Introductory Course on Non-smooth Optimisation

Lecture 07

Other operator splitting methods

Outline

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

Outline

Sum of three operators

Problem

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

Assumptions

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

Solution characterisation

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

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

- apply the resolvent

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

- equivalent formulation

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

- iteration

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

Three-operator splitting

Three-operator splitting

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

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

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

Fixed-point characterisartion

Fixed-point formulation

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

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

Property

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

Outline

Subspace constrained monotone inclusion

Problem

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

Assumptions

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

Forward–Douglas–Rachford splitting

Forward–Douglas–Rachford splitting

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

$$\begin{aligned} u_{k+1} &= \mathcal{J}_{\gamma A}\big(2x_k - z_k - \gamma \mathcal{P}_V \circ C(x_k)\big) \\ z_{k+1} &= (1-\lambda)z_k + \lambda(z_k + u_{k+1} - x_k) \\ x_{k+1} &= \mathcal{P}_V(z_{k+1}) \end{aligned}$$

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

Fixed-point characterisartion

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

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

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

- For z_k,

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

Hint: $\mathrm{Id} - \mathcal{P}_V = \frac{1}{2}(\mathrm{Id} - \gamma C_V) - \mathcal{P}_V(\mathrm{Id} - \gamma C_V) + \frac{1}{2}(\mathrm{Id} - \gamma C_V)$

Property

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

Outline

A general monotone inclusion

Problem $r \geq 2$

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

Assumptions

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

Generalised Forward–Backward splitting

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 $x_0 = \sum_i \omega_i z_{i,0}$:

$$\begin{aligned}
&\text{For } i \in \{1, \cdots, r\} \\
&\left\lfloor \begin{aligned} u_{i,k+1} &= \mathcal{J}_{\frac{\gamma}{\omega_i} A_i}(2x_k - z_{i,k} - \gamma B(x_k)) \\ z_{i,k+1} &= (1-\lambda) z_{i,k} + \lambda\big(z_{i,k} + u_{i,k+1} - x_k\big), \end{aligned} \right. \\
&x_{k+1} = \textstyle\sum_i \omega_i z_{i,k+1}.
\end{aligned}$$

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

Product space

- 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

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

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

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

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

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

  For B, define

$$\boldsymbol{B} : \mathcal{H} \to \mathcal{H},\ \boldsymbol{x} = (x_i)_i \mapsto (B(x_i))_i$$

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

Fixed-point characterisartion

Fixed-point formulation

- For $\boldsymbol{u}_{k+1}$,

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

- Identify: $\mathbf{Id} - \mathcal{P}_{\boldsymbol{\mathcal{S}}} = \frac{1}{2}(\mathbf{Id} - \gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}}) - \mathcal{P}_{\boldsymbol{\mathcal{S}}}(\mathbf{Id} - \gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}}) + \frac{1}{2}(\mathrm{Id} - \gamma\boldsymbol{B}_{\boldsymbol{\mathcal{S}}})$
- For z_k,

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

Property

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

Remarks

- 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 difficult

Reference

- 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.