Stochastic Block-Iterative Projections Method for Convex Feasibility

Javier I. Madariaga

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

NLA Student Seminar
March 24, 2025

NC STATE UNIVERSITY

Background

- Throughout, $\mathcal{H}$ is a separable real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and norm $\|\cdot\|$.

Background

- Throughout, H is a separable real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and norm $\|\cdot\|$.
- $(\Omega, \mathcal{F}, \mathsf{P})$ is the probability space that defines all the random variables.

Background

- Throughout, H is a separable real Hilbert space with scalar product $\langle\cdot \mid \cdot\rangle$ and norm $\|\cdot\|$.
- $(\Omega, \mathcal{F}, \mathsf{P})$ is the probability space that defines all the random variables.
- We use sans-serif letters for deterministic variables and italicized serif letters for random variables.

General problem

Problem 1

Let Z be a nonempty closed convex subset of H. The task is to

$$\text{find } \bar{x} \in H \text{ such that } \bar{x} \in Z. \tag{1}$$

General problem

Problem 1

Let Z be a nonempty closed convex subset of H. The task is to

$$\text{find } \bar{x} \in H \ \text{ such that } \ \bar{x} \in Z. \tag{1}$$

Problem 1 covers many problems in analysis and optimization.

Example

Let $f\colon \mathcal{H} \to]-\infty, +\infty]$ be a proper, lower semicontinuous convex function.

Let C be a nonempty closed and convex subset of $\mathcal{H}$. The minimization problem

$$\operatorname*{minimize}_{x \in C} \; f(x).$$

is an example of Problem 1.

How to solve Problem 1 using iterative methods. Take $x_0 \in H$.

Results

How to solve Problem 1 using iterative methods. Take $x_0 \in H$.

0.- Compute $\bar{x} = \mathrm{proj}_Z\, x_0$.

Results

How to solve Problem 1 using iterative methods. Take $x_0 \in H$.

0.- Compute $\bar{x} = \mathrm{proj}_Z x_0$.

- Convergence in one iteration! ($\bar{x} \in Z$).

Results

How to solve Problem 1 using iterative methods. Take $x_0 \in H$.

0.- Compute $\bar{x} = \mathrm{proj}_Z\, x_0$.

- Convergence in one iteration! ($\bar{x} \in Z$).
- It is not realist to know proj_Z.

Results

How to solve Problem 1 using iterative methods. Take $x_0 \in H$.

0.- Compute $\bar{x} = \mathrm{proj}_Z\, x_0$.

- Convergence in one iteration! ($\bar{x} \in Z$).
- It is not realist to know proj_Z.

We will assume that $Z = \bigcap_{1 \leqslant k \leqslant p} C_k$ and each proj_{C_k} is easy to compute.

Results

1.- Successive projections **(Von Neumann 1933)**.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \mathrm{proj}_{C_{n \bmod p}} x_n. \tag{1}$$

Results

1.- Successive projections **(Von Neumann 1933)**.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \mathrm{proj}_{C_{n \bmod p}} x_n. \tag{1}$$

2.- Barycentric method **(Auslender 1969)**.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \frac{1}{p} \sum_{k=1}^{p} \mathrm{proj}_{C_k} x_n. \tag{2}$$

Results

1.- Successive projections **(Von Neumann 1933)**.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \mathrm{proj}_{C_{n \bmod p}} x_n. \tag{1}$$

2.- Barycentric method **(Auslender 1969)**.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \frac{1}{p} \sum_{k=1}^{p} \mathrm{proj}_{C_k} x_n. \tag{2}$$

Weak convergence to a point in Z is guaranteed.

Results

3.- Parallel projections (**Combettes 1997**).

- Set $K_n \subset \mathbb{N}$ and weights $\beta_{k,n} \in [0,1]$ with $\sum_{k \in K_n} \beta_{k,n} = 1$.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \sum_{k \in K_n} \beta_{k,n} \operatorname{proj}_{C_k} x_n. \tag{3}$$

3.- Parallel projections (**Combettes 1997**).

