

Introductory Course on Non-smooth Optimisation

Lecture 07 - Other operator splitting methods

Jingwei Liang

Department of Applied Mathematics and Theoretical Physics

- 1 Three-operator splitting
- 2 Forward-Douglas-Rachford splitting
- 3 Generalised Forward-Backward splitting

Problem

Find $x \in \mathbb{R}^n$ such that $0 \in A(x) + B(x) + C(x)$.

Assumptions

- $A, B : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ are maximal monotone.
- $C : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is β -cocoercive.
- $\text{zer}(A + B + C) \neq \emptyset$.

- given $x^* \in \text{zer}(A + B + C)$, there exists $z^* \in \mathbb{R}^n$ such that

$$\begin{cases} x^* - z^* \in \gamma A(x^*) + \gamma C(x^*) \\ z^* - x^* \in \gamma B(x^*) \end{cases} \implies \begin{cases} 2x^* - z^* - \gamma C(x^*) \in x^* + \gamma A(x^*), \\ z^* \in x^* + \gamma B(x^*). \end{cases}$$

- apply the resolvent

$$\begin{cases} x^* = \mathcal{J}_{\gamma A}(2x^* - z^* - \gamma C(x^*)), \\ x^* = \mathcal{J}_{\gamma B}(z^*). \end{cases}$$

- equivalent formulation

$$\begin{cases} z^* = z^* + \mathcal{J}_{\gamma A}(2x^* - z^* - \gamma C(x^*)) - x^*, \\ x^* = \mathcal{J}_{\gamma B}(z^*). \end{cases}$$

- fixed-point iteration

$$\begin{cases} z_{k+1} = z_k + (\mathcal{J}_{\gamma A}(2x_k - z_k - \gamma C(x_k)) - x_k), \\ x_{k+1} = \mathcal{J}_{\gamma B}(z_{k+1}). \end{cases}$$

Three-operator splitting

Let $z_0 \in \mathbb{R}^n$, $\gamma \in]0, 2\beta[$ and $x_0 = \mathcal{J}_{\gamma B}(z_0)$, $\lambda \in]0, \frac{4\beta - \gamma}{2\beta}[$:

$$u_{k+1} = \mathcal{J}_{\gamma A}(2x_k - z_k - \gamma C(x_k)),$$

$$z_{k+1} = (1 - \lambda)z_k + \lambda(z_k + u_{k+1} - x_k),$$

$$x_{k+1} = \mathcal{J}_{\gamma B}(z_{k+1}).$$

- Recovers Douglas–Rachford when $C = 0$.
- Recovers Forward–Backward when $B = 0$.

Fixed-point formulation

- $u_{k+1} = \mathcal{J}_{\gamma A}(2x_k - z_k - \gamma C(x_k)) = \mathcal{J}_{\gamma A} \circ (2\mathcal{J}_{\gamma B} - \text{Id} - \gamma C \circ \mathcal{J}_{\gamma B})(z_k).$
- For z_k ,

$$\begin{aligned} z_{k+1} &= (1 - \lambda)z_k + \lambda(z_k + u_{k+1} - x_k) \\ &= (1 - \lambda)z_k + \lambda(\text{Id} - \mathcal{J}_{\gamma B} + \mathcal{J}_{\gamma A} \circ (2\mathcal{J}_{\gamma B} - \text{Id} - \gamma C \circ \mathcal{J}_{\gamma B}))(z_k). \end{aligned}$$

Property

- $\mathcal{T}_{\text{TOS}} \stackrel{\text{def}}{=} \text{Id} - \mathcal{J}_{\gamma B} + \mathcal{J}_{\gamma A} \circ (2\mathcal{J}_{\gamma B} - \text{Id} - \gamma C \circ \mathcal{J}_{\gamma B})$ is $\frac{2\beta}{4\beta - \gamma}$ -averaged.

- 1 Three-operator splitting
- 2 Forward–Douglas–Rachford splitting**
- 3 Generalised Forward–Backward splitting

Problem

Find $x \in \mathbb{R}^n$ such that $0 \in A(x) + \mathcal{N}_V(x) + C(x)$.

Assumptions

- $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ are maximal monotone.
- $V \subseteq \mathbb{R}^n$ is a closed subspace.
- $C : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is β -cocoercive.
- $\text{zer}(A + \mathcal{N}_V + C) \neq \emptyset$.

