*) \in \operatorname{zer} \underline{\mathcal{S}} \Rightarrow \mathbf{x} \text{ solves Problem 1.}$$

- **Strategy:** Solve the saddle form.
    - $\underline{\mathcal{S}} = \mathbf{W} + \mathbf{C}$.
    - We use the framework without errors:

Proposed method II

In the context of M = W + C, let $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ and iterate

$$
\begin{array}{l}
\text{for } \mathsf{n} = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\text{take } \{w_{\mathsf{n}}, w_{\mathsf{n}}^*\} \subset L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \\
\qquad\qquad \text{such that } (w_{\mathsf{n}}, w_{\mathsf{n}}^*) \in \operatorname{gra} \mathsf{W} \ \mathsf{P}\text{-a.s.} \\
\text{take } q_{\mathsf{n}} \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \\
t_{\mathsf{n}}^* = w_{\mathsf{n}}^* + \mathsf{C}q_{\mathsf{n}}^* \\
\Delta_{\mathsf{n}} = \langle x_{\mathsf{n}} - w_{\mathsf{n}} \mid t_{\mathsf{n}}^* \rangle - (4\alpha)^{-1} \|w_{\mathsf{n}} - q_{\mathsf{n}}\|^2 \\
\theta_{\mathsf{n}} = \dfrac{1_{[t_{\mathsf{n}}^* \neq 0]} 1_{[\Delta_{\mathsf{n}} > 0]} \Delta_{\mathsf{n}}}{\|t_{\mathsf{n}}^*\|^2 + 1_{[t_{\mathsf{n}}^* = 0]}} \\
d_{\mathsf{n}} = \theta_{\mathsf{n}} t_{\mathsf{n}}^* \\
\text{take } \lambda_{\mathsf{n}} \in \left]0, 2\right[\\
x_{\mathsf{n}+1} = x_{\mathsf{n}} - \lambda_{\mathsf{n}} d_{\mathsf{n}}.
\end{array}
\right.
\end{array}
\tag{7}
$$

Algorithm

Solving the saddle form:

Let,

- for every $i \in I$, $x_{i,0} \in L^2(\Omega, \mathcal{F}, P; H_i)$.
- for every $k \in K$, $\{y_{k,0}, z_{k,0}, v^*_{k,0}\} \subset L^2(\Omega, \mathcal{F}, P; G_k)$.

