Geometrical scheme for stochastic Fejér monotonicity

Javier I. Madariaga

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

NLA Student Seminar
November 14, 2024

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 $\|\cdot\|$.
- The sets of strong and weak sequential cluster points of a sequence $(x_n)_{n\in\mathbb{N}}$ in H are denoted by $\mathfrak{S}(x_n)_{n\in\mathbb{N}}$ and $\mathfrak{W}(x_n)_{n\in\mathbb{N}}$, respectively.

General problem

Problem 1

Let Z be a nonempty closed convex subset of H.

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 I

For every $n \in \mathbb{N}$, let $T_n : \mathsf{H} \to \mathsf{H}$ be a *firmly nonexpansive operator*, i.e.,

$$(\forall (x, y) \in \mathsf{H} \times \mathsf{H}) \quad \|T_n x - T_n y\|^2 \leqslant \langle T_n x - T_n y \mid x - y \rangle. \tag{1}$$

Example I

For every $n \in \mathbb{N}$, let $T_n : \mathcal{H} \to \mathcal{H}$ be a *firmly nonexpansive operator*, i.e.,

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{H}) \quad \|T_n x - T_n y\|^2 \leqslant \langle T_n x - T_n y \mid x - y \rangle. \tag{1}$$

The common fixed point problem, i.e.,

$$\text{Find } \bar{x} \in \mathcal{H} \text{ such that } (\forall n \in \mathbb{N})\ T_n \bar{x} = \bar{x},$$

is an example of Problem 1.

Example II

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

Example II

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

$$\underset{x \in \mathcal{H}}{\text{minimize}} \ \ f(x).$$

is an example of Problem 1.

Fejér monotonicity

A sequence $(x_n)_{n\in\mathbb{N}}$ of vectors in $\mathcal{H}$ is said to be Fejér monotone with respect to Z if

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|.$$

Fejér monotonicity

A sequence $(x_n)_{n\in\mathbb{N}}$ of vectors in H is said to be Fejér monotone with respect to Z if

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|.$$

Some consequences are:

Fejér monotonicity

A sequence $(x_n)_{n\in\mathbb{N}}$ of vectors in H is said to be Fejér monotone with respect to Z if

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|.$$

Some consequences are:

1. $(x_n)_{n\in\mathbb{N}}$ is bounded.

Fejér monotonicity

A sequence $(x_n)_{n\in\mathbb{N}}$ of vectors in H is said to be Fejér monotone with respect to Z if

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|.$$

Some consequences are:

1. $(x_n)_{n\in\mathbb{N}}$ is bounded.
2. $(\forall z \in Z)$ $\left(\|x_n - z\|\right)_{n\in\mathbb{N}}$ converges.

Fejér monotonicity

A sequence $(x_n)_{n\in\mathbb{N}}$ of vectors in $\mathcal{H}$ is said to be Fejér monotone with respect to Z if

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|.$$

Some consequences are:

1. $(x_n)_{n\in\mathbb{N}}$ is bounded.
2. $(\forall z \in Z)$ $\big(\|x_n - z\|\big)_{n\in\mathbb{N}}$ converges.
3. Suppose that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset Z$.

Fejér monotonicity

A sequence $(x_n)_{n\in\mathbb{N}}$ of vectors in $\mathcal{H}$ is said to be Fejér monotone with respect to Z if

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - z\| \leqslant \|x_n - z\|.$$

Some consequences are:

1. $(x_n)_{n\in\mathbb{N}}$ is bounded.
2. $(\forall z \in Z)$ $\big(\|x_n - z\|\big)_{n\in\mathbb{N}}$ converges.
3. Suppose that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset Z$. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to some vector in Z.