Forward–Douglas–Rachford splitting

Let $z_0 \in \mathbb{R}^n$, $\gamma \in]0, 2\beta[$ and $x_0 = \mathcal{J}_{\gamma B}(z_0)$, $\lambda \in]0, \frac{4\beta - \gamma}{2\beta}[$:

$$u_{k+1} = \mathcal{J}_{\gamma A}(2x_k - z_k - \gamma \mathcal{P}_V \circ C(x_k)),$$

$$z_{k+1} = (1 - \lambda)z_k + \lambda(z_k + u_{k+1} - x_k),$$

$$x_{k+1} = \mathcal{P}_V(z_{k+1}).$$

- FDR was proposed before TOS.
- Recovers Douglas–Rachford when $C = 0$.
- Recovers Forward–Backward when $V = \mathbb{R}^n$.

Fixed-point formulation Denote $C_V = \mathcal{P}_V \circ C \circ \mathcal{P}_V$,

- For u_{k+1} : $\mathcal{R}_V \circ C_V = (2\mathcal{P}_V - \text{Id})C_V = C_V$

$$\begin{aligned}u_{k+1} &= \mathcal{J}_{\gamma A} \circ (2\mathcal{P}_V - \text{Id} - \gamma C_V)(z_k) \\ &= \mathcal{J}_{\gamma A} \circ \mathcal{R}_V \circ (\text{Id} - \gamma C_V)(z_k).\end{aligned}$$

- For z_k ,

$$\begin{aligned}z_{k+1} &= (1 - \lambda)z_k + \lambda(z_k + u_{k+1} - x_k) \\ &= (1 - \lambda)z_k + \lambda(\mathcal{J}_{\gamma A} \circ \mathcal{R}_V \circ (\text{Id} - \gamma C_V) + \text{Id} - \mathcal{P}_V)(z_k) \\ &= (1 - \lambda)z_k + \lambda \frac{1}{2}(\text{Id} + \mathcal{R}_{\gamma R} \mathcal{R}_V)(\text{Id} - \gamma C_V)(z_k).\end{aligned}$$

Identify: $\text{Id} - \mathcal{P}_V = \frac{1}{2}(\text{Id} - \gamma C_V) - \mathcal{P}_V(\text{Id} - \gamma C_V) + \frac{1}{2}(\text{Id} - \gamma C_V)$.

Property

- $\mathcal{T}_{\text{FDR}} \stackrel{\text{def}}{=} \frac{1}{2}(\text{Id} + \mathcal{R}_{\gamma R} \mathcal{R}_V)(\text{Id} - \gamma C_V)$ is $\frac{2\beta}{4\beta - \gamma}$ -averaged.

- 1 Three-operator splitting
- 2 Forward–Douglas–Rachford splitting
- 3 Generalised Forward–Backward splitting**

Problem $r \geq 2$

Find $x \in \mathbb{R}^n$ such that $0 \in \sum_{i=1}^r A_i(x) + B(x)$.

Assumptions

- for each i , $A_i : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximal monotone.
- $B : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is β -cocoercive.
- $\text{zer}(\sum_i A + B) \neq \emptyset$.

Generalised Forward-Backward splitting

Let $(\omega_i)_i \in]0, 1[^r$ s.t. $\sum_i \omega_i = 1$, $\gamma \in]0, 2\beta[$, $\lambda \in]0, \frac{4\beta - \gamma}{2\beta}[$. $z_{i,0} \in \mathbb{R}^n$ and $x_0 = \sum_i \omega_i z_{i,0}$:

For $i \in \{1, \dots, r\}$

$$\begin{cases} u_{i,k+1} = \mathcal{J}_{\frac{\gamma}{\omega_i} A_i}(2x_k - z_{i,k} - \gamma B(x_k)) \\ z_{i,k+1} = (1 - \lambda)z_{i,k} + \lambda(z_{i,k} + u_{i,k+1} - x_k), \end{cases}$$
$$x_{k+1} = \sum_i \omega_i z_{i,k+1}.$$

- Earliest of the three methods.
- Recovers Douglas-Rachford in product space when $B = 0$.
- Recovers Forward-Backward when $r = 1$.