Iterate as follows

Algorithm

$$
\begin{array}{l}
\text{for } \mathsf{n} = 0,1,\ldots\\
\quad \text{for every } \mathsf{i} \in I_{\mathsf{n}}\\
\quad\quad l^*_{\mathsf{i,n}} = \mathsf{Q_i} x_{\mathsf{i,n}} + \mathsf{R_i}\boldsymbol{x}_{\mathsf{n}} + \sum_{\mathsf{k}\in K} \mathsf{L}^*_{\mathsf{ki}} v^*_{\mathsf{k,n}};\ a_{\mathsf{i,n}} = \mathsf{J}_{\gamma_{\mathsf{i,n}}\mathsf{A_i}}\big(x_{\mathsf{i,n}} + \gamma_{\mathsf{i,n}}(\mathsf{s}^*_{\mathsf{i}} - l^*_{\mathsf{i,n}} - \mathsf{C_i} x_{\mathsf{i,n}})\big);\\
\quad\quad a^*_{\mathsf{i,n}} = \gamma^{-1}_{\mathsf{i,n}}(x_{\mathsf{i,n}} - a_{\mathsf{i,n}}) - l^*_{\mathsf{i,n}} + \mathsf{Q_i} a_{\mathsf{i,n}};\ \xi_{\mathsf{i,n}} = \|a_{\mathsf{i,n}} - x_{\mathsf{i,n}}\|^2;\\
\quad \text{for every } \mathsf{i} \in I \smallsetminus I_{\mathsf{n}}\\
\quad\quad a_{\mathsf{i,n}} = a_{\mathsf{i,n-1}};\ a^*_{\mathsf{i,n}} = a^*_{\mathsf{i,n-1}};\ \xi_{\mathsf{i,n}} = \xi_{\mathsf{i,n-1}};\\
\quad \text{for every } \mathsf{k} \in K_{\mathsf{n}}\\
\quad\quad u^*_{\mathsf{k,n}} = v^*_{\mathsf{k,n}} - \mathsf{B}^l_{\mathsf{k}} y_{\mathsf{k,n}};\ w^*_{\mathsf{k,n}} = v^*_{\mathsf{k,n}} - \mathsf{D}^l_{\mathsf{k}} z_{\mathsf{k,n}};\ b_{\mathsf{k,n}} = \mathsf{J}_{\mu_{\mathsf{k,n}}\mathsf{B}^m_{\mathsf{k}}}\big(y_{\mathsf{k,n}} + \mu_{\mathsf{k,n}}(u^*_{\mathsf{k,n}} - \mathsf{B}^c_{\mathsf{k}} y_{\mathsf{k,n}})\big);\\
\quad\quad d_{\mathsf{k,n}} = \mathsf{J}_{\nu_{\mathsf{k,n}}\mathsf{D}^m_{\mathsf{k}}}\big(z_{\mathsf{k,n}} + \nu_{\mathsf{k,n}}(w^*_{\mathsf{k,n}} - \mathsf{D}^c_{\mathsf{k}} z_{\mathsf{k,n}})\big);\ e^*_{\mathsf{k,n}} = \sigma_{\mathsf{k,n}}\big(\sum_{\mathsf{i}\in I} \mathsf{L_{ki}} x_{\mathsf{i,n}} - y_{\mathsf{k,n}} - z_{\mathsf{k,n}} - \mathsf{r_k}\big) + v^*_{\mathsf{k,n}};\\
\quad\quad q^*_{\mathsf{k,n}} = \mu^{-1}_{\mathsf{k,n}}(y_{\mathsf{k,n}} - b_{\mathsf{k,n}}) + u^*_{\mathsf{k,n}} + \mathsf{B}^l_{\mathsf{k}} b_{\mathsf{k,n}} - e^*_{\mathsf{k,n}};\ t^*_{\mathsf{k,n}} = \nu^{-1}_{\mathsf{k,n}}(z_{\mathsf{k,n}} - d_{\mathsf{k,n}}) + w^*_{\mathsf{k,n}} + \mathsf{D}^l_{\mathsf{k}} d_{\mathsf{k,n}} - e^*_{\mathsf{k,n}};\\
\quad\quad \eta_{\mathsf{k,n}} = \|b_{\mathsf{k,n}} - y_{\mathsf{k,n}}\|^2 + \|d_{\mathsf{k,n}} - z_{\mathsf{k,n}}\|^2;\ e_{\mathsf{k,n}} = \mathsf{r_k} + b_{\mathsf{k,n}} + d_{\mathsf{k,n}} - \sum_{\mathsf{i}\in I} \mathsf{L_{ki}} a_{\mathsf{i,n}};\\
\quad \text{for every } \mathsf{k} \in K \smallsetminus K_{\mathsf{n}}\\
\quad\quad b_{\mathsf{k,n}} = b_{\mathsf{k,n-1}};\ d_{\mathsf{k,n}} = d_{\mathsf{k,n-1}};\ e^*_{\mathsf{k,n}} = e^*_{\mathsf{k,n-1}};\ q^*_{\mathsf{k,n}} = q^*_{\mathsf{k,n-1}};\ t^*_{\mathsf{k,n}} = t^*_{\mathsf{k,n-1}};\\
\quad\quad \eta_{\mathsf{k,n}} = \eta_{\mathsf{k,n-1}};\ e_{\mathsf{k,n}} = \mathsf{r_k} + b_{\mathsf{k,n}} + d_{\mathsf{k,n}} - \sum_{\mathsf{i}\in I} \mathsf{L_{ki}} a_{\mathsf{i,n}};\\
\quad \text{for every } \mathsf{i} \in I\\
\quad\quad p^*_{\mathsf{i,n}} = a^*_{\mathsf{i,n}} + \mathsf{R_i}\boldsymbol{a}_{\mathsf{n}} + \sum_{\mathsf{k}\in K} \mathsf{L}^*_{\mathsf{ki}} e^*_{\mathsf{k,n}};\\
\quad \Delta_{\mathsf{n}} = -(4\alpha)^{-1}\big(\sum_{\mathsf{i}\in I} \xi_{\mathsf{i,n}} + \sum_{\mathsf{k}\in K} \eta_{\mathsf{k,n}}\big) + \sum_{\mathsf{i}\in I} \langle x_{\mathsf{i,n}} - a_{\mathsf{i,n}} \mid p^*_{\mathsf{i,n}}\rangle\\
\quad\quad\quad + \sum_{\mathsf{k}\in K} \big(\langle y_{\mathsf{k,n}} - b_{\mathsf{k,n}} \mid q^*_{\mathsf{k,n}}\rangle + \langle z_{\mathsf{k,n}} - d_{\mathsf{k,n}} \mid t^*_{\mathsf{k,n}}\rangle + \langle e_{\mathsf{k,n}} \mid v^*_{\mathsf{k,n}} - e^*_{\mathsf{k,n}}\rangle\big);\\
\quad \theta_{\mathsf{n}} = 1_{[\Delta_{\mathsf{n}}>0]}\Delta_{\mathsf{n}} \Big/ \Big(\sum_{\mathsf{i}\in I} \|p^*_{\mathsf{i,n}}\|^2 + \sum_{\mathsf{k}\in K} \big(\|q^*_{\mathsf{k,n}}\|^2 + \|t^*_{\mathsf{k,n}}\|^2 + \|e_{\mathsf{k,n}}\|^2\big) + 1_{[\Delta_{\mathsf{n}}\leqslant 0]}\Big);\\
\quad \lambda_{\mathsf{n}} \in \left]0,2\right[\\
\quad \text{for every } \mathsf{i} \in I\\
\quad\quad x_{\mathsf{i,n+1}} = x_{\mathsf{i,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} p^*_{\mathsf{i,n}};\\
\quad \text{for every } \mathsf{k} \in K\\
\quad\quad y_{\mathsf{k,n+1}} = y_{\mathsf{k,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} q^*_{\mathsf{k,n}};\ z_{\mathsf{k,n+1}} = z_{\mathsf{k,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} t^*_{\mathsf{k,n}};\ v^*_{\mathsf{k,n+1}} = v^*_{\mathsf{k,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} e_{\mathsf{k,n}};
\end{array}
$$