- Set $K_n \subset \mathbb{N}$ and weights $\beta_{k,n} \in [0,1]$ with $\sum_{k \in K_n} \beta_{k,n} = 1$.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \sum_{k \in K_n} \beta_{k,n} \operatorname{proj}_{C_k} x_n. \tag{3}$$

- Add relaxation parameters $\lambda_n \in [\varepsilon, 2-\varepsilon]$ $(\star)$.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n \left(\sum_{k \in K_n} \beta_{k,n} \operatorname{proj}_{C_k} x_n - x_n \right). \tag{4}$$

Results

3.- Parallel projections (**Combettes 1997**).

- Add extrapolation.

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n L_n \left(\sum_{k \in K_n} \beta_{k,n} \operatorname{proj}_{C_k} x_n - x_n \right), \tag{5}$$

where

$$L_n = \begin{cases} \dfrac{\sum\limits_{k \in K_n} \beta_{k,n} \left\| \operatorname{proj}_{C_k} x_n - x_n \right\|^2}{\left\| \sum\limits_{k \in K_n} \beta_{k,n} \operatorname{proj}_{C_k} x_n - x_n \right\|^2}, & \text{if } x_n \notin \bigcap_{k \in K_n} C_k; \\ 1, & \text{otherwise.} \end{cases} \tag{6}$$

Results

3.- Parallel projections (**Combettes 1997)**.

- Weak convergence guaranteed under some assumptions over K_n, e.g.,

$$(\forall k \in \mathbb{N})(\exists M_k \in \mathbb{N} \setminus \{0\})(\forall n \in \mathbb{N}) \quad k \in \bigcup_{j=n}^{n+M_k-1} K_j. \tag{7}$$

Our setting

Problem

Let $(\mathsf{K}, \mathcal{K})$ *be a measurable space and* $(\mathsf{C}_\mathsf{k})_{\mathsf{k}\in\mathsf{K}}$ *a family of closed and convex subsets of* H. *Let* k *be a* K*-valued random variable.*

Our setting

Problem

Let $(\mathsf{K}, \mathcal{K})$ *be a measurable space and* $(\mathsf{C}_\mathsf{k})_{\mathsf{k}\in\mathsf{K}}$ *a family of closed and convex subsets of* H. *Let* k *be a* K*-valued random variable. The task is to*

$$\textit{find} \;\; \mathsf{x} \in \mathsf{Z} = \big\{\mathsf{z} \in \mathsf{H} \mid \mathsf{z} \in \mathsf{C}_{k} \;\; \mathsf{P}\text{-a.s.}\big\}, \tag{8}$$

under the assumption that $\mathsf{Z} \neq \varnothing$.

Our setting

We propose a stochastic extension of the parallel projections algorithm with stochasticity at four levels:

- The sets are indexed on a general measurable space rather than a countable set.

Our setting

We propose a stochastic extension of the parallel projections algorithm with stochasticity at four levels:

- The sets are indexed on a general measurable space rather than a countable set.
- The block of activated operators are randomly selected at each iteration.

Our setting

We propose a stochastic extension of the parallel projections algorithm with stochasticity at four levels:

- The sets are indexed on a general measurable space rather than a countable set.
- The block of activated operators are randomly selected at each iteration.
- The evaluations of the projections at iteration n are averaged and extrapolated with random weights $(\beta_{i,n})_{1\leqslant i\leqslant M}$.

Our setting

We propose a stochastic extension of the parallel projections algorithm with stochasticity at four levels:

- The sets are indexed on a general measurable space rather than a countable set.
- The block of activated operators are randomly selected at each iteration.
- The evaluations of the projections at iteration n are averaged and extrapolated with random weights $(\beta_{i,n})_{1\leqslant i\leqslant M}$.
- The relaxation parameter λ_n at iteration n is random and **not confined to the interval** $]0,2[$**.**

Our setting

Theorem (1)