- Let $\mathcal{H} = \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ be the product space endowed with the scalar product and norm defined by

$$\forall \mathbf{x}, \mathbf{x}' \in \mathcal{H}, \langle \mathbf{x}, \mathbf{x}' \rangle = \sum_{i=1}^r \omega_i \langle x_i, x'_i \rangle, \quad \|\mathbf{x}\| = \sqrt{\sum_{i=1}^r \omega_i \|x_i\|^2}.$$

- Let $\mathcal{S} = \{\mathbf{x} = (x_i)_i \in \mathcal{H} \mid x_1 = \cdots = x_r\}$ and $\mathcal{S}^\perp = \{\mathbf{x} = (x_i)_i \in \mathcal{H} \mid \sum_{i=1}^r \omega_i x_i = 0\}$. Define the canonical isometry $\mathbf{C} : \mathcal{H} \rightarrow \mathcal{S}$, $\mathbf{x} \mapsto (x, \cdots, x)$, then

$$\mathcal{P}_{\mathcal{S}}(\mathbf{z}) \stackrel{\text{def}}{=} \mathbf{C}(\sum_{i=1}^r \omega_i z_i), \quad \forall \mathbf{z} \in \mathcal{H}.$$

- Let $\gamma = (\gamma_i)_i \in]0, +\infty[^r$. For $A_i, i = 1, \dots, r$, define

$$\gamma \mathbf{A} : \mathcal{H} \rightrightarrows \mathcal{H}, \mathbf{x} = (x_i)_i \mapsto \bigtimes_{i=1}^r \gamma_i A_i(x_i).$$

For B , define

$$\mathbf{B} : \mathcal{H} \rightarrow \mathcal{H}, \mathbf{x} = (x_i)_i \mapsto (B(x_i))_i.$$

- Define $\mathbf{B}_{\mathcal{S}} = \mathbf{B} \circ \mathcal{P}_{\mathcal{S}}$ and $\mathcal{J}_{\gamma \mathbf{A}} = (\mathcal{J}_{\gamma_i A_i})_i$.

Fixed-point formulation

- For \mathbf{u}_{k+1} ,

$$\begin{aligned}\mathbf{u}_{k+1} &= \mathcal{J}_{\gamma\mathbf{A}}(2\mathcal{P}_{\mathcal{S}}(\mathbf{z}_k) - \mathbf{z}_k - \gamma\mathbf{B}_{\mathcal{S}}(\mathbf{z}_k)) \\ &= \mathcal{J}_{\gamma\mathbf{A}} \circ \mathcal{R}_{\mathcal{S}} \circ (\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}})(\mathbf{z}_k).\end{aligned}$$

- Identify: $\mathbf{Id} - \mathcal{P}_{\mathcal{S}} = \frac{1}{2}(\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}}) - \mathcal{P}_{\mathcal{S}} \circ (\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}}) + \frac{1}{2}(\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}})$.

- For \mathbf{z}_k ,

$$\begin{aligned}\mathbf{z}_{k+1} &= (1 - \lambda)\mathbf{z}_k + (\mathbf{z}_k + \mathcal{J}_{\gamma\mathbf{A}}(2\mathcal{P}_{\mathcal{S}}(\mathbf{z}_k) - \mathbf{z}_k - \gamma\mathbf{B}_{\mathcal{S}}(\mathbf{z}_k)) - \mathcal{P}_{\mathcal{S}}(\mathbf{z}_k)) \\ &= (1 - \lambda)\mathbf{z}_k + (\mathcal{J}_{\gamma\mathbf{A}} \circ \mathcal{R}_{\mathcal{S}} \circ (\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}})(\mathbf{z}_k) + (\mathbf{Id} - \mathcal{P}_{\mathcal{S}})(\mathbf{z}_k)) \\ &= (1 - \lambda)\mathbf{z}_k + \lambda\frac{1}{2}(\mathbf{Id} + \mathcal{R}_{\gamma\mathbf{A}}\mathcal{R}_{\mathcal{S}}) \circ (\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}})(\mathbf{z}_k).\end{aligned}$$

Property

- $\mathcal{T}_{\text{GFB}} \stackrel{\text{def}}{=} \frac{1}{2}(\mathbf{Id} + \mathcal{R}_{\gamma\mathbf{A}}\mathcal{R}_{\mathcal{S}}) \circ (\mathbf{Id} - \gamma\mathbf{B}_{\mathcal{S}})$ is $\frac{2\beta}{4\beta - \gamma}$ -averaged.

- Structure and splitting are the key to design first-order methods.
- Convergence analysis via Krasnosel'skiĭ-Mann iteration.
- Most common structure for Krasnosel'skiĭ-Mann operator: PPA and FB.
- Acceleration in general is difficult.

- D. Davis and W. Yin. “A three-operator splitting scheme and its optimization applications”. Set-Valued and Variational Analysis, 2017.
- L. M. Briceño-Arias. “Forward-douglas-rachford splitting and forward-partial inverse method for solving monotone inclusions”. Optimization, 64(5):1239–1261, 2015.
- H. Raguet, M. J. Fadili, and G. Peyré. “Generalized forward-backward splitting”. SIAM Journal on Imaging Sciences, 6(3):1199–1226, 2013.