Algorithm

for $\mathsf{n} = 0, 1, \ldots$

for every $\mathsf{i} \in I_{\mathsf{n}}$

$l^*_{\mathsf{i,n}} = \mathsf{Q}_{\mathsf{i}} x_{\mathsf{i,n}} + \mathsf{R}_{\mathsf{i}} \boldsymbol{x}_{\mathsf{n}} + \sum_{\mathsf{k}\in K} \mathsf{L}^*_{\mathsf{ki}} v^*_{\mathsf{k,n}}$; $a_{\mathsf{i,n}} = \mathsf{J}_{\gamma_{\mathsf{i,n}}\mathsf{A}_{\mathsf{i}}}\big(x_{\mathsf{i,n}} + \gamma_{\mathsf{i,n}}(\mathsf{s}^*_{\mathsf{i}} - l^*_{\mathsf{i,n}} - \mathsf{C}_{\mathsf{i}} x_{\mathsf{i,n}})\big)$;

$a^*_{\mathsf{i,n}} = \gamma^{-1}_{\mathsf{i,n}}(x_{\mathsf{i,n}} - a_{\mathsf{i,n}}) - l^*_{\mathsf{i,n}} + \mathsf{Q}_{\mathsf{i}} a_{\mathsf{i,n}}$; $\xi_{\mathsf{i,n}} = \|a_{\mathsf{i,n}} - x_{\mathsf{i,n}}\|^2$;

for every $\mathsf{i} \in I \smallsetminus I_{\mathsf{n}}$

$a_{\mathsf{i,n}} = a_{\mathsf{i,n-1}}$; $a^*_{\mathsf{i,n}} = a^*_{\mathsf{i,n-1}}$; $\xi_{\mathsf{i,n}} = \xi_{\mathsf{i,n-1}}$;

for every $\mathsf{k} \in K_{\mathsf{n}}$

$u^*_{\mathsf{k,n}} = v^*_{\mathsf{k,n}} - \mathsf{B}^l_{\mathsf{k}} y_{\mathsf{k,n}}$; $w^*_{\mathsf{k,n}} = v^*_{\mathsf{k,n}} - \mathsf{D}^l_{\mathsf{k}} z_{\mathsf{k,n}}$; $b_{\mathsf{k,n}} = \mathsf{J}_{\mu_{\mathsf{k,n}}\mathsf{B}^m_{\mathsf{k}}}\big(y_{\mathsf{k,n}} + \mu_{\mathsf{k,n}}(u^*_{\mathsf{k,n}} - \mathsf{B}^c_{\mathsf{k}} y_{\mathsf{k,n}})\big)$;