In the setting of Problem 1, let $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$*,* $0 < \mathsf{M} \in \mathbb{N}$*,* $\delta \in \left]0, 1/\mathsf{M}\right[$*, and* $\rho \in \left[2, +\infty\right[$*. Iterate*

$$
\begin{array}{l}
\textit{for } \mathsf{n} = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\mathcal{X}_{\mathsf{n}} = \sigma(x_0, \ldots, x_{\mathsf{n}}) \\
\textit{for } \mathsf{i} = 1, \ldots, \mathsf{M} : k_{\mathsf{i},\mathsf{n}} \textit{ is distributed as } k \textit{ and independent of } \mathcal{X}_{\mathsf{n}} \\
(\beta_{\mathsf{i},\mathsf{n}})_{1 \leqslant \mathsf{i} \leqslant \mathsf{M}} \textit{ are } [\delta, 1] \textit{-valued r. v. such that } \sum_{\mathsf{i}=1}^{\mathsf{M}} \beta_{\mathsf{i},\mathsf{n}} = 1 \ \mathsf{P}\text{-a.s.} \\
p_{\mathsf{n}} = \sum_{\mathsf{i}=1}^{\mathsf{M}} \beta_{\mathsf{i},\mathsf{n}} \operatorname{proj}_{\mathsf{C}_{k_{\mathsf{i},\mathsf{n}}}} x_{\mathsf{n}} \\
L_{\mathsf{n}} = \dfrac{\sum_{\mathsf{i}=1}^{\mathsf{M}} \beta_{\mathsf{i},\mathsf{n}} \|\operatorname{proj}_{\mathsf{C}_{k_{\mathsf{i},\mathsf{n}}}} x_{\mathsf{n}} - x_{\mathsf{n}}\|^2 + 1_{[p_{\mathsf{n}} = x_{\mathsf{n}}]}}{\|p_{\mathsf{n}} - x_{\mathsf{n}}\|^2 + 1_{[p_{\mathsf{n}} = x_{\mathsf{n}}]}} \\
\lambda_{\mathsf{n}} \in L^\infty(\Omega, \mathcal{F}, \mathsf{P}; \left]0, \rho\right]) \\
x_{\mathsf{n}+1} = x_{\mathsf{n}} + \lambda_{\mathsf{n}} L_{\mathsf{n}} (p_{\mathsf{n}} - x_{\mathsf{n}}).
\end{array}
\right.
\end{array}
\tag{9}
$$

Our setting

Theorem (1)

Suppose that, there exists $\mu \in]0,1[$ *such that*

$$\inf_{n \in \mathbb{N}} \mathsf{E}(\lambda_n(2-\lambda_n)) \geqslant \mu. \tag{10}$$

Then $(x_n)_{n\in\mathbb{N}}$ *converges weakly in* $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ *and* P-a.s. *to some* $x \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$.

Super relaxations

From Theorem 1 we see that

$$\mathsf{E}(\lambda_n(2-\lambda_n)) \geqslant \mu \implies \mathsf{E}\lambda_n \in [\varepsilon, 2-\varepsilon]. \tag{11}$$

Super relaxations

From Theorem 1 we see that

$$\mathsf{E}(\lambda_n(2-\lambda_n)) \geqslant \mu \implies \mathsf{E}\lambda_n \in [\varepsilon, 2-\varepsilon]. \tag{11}$$

When $(\lambda_n)_{n\in\mathbb{N}}$ is deterministic we obtain,

$$\lambda_n \in [\varepsilon, 2-\varepsilon], \tag{12}$$

and it recover the standard range found in the literature.

Super relaxations

Example

- Let $\chi \in]0,1[$, $\alpha \in]0,+\infty[$, and $\beta \in]0,+\infty[$.

Super relaxations

Example

- Let $\chi \in \left]0,1\right[$, $\alpha \in \left]0,+\infty\right[$, and $\beta \in \left]0,+\infty\right[$.
- We assume that $\mathsf{P}([\lambda_{\mathsf{n}} = \alpha]) = \chi$ and $\mathsf{P}([\lambda_{\mathsf{n}} = \beta]) = 1 - \chi$.