Fejérian scheme

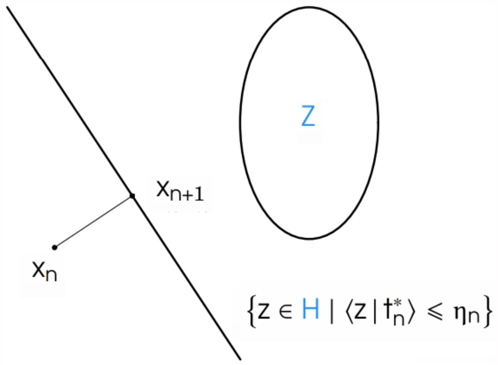
Let $x_0 \in \mathcal{H}$. Iterate

Fejérian scheme

Let $x_0 \in \mathcal{H}$. Iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
t_n^* \in \mathcal{H} \text{ and } \eta_n \in \mathbb{R} \text{ satisfy} \\
\qquad \begin{cases}
\alpha_n = \begin{cases} \dfrac{\langle x_n \mid t_n^* \rangle - \eta_n}{\|t_n^*\|^2}, & \text{if } t_n^* \neq 0 \text{ and } \langle x_n \mid t_n^* \rangle > \eta_n; \\ 0, & \text{otherwise.} \end{cases} \\
(\forall z \in Z) \quad \left\langle z \,\middle|\, \alpha_n t_n^* \right\rangle \leqslant \alpha_n \eta_n
\end{cases} \\
x_{n+1} = x_n - \alpha_n t_n^*.
\end{array}
\right.
\end{array}
$$

Fejérian scheme



Fejérian scheme

Main issue: How to construct/select t_n^* and η_n such that

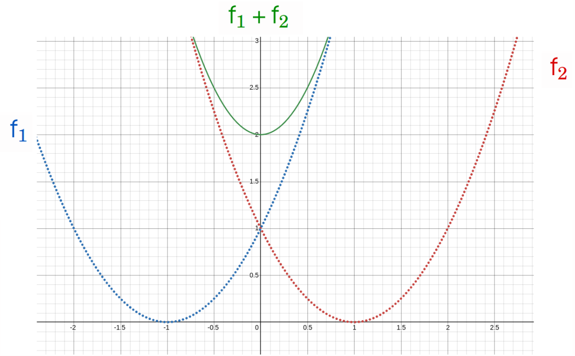
$$\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset Z \quad ?$$

Fejérian scheme

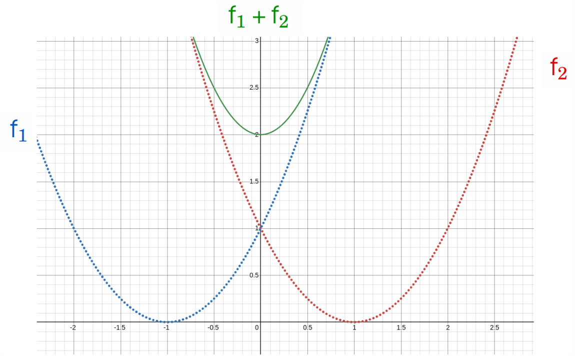
Stochastic algorithms

When transforming those algorithms into stochastic algorithms, **Fejér monotonicity is no longer true**.

Let $f_1 \colon x \mapsto (x+1)^2$ and $f_2 \colon x \mapsto (x-1)^2$. The task is to minimize $f_1 + f_2$ using the stochastic gradient method.



Let $f_1 : x \mapsto (x+1)^2$ and $f_2 : x \mapsto (x-1)^2$. The task is to minimize $f_1 + f_2$ using the stochastic gradient method.



No Fejér monotonicity (in general)!

Background

- The underlying probability space is $(\Omega, \mathcal{F}, \mathsf{P})$. Denote as $\mathcal{B}_{\mathsf{H}}$ the Borel σ-algebra of H. An H-valued random variable is a measurable mapping $x\colon (\Omega, \mathcal{F}) \to (\mathsf{H}, \mathcal{B}_{\mathsf{H}})$.

Background

- The underlying probability space is $(\Omega, \mathcal{F}, \mathsf{P})$. Denote as $\mathcal{B}_{\mathsf{H}}$ the Borel σ-algebra of H. An H-valued random variable is a measurable mapping $x\colon (\Omega, \mathcal{F}) \to (\mathsf{H}, \mathcal{B}_{\mathsf{H}})$.
- Given $\mathsf{p} \in [1, +\infty[$, $L^{\mathsf{p}}(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ denotes the space of H-valued random variables $x\colon (\Omega, \mathcal{F}) \to (\mathsf{H}, \mathcal{B}_{\mathsf{H}})$ such that

$$\mathsf{E}\|x\|^{\mathsf{p}} < +\infty.$$

Background

- The underlying probability space is $(\Omega, \mathcal{F}, \mathsf{P})$. Denote as $\mathcal{B}_{\mathsf{H}}$ the Borel σ-algebra of H. An H-valued random variable is a measurable mapping $x\colon (\Omega, \mathcal{F}) \to (\mathsf{H}, \mathcal{B}_{\mathsf{H}})$.
- Given $\mathsf{p} \in [1, +\infty[$, $L^{\mathsf{p}}(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ denotes the space of H-valued random variables $x\colon (\Omega, \mathcal{F}) \to (\mathsf{H}, \mathcal{B}_{\mathsf{H}})$ such that

$$\mathsf{E}\|x\|^{\mathsf{p}} < +\infty.$$

Further,

$$(\forall \mathsf{A} \in \mathcal{B}_{\mathsf{H}}) \quad L^{\mathsf{p}}(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{A}) = \{x \in L^{\mathsf{p}}(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}) \mid x \in \mathsf{A}\ \mathsf{P}\text{-a.s.}\}$$

Stochastic quasi-Fejér monotonicity

- In the late 60s, Ermol'ev introduce some notion of Random Fejérity.