$d_{\mathsf{k,n}} = \mathsf{J}_{\nu_{\mathsf{k,n}}\mathsf{D}^m_{\mathsf{k}}}\big(z_{\mathsf{k,n}} + \nu_{\mathsf{k,n}}(w^*_{\mathsf{k,n}} - \mathsf{D}^c_{\mathsf{k}} z_{\mathsf{k,n}})\big)$; $e^*_{\mathsf{k,n}} = \sigma_{\mathsf{k,n}}\big(\sum_{\mathsf{i}\in I} \mathsf{L}_{\mathsf{ki}} x_{\mathsf{i,n}} - y_{\mathsf{k,n}} - z_{\mathsf{k,n}} - \mathsf{r}_{\mathsf{k}}\big) + v^*_{\mathsf{k,n}}$;

$q^*_{\mathsf{k,n}} = \mu^{-1}_{\mathsf{k,n}}(y_{\mathsf{k,n}} - b_{\mathsf{k,n}}) + u^*_{\mathsf{k,n}} + \mathsf{B}^l_{\mathsf{k}} b_{\mathsf{k,n}} - e^*_{\mathsf{k,n}}$; $t^*_{\mathsf{k,n}} = \nu^{-1}_{\mathsf{k,n}}(z_{\mathsf{k,n}} - d_{\mathsf{k,n}}) + w^*_{\mathsf{k,n}} + \mathsf{D}^l_{\mathsf{k}} d_{\mathsf{k,n}} - e^*_{\mathsf{k,n}}$;

$\eta_{\mathsf{k,n}} = \|b_{\mathsf{k,n}} - y_{\mathsf{k,n}}\|^2 + \|d_{\mathsf{k,n}} - z_{\mathsf{k,n}}\|^2$; $e_{\mathsf{k,n}} = \mathsf{r}_{\mathsf{k}} + b_{\mathsf{k,n}} + d_{\mathsf{k,n}} - \sum_{\mathsf{i}\in I} \mathsf{L}_{\mathsf{ki}} a_{\mathsf{i,n}}$;

for every $\mathsf{k} \in K \smallsetminus K_{\mathsf{n}}$

$b_{\mathsf{k,n}} = b_{\mathsf{k,n-1}}$; $d_{\mathsf{k,n}} = d_{\mathsf{k,n-1}}$; $e^*_{\mathsf{k,n}} = e^*_{\mathsf{k,n-1}}$; $q^*_{\mathsf{k,n}} = q^*_{\mathsf{k,n-1}}$; $t^*_{\mathsf{k,n}} = t^*_{\mathsf{k,n-1}}$;

$\eta_{\mathsf{k,n}} = \eta_{\mathsf{k,n-1}}$; $e_{\mathsf{k,n}} = \mathsf{r}_{\mathsf{k}} + b_{\mathsf{k,n}} + d_{\mathsf{k,n}} - \sum_{\mathsf{i}\in I} \mathsf{L}_{\mathsf{ki}} a_{\mathsf{i,n}}$;

for every $\mathsf{i} \in I$

$p^*_{\mathsf{i,n}} = a^*_{\mathsf{i,n}} + \mathsf{R}_{\mathsf{i}} \boldsymbol{a}_{\mathsf{n}} + \sum_{\mathsf{k}\in K} \mathsf{L}^*_{\mathsf{ki}} e^*_{\mathsf{k,n}}$;

$$\Delta_{\mathsf{n}} = -(4\alpha)^{-1}\big(\textstyle\sum_{\mathsf{i}\in I} \xi_{\mathsf{i,n}} + \sum_{\mathsf{k}\in K} \eta_{\mathsf{k,n}}\big) + \sum_{\mathsf{i}\in I} \langle x_{\mathsf{i,n}} - a_{\mathsf{i,n}} \mid p^*_{\mathsf{i,n}}\rangle$$
$$+ \textstyle\sum_{\mathsf{k}\in K} \big(\langle y_{\mathsf{k,n}} - b_{\mathsf{k,n}} \mid q^*_{\mathsf{k,n}}\rangle + \langle z_{\mathsf{k,n}} - d_{\mathsf{k,n}} \mid t^*_{\mathsf{k,n}}\rangle + \langle e_{\mathsf{k,n}} \mid v^*_{\mathsf{k,n}} - e^*_{\mathsf{k,n}}\rangle\big);$$