Super relaxations

Example

- Let $\chi \in]0, 1[$, $\alpha \in]0, +\infty[$, and $\beta \in]0, +\infty[$.
- We assume that $\mathsf{P}([\lambda_n = \alpha]) = \chi$ and $\mathsf{P}([\lambda_n = \beta]) = 1 - \chi$.
- Then λ_n satisfies (10) if

$$\chi\big(\alpha(2-\alpha)\big) + (1-\chi)\big(\beta(2-\beta)\big) > 0. \tag{13}$$

Super relaxations

Example

- Let $\chi \in]0, 1[$, $\alpha \in]0, +\infty[$, and $\beta \in]0, +\infty[$.
- We assume that $\mathsf{P}([\lambda_n = \alpha]) = \chi$ and $\mathsf{P}([\lambda_n = \beta]) = 1 - \chi$.
- Then λ_n satisfies (10) if

$$\chi\big(\alpha(2-\alpha)\big) + (1-\chi)\big(\beta(2-\beta)\big) > 0. \tag{13}$$

- This holds for $\chi = 1/7$, $\alpha = 2.5$, and $\beta = 1.8$.
  In such a case, $\mathsf{E}\lambda_n = 1.9$.

Super relaxations

Example

- Let $\alpha \in \,]0, +\infty[$ and $\beta \in \,]\alpha, +\infty[$, and we assume that $\lambda_n \sim \mathbf{uniform}([\alpha, \beta])$.

Super relaxations

Example

- Let $\alpha \in]0, +\infty[$ and $\beta \in]\alpha, +\infty[$, and we assume that $\lambda_n \sim \mathbf{uniform}([\alpha, \beta])$.
- Then λ_n satisfies (10) if

$$3\alpha + 3\beta - (\alpha^2 + \alpha\beta + \beta^2) > 0. \tag{14}$$

Super relaxations

Example

- Let $\alpha \in]0, +\infty[$ and $\beta \in]\alpha, +\infty[$, and we assume that $\lambda_n \sim \mathbf{uniform}([\alpha, \beta])$.
- Then λ_n satisfies (10) if

$$3\alpha + 3\beta - (\alpha^2 + \alpha\beta + \beta^2) > 0. \tag{14}$$

- This condition holds for $\alpha = 1.5$, and $\beta = 2.3$, where $E\lambda_n = 1.9$.

Numerical results

The goal is to recover $\bar{x} \in \mathbb{R}^N$ ($N = 1024$) from 20 noisy observations given by

$$(\forall k \in \{1, \ldots, 20\}) \quad r_k = L_k \bar{x} + w_k \tag{15}$$

where $L_k \colon \mathbb{R}^N \to \mathbb{R}^N$ is a Gaussian convolution filter with zero mean and standard deviation taken uniformly in $[10, 30]$, and

$$w_k \sim \mathrm{uniform}([-0.1, 0.1]^N) \tag{16}$$

is a bounded random noise vector.

Numerical results

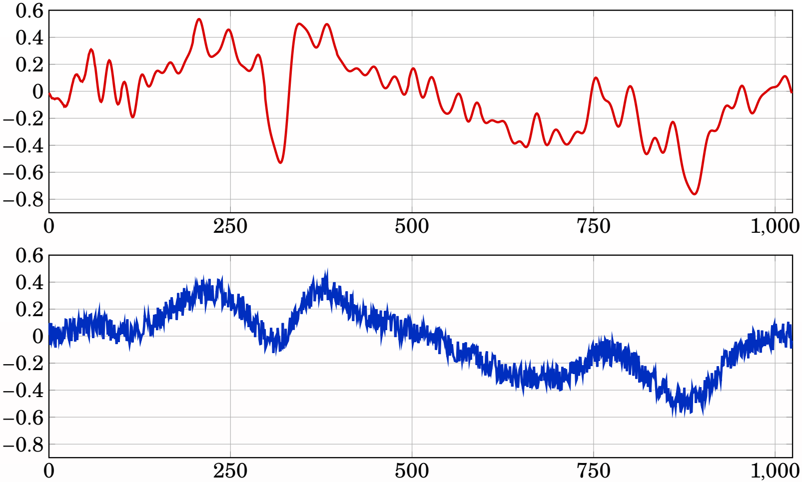

Figure 1: Original signal $\bar{x}$ and noisy observation r_1.