Stochastic quasi-Fejér monotonicity

- In the late 60s, Ermol'ev introduce some notion of Random Fejérity.
- This was highly ignored until 10 years ago.

Stochastic quasi-Fejér monotonicity

- In the late 60s, Ermol'ev introduce some notion of Random Fejérity.
- This was highly ignored until 10 years ago.
- One of the biggest contribuitor is (Combettes and Pesquet 2015)

Stochastic quasi-Fejér monotonicity

- In the late 60s, Ermol'ev introduce some notion of Random Fejérity.
- This was highly ignored until 10 years ago.
- One of the biggest contribuitor is (Combettes and Pesquet 2015)

A sequence $(x_n)_{n\in\mathbb{N}}$ of H-valued random variables is said to be stochastic quasi-Fejér monotone with respect to Z if there exists a sequence of $[0,+\infty[$-valued random variables $(\varepsilon_n)_{n\in\mathbb{N}}$ such that

$$(\forall z \in Z)(\forall n \in \mathbb{N}) \quad \begin{cases} \mathcal{X}_n = \sigma(x_0,\ldots,x_n) \\ \mathsf{E}\big(\|x_{n+1}-z\| \,\big|\, \mathcal{X}_n\big) \leqslant \|x_n - z\| + 2\varepsilon_n. \end{cases}$$

Stochastic quasi-Fejér monotonicity

No geometrical interpretation of this!

Stochastic quasi-Fejér monotonicity

No geometrical interpretation of this!

We propose a geometrical framework...

Let $x_0 \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$. Iterate

$$
\begin{array}{l}
\text{for } \mathsf{n} = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
\mathfrak{X}_\mathsf{n} = \sigma(x_0, \ldots, x_\mathsf{n}) \\
t_\mathsf{n}^* \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H}),\ \eta_\mathsf{n} \in L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R}),\ \text{and } \varepsilon_\mathsf{n} \in L^1(\Omega, \mathcal{F}, \mathsf{P}; [0, +\infty[) \\
\quad \text{satisfy} \\
\qquad \begin{cases}
1_{[t_\mathsf{n}^* \neq 0]} \eta_\mathsf{n} / (\|t_\mathsf{n}^*\| + 1_{[t_\mathsf{n}^* = 0]}) \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathbb{R}) \\
\alpha_\mathsf{n} = \dfrac{1_{[t_\mathsf{n}^* \neq 0 \text{ and } \langle x_\mathsf{n} | t_\mathsf{n}^* \rangle > \eta_\mathsf{n}]} \big(\langle x_\mathsf{n} \mid t_\mathsf{n}^* \rangle - \eta_\mathsf{n}\big)}{\|t_\mathsf{n}^*\|^2 + 1_{[t_\mathsf{n}^* = 0]}} \\
(\forall \mathsf{z} \in \mathsf{Z}) \ \ \langle \mathsf{z} \mid \mathsf{E}(\alpha_\mathsf{n} t_\mathsf{n}^* \mid \mathfrak{X}_\mathsf{n}) \rangle \leqslant \mathsf{E}(\alpha_\mathsf{n} \eta_\mathsf{n} \mid \mathfrak{X}_\mathsf{n}) + \varepsilon_\mathsf{n} \ \text{P-a.s.}
\end{cases} \\
x_{\mathsf{n}+1} = x_\mathsf{n} - \alpha_\mathsf{n} t_\mathsf{n}^*.
\end{array}
\right.
\end{array}
$$

Some results

1. Let $n \in \mathbb{N}$ and $z \in Z$. Then
   $E(\|x_{n+1} - z\|^2 \mid \mathcal{X}_n) \leqslant \|x_n - z\|^2 + 2\varepsilon_n$ P-a.s.

Some results

1. Let $n \in \mathbb{N}$ and $z \in Z$. Then
   $E(\|x_{n+1} - z\|^2 \mid \mathcal{X}_n) \leqslant \|x_n - z\|^2 + 2\varepsilon_n$ P-a.s.