$$\theta_{\mathsf{n}} = 1_{[\Delta_{\mathsf{n}}>0]}\Delta_{\mathsf{n}} \Big/ \Big(\textstyle\sum_{\mathsf{i}\in I} \|p^*_{\mathsf{i,n}}\|^2 + \sum_{\mathsf{k}\in K}\big(\|q^*_{\mathsf{k,n}}\|^2 + \|t^*_{\mathsf{k,n}}\|^2 + \|e_{\mathsf{k,n}}\|^2\big) + 1_{[\Delta_{\mathsf{n}}\leqslant 0]}\Big);$$

$\lambda_{\mathsf{n}} \in \,]0, 2[$

for every $\mathsf{i} \in I$

$x_{\mathsf{i,n+1}} = x_{\mathsf{i,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} p^*_{\mathsf{i,n}}$;

for every $\mathsf{k} \in K$

$y_{\mathsf{k,n+1}} = y_{\mathsf{k,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} q^*_{\mathsf{k,n}}$; $z_{\mathsf{k,n+1}} = z_{\mathsf{k,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} t^*_{\mathsf{k,n}}$; $v^*_{\mathsf{k,n+1}} = v^*_{\mathsf{k,n}} - \lambda_{\mathsf{n}}\theta_{\mathsf{n}} e_{\mathsf{k,n}}$;

Control rule

- **Assumption:** There exists $N \in \mathbb{N} \setminus \{0\}$ such that $\mathit{I}_0 = \mathit{I}, \mathit{K}_0 = \mathit{K}$, and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} (\forall i \in \mathit{I}) \;\; \mathsf{P}\left(\left[i \in \bigcup_{j=n}^{n+N-1} \mathit{I}_j \right]\right) \geqslant \pi_i > 0 \\ (\forall k \in \mathit{K}) \;\; \mathsf{P}\left(\left[k \in \bigcup_{j=n}^{n+N-1} \mathit{K}_j \right]\right) \geqslant \zeta_k > 0. \end{cases} \tag{8}$$

- Particularly true if $\mathit{I}_n = \{i_n\}$ and $\mathit{K}_n = \{k_n\}$ with $i_n = \mathrm{uniform}(\mathit{I})$ and $k_n = \mathrm{uniform}(\mathit{K})$.

Denote by $\mathscr{P}$ the set of solutions to Problem 1. Then there exists a $\mathscr{P}$-valued random variable $\overline{\boldsymbol{x}}$ such that

$$(\forall \mathrm{i} \in I) \quad \begin{cases} x_{\mathrm{i},\mathrm{n}} \to \overline{x}_{\mathrm{i}} \ \ \text{P-a.s.} \\ x_{\mathrm{i},\mathrm{n}} \to \overline{x}_{\mathrm{i}} \ \ \text{in } \ L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R}^{\mathsf{N}}). \end{cases} \tag{9}$$

Example on minimization

Product space stochastic gradient algorithm

Let $\alpha \in]0,+\infty[$ and, for every $k \in \{1,\ldots,p\}$, let $g_k \colon \mathbb{R}^N \to \mathbb{R}$ be convex, differentiable, and such that ∇g_k is $1/\alpha$-Lipschitzian. The task is to

$$\underset{x\in\mathbb{R}^N}{\text{minimize}} \quad \frac{1}{p}\sum_{k=1}^{p} g_k(x). \tag{10}$$

- If p is large, we may want to use a stochastic method to split the sum.
- **First idea:** Stochastic gradient method with uniform sampling from $\{1,\ldots,p\}$.
    - It requires control over the variance.
    - Only convergence of the function values.

Product space stochastic gradient algorithm

Let $\alpha \in]0, +\infty[$ and, for every $k \in \{1, \ldots, p\}$, let $g_k \colon \mathbb{R}^N \to \mathbb{R}$ be convex, differentiable, and such that ∇g_k is $1/\alpha$-Lipschitzian. The task is to

$$\underset{x \in \mathbb{R}^N}{\text{minimize}} \; \frac{1}{p} \sum_{k=1}^{p} g_k(x). \tag{10}$$

- If p is large, we may want to use a stochastic method to split the sum.
- **First idea:** Stochastic gradient method with uniform sampling from $\{1, \ldots, p\}$.
  - It requires control over the variance.
  - Only convergence of the function values.
- **Second idea:** The saddle form.
  - Convergence to a solution without extra assumptions.