Numerical results

Set, for every $k \in \{1, \ldots, 20\}$ and every $j \in \{1, \ldots, N\}$,

$$C_{k,j} = \left\{x \in \mathbb{R}^N \mid -0.1 \leqslant \langle L_k x - r_k \mid e_j \rangle \leqslant 0.1\right\}. \tag{17}$$

Numerical results

Set, for every $k \in \{1, \ldots, 20\}$ and every $j \in \{1, \ldots, N\}$,

$$C_{k,j} = \left\{x \in \mathbb{R}^N \mid -0.1 \leqslant \langle L_k x - r_k \mid e_j \rangle \leqslant 0.1\right\}. \tag{17}$$

We find a point on the intersection by using

- $K = \{1, \ldots, 20\} \times \{1, \ldots, N\}$.
- $\mathcal{K} = 2^K$.
- $k \sim \mathrm{uniform}(K)$.

Numerical results

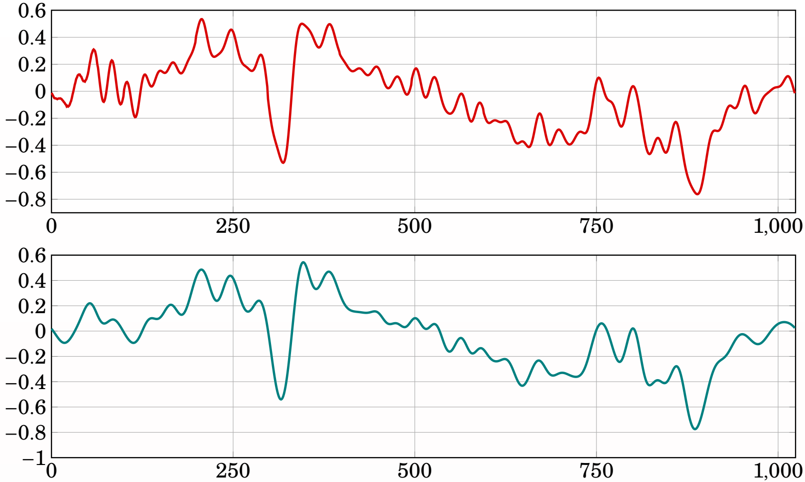

Figure 2: Original signal $\bar{x}$ and solution produced.

Numerical results

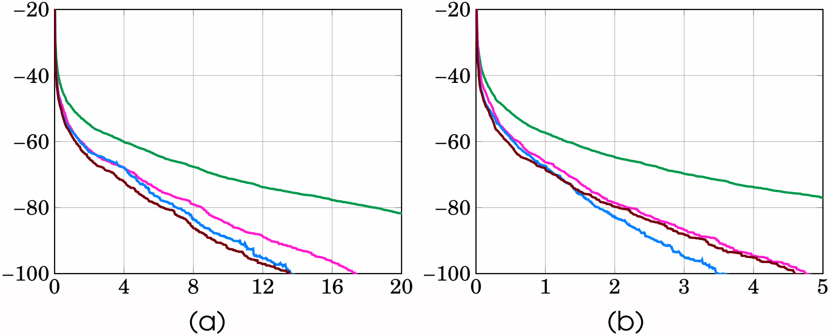

Figure 3: Normalized error $20\log(\|x_n - x_\infty\|/\|x_0 - x_\infty\|)$ (dB) versus execution time (s). Green: $\lambda_n \equiv 1$. Magenta: $\lambda_n \equiv 1.9$. Blue: $P([\lambda_n = 1.5]) = 1/2$ and $P([\lambda_n = 2.3]) = 1/2$. Brown: $\lambda_n \sim \mathrm{uniform}([1.5, 2.3])$. (a): $M = 1$. (b): $M = 128$.

Stochastic Block-Iterative Projections Method for Convex Feasibility

Javier I. Madariaga

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

NLA Student Seminar
March 24, 2025

NC STATE UNIVERSITY