2. Let $n \in \mathbb{N}$ and $z \in L^2(\Omega, \mathcal{X}_n, P; Z)$. Then
   $E(\|x_{n+1} - z\|^2 \mid \mathcal{X}_n) \leqslant \|x_n - z\|^2 + 2\varepsilon_n$ P-a.s.

Some results

1. Let $n \in \mathbb{N}$ and $z \in Z$. Then
$E(\|x_{n+1} - z\|^2 \mid \mathcal{X}_n) \leqslant \|x_n - z\|^2 + 2\varepsilon_n$ P-a.s.

2. Let $n \in \mathbb{N}$ and $z \in L^2(\Omega, \mathcal{X}_n, P; Z)$. Then
$E(\|x_{n+1} - z\|^2 \mid \mathcal{X}_n) \leqslant \|x_n - z\|^2 + 2\varepsilon_n$ P-a.s.

3. Let $n \in \mathbb{N}$ and $z \in L^2(\Omega, \mathcal{X}_n, P; Z)$. Then
$\|x_{n+1} - z\|^2_{L^2(\Omega,\mathcal{F},P;H)} \leqslant \|x_n - z\|^2_{L^2(\Omega,\mathcal{F},P;H)} + 2E\varepsilon_n$.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \varepsilon_n < +\infty$ P-a.s. Then the following hold:

1. The sequence $(\|x_n\|)_{n\in\mathbb{N}}$ is bounded P-a.s.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \varepsilon_n < +\infty$ P-a.s. Then the following hold:

1. The sequence $(\|x_n\|)_{n\in\mathbb{N}}$ is bounded P-a.s.

2. Let $z \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$. Then $(\|x_n - z\|)_{n\in\mathbb{N}}$ converges P-a.s.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \varepsilon_n < +\infty$ P-a.s. Then the following hold:

1. The sequence $(\|x_n\|)_{n\in\mathbb{N}}$ is bounded P-a.s.
2. Let $z \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$. Then $(\|x_n - z\|)_{n\in\mathbb{N}}$ converges P-a.s.
3. Suppose that $\mathfrak{W}(x_n)_{n\in\mathbb{N}} \subset \mathsf{Z}$ P-a.s. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to a Z-valued random variable.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n < +\infty$. Then the following hold:

Some results

Suppose that $\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n < +\infty$. Then the following hold:

1. The sequence $(\|x_n\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{n\in\mathbb{N}}$ is bounded.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n < +\infty$. Then the following hold:

1. The sequence $(\|x_n\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{n\in\mathbb{N}}$ is bounded.
2. Let $z \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})$. Then $(\mathsf{E}\|x_n - z\|)_{n\in\mathbb{N}}$ converges.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n < +\infty$. Then the following hold:

1. The sequence $(\|x_n\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{n\in\mathbb{N}}$ is bounded.
2. Let $z \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$. Then $(\mathsf{E}\|x_n - z\|)_{n\in\mathbb{N}}$ converges.
3. Let x be an H-valued random variable and suppose that $(x_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to x.

Some results

Suppose that $\sum_{n\in\mathbb{N}} \mathsf{E}\varepsilon_n < +\infty$. Then the following hold:

1. The sequence $(\|x_n\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{n\in\mathbb{N}}$ is bounded.
2. Let $z \in L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{Z})$. Then $(\mathsf{E}\|x_n - z\|)_{n\in\mathbb{N}}$ converges.
3. Let x be an H-valued random variable and suppose that $(x_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to x. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})$ to x.

Some results

Suppose that $\sum_{n \in \mathbb{N}} \mathsf{E}\varepsilon_n < +\infty$. Then the following hold:

1. The sequence $(\|x_n\|_{L^2(\Omega,\mathcal{F},\mathsf{P};\mathsf{H})})_{n\in\mathbb{N}}$ is bounded.
2. Let $z \in L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{Z})$. Then $(\mathsf{E}\|x_n - z\|)_{n\in\mathbb{N}}$ converges.
3. Let x be an H-valued random variable and suppose that $(x_n)_{n\in\mathbb{N}}$ converges weakly P-a.s. to x. Then $(x_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ to x.
4. Let x be an H-valued random variable. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly P-a.s. to x if and only if $(x_n)_{n\in\mathbb{N}}$ converges strongly in $L^1(\Omega, \mathcal{F}, \mathsf{P}; \mathsf{H})$ to x.

Geometrical scheme for stochastic Fejér monotonicity

Javier I. Madariaga

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

NLA Student Seminar
November 14, 2024

NC STATE UNIVERSITY