Product space stochastic gradient algorithm

Let x_0, $(y_{k,0})_{k\in K}$, $(z_{k,0})_{k\in K}$, and $(v^*_{k,0})_{k\in K}$ be in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathbb{R}^N)$. Iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
l^*_n = \sum_{k\in K} v^*_{k,n};\\
\xi_n = \|\gamma_n l^*_n\|^2;\\
\text{if } k = k_n\\
\left\lfloor
\begin{array}{l}
b_{k,n} = y_{k,n} + \mu_{k,n}\big(v^*_{k,n} - \nabla g_k(y_{k,n})\big);\\
q^*_{k,n} = \mu^{-1}_{k,n}(y_{k,n} - b_{k,n}) - \sigma_{k,n}(x_n - y_{k,n} - z_{k,n});\ t^*_{k,n} = -v^*_{k,n} - \sigma_{k,n}(x_n - y_{k,n} - z_{k,n});\\
\eta_{k,n} = \|b_{k,n} - y_{k,n}\|^2 + \|\nu_{k,n} v^*_{k,n}\|^2;\ e_{k,n} = b_{k,n} + z_{k,n} + \nu_{k,n} v^*_{k,n} + \gamma_n l^*_n - x_n;
\end{array}
\right.\\
\text{else}\\
\left\lfloor
\begin{array}{l}
b_{k,n} = b_{k,n-1};\ q^*_{k,n} = q^*_{k,n-1};\ t^*_{k,n} = t^*_{k,n-1};\ \eta_{k,n} = \eta_{k,n-1};\ e_{k,n} = b_{k,n} + z_{k,n} + \nu_{k,n} v^*_{k,n} + \gamma_n l^*_n - x_n;
\end{array}
\right.\\
p^*_n = \sum_{k\in K}\big(\sigma_{k,n}(x_n - y_{k,n} - z_{k,n}) + v^*_{k,n}\big);\\
\Delta_n = -(4\alpha)^{-1}\big(\xi_n + \sum_{k\in K}\eta_{k,n}\big) + \langle \gamma_n l^*_n \mid p^*_n\rangle\\
\qquad + \sum_{k\in K}\Big(\langle y_{k,n} - b_{k,n} \mid q^*_{k,n}\rangle - \langle \nu_{k,n} v^*_{k,n} \mid t^*_{k,n}\rangle - \big\langle e_{k,n} \mid \sigma_{k,n}(x_n - y_{k,n} - z_{k,n})\big\rangle\Big);\\
\theta_n = 1_{[\Delta_n>0]}\Delta_n\Big/\Big(\sum_{i\in I}\|p^*_{i,n}\|^2 + \sum_{k\in K}\big(\|q^*_{k,n}\|^2 + \|t^*_{k,n}\|^2 + \|e_{k,n}\|^2\big) + 1_{[\Delta_n\leqslant 0]}\Big);\\
\lambda_n \in L^\infty(\Omega,\mathcal{F},\mathsf{P};\,]0,+\infty[)\\
x_{n+1} = x_n - \lambda_n\theta_n p^*_n;\\
\text{for every } k \in K\\
\left\lfloor
\begin{array}{l}
y_{k,n+1} = y_{k,n} - \lambda_n\theta_n q^*_{k,n};\ z_{k,n+1} = z_{k,n} - \lambda_n\theta_n t^*_{k,n};\ v^*_{k,n+1} = v^*_{k,n} - \lambda_n\theta_n e_{k,n};
\end{array}
\right.
\end{array}
\right.
\end{array}
$$

Then $x_n \to \bar{x} \in L^1(\Omega,\mathcal{F},\mathsf{P};\mathscr{P})$.

References

P. L. Combettes and J. I. Madariaga,
A geometric framework for stochastic iterations,
preprint, 2025.

P. L. Combettes and J. I. Madariaga,
Asymptotic analysis of stochastic splitting methods for multivariate monotone inclusions,
preprint, 2025.

jimadari@ncsu